UNIVERSIDADE DE SÃO PAULO INSTITUTO DE QUÍMICA

Programa de Pós-Graduação em Ciências Biológicas (Bioquímica)

JOÃO BATISTA PLACIDO DO NASCIMENTO

Identificação de RNAs não codificadores expressos no epitélio olfatório

Versão corrigida da Tese defendida

São Paulo Data do Depósito na SPG: 01/02/2018

JOÃO BATISTA PLACIDO DO NASCIMENTO

Identificação de RNAs não codificadores expressos no epitélio olfatório

Tese apresentada ao Instituto de Química da Universidade de São Paulo para obtenção do Título de Doutor em Ciências, área Bioquímica

Orientadora: Profa. Dra. Bettina Malnic

São Paulo 2018 Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletronico, 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

Nascimento, João Batista Placido do Identificação de RNAs não codificadores expressos no epitélio olfatório / João Batista Placido do Nascimento. - São Paulo, 2018. 103 p.
Tese (doutorado) - Instituto de Química da Universidade de São Paulo. Departamento de Bioquímica. Orientador: Malnic, Bettina
1. Epitélio olfatório. 2. RNAs não codificadores (ncRNAs). 3. micro RNAs (miRNAs). 4. Long Intergenic Non-coding RNAs (lincRNAs). 5. RNA-seq. I. T. II. Malnic, Bettina, orientador.



JOÃO BATISTA PLACIDO DO NASCIMENTO

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:

Profa, Dra. Bettina Malnic (Orientadora e Presidente)

Prof. Dr. Sergio Verjovski de Almeida IQ - USP

Profa. Dra. Aline Maria da Silva IQ - USP

Profa. Dra. Chae Yun Irene Yan ICB - USP

> Prof. Dr. Fabio Papes IB - UNICAMP

SÃO PAULO 15 de maio de 2018

Dedico este trabalho ao meu falecido pai José Placido do Nascimento.

AGRADECIMENTOS

Agradeço à Profa. Dra. Bettina Malnic por aceitar me orientar e por todo o suporte e orientação que me deu durante todo o período de meu doutorado.

Ao meu colaborador Pedro Galante.

A todos os professores que pude ter contato em disciplinas e seminários, dentro e fora do Instituto de Química.

À Dra. Layla Farage, responsável pelo Centro Avançado de Tecnologias em Genômica (CATG).

Aos técnicos dos laboratórios em que frequentei e principalmente de meu laboratório, Erika e Edson, por todo serviço prestado.

Aos colegas de laboratório Daniela, Maíra, Lúcia, Cleiton, Artur pela amizade e ajuda em alguns experimentos e companhia durante o Doutorado.

Aos funcionários do IQUSP por proporcionarem um ambiente agradável de trabalho e aos serviços prestados.

Aos meus amigos de fora da USP, por todo o apoio e interesse durante o período de meu doutorado.

À minha esposa Ana Paula, por seu companheirismo, paciência e por todo apoio, carinho e amor em todos os momentos dessa minha longa trajetória.

Ao CNPq, CAPES e FAPESP pelo apoio financeiro e à FAPESP pela bolsa de Doutorado.

"Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes".

Marthin Luther King

RESUMO

Nascimento, João Batista Placido do. Identificação de RNAs não codificadores expressos no epitélio olfatório. 2018. 103 páginas. Tese de Doutorado - Programa de Pós-Graduação em Ciências Biológicas (Bioquímica). Instituto de Química, Universidade de São Paulo, São Paulo.

Odorantes são detectados por centenas de receptores olfatórios (ORs) que pertencem à superfamília dos receptores acoplados à proteína G. Estes receptores são expressos nos neurônios sensoriais olfatórios localizados na cavidade nasal. Cada neurônio sensorial olfatório expressa um único alelo de gene OR de uma grande família de genes OR. Este padrão característico da expressão de genes OR resulta na formação de um mapa olfatório espacial no bulbo olfatório, que é necessário para a discriminação de odorantes pelo sistema olfatório. Os mecanismos envolvidos nesta regulação ainda não são bem conhecidos. O DNA genômico em neurônios olfatórios é coberto com marcas repressivas de metilação de histonas, indicando que a regulação da estrutura da cromatina deve desempenhar um papel importante na regulação da expressão de genes OR. Trabalhos anteriores demonstraram que RNAs não codificadores (ncRNAs) estão envolvidos na deposição de marcas de histonas em determinados genes. No entanto, os ncRNAs expressos no epitélio olfatório ainda não são conhecidos. Neste trabalho, identificamos e catalogamos o repertório completo de ncRNAs anotados, incluindo os miRNAs, expressos no epitélio olfatório de camundongos recémnascidos e adultos. Muitos destes, apesar de já anotados como ncRNAs, ainda não foram descritos na literatura como expressos no MOE. Identificamos ao todo 1161 miRNAs e 295 lincRNAs expressos no epitélio olfatório, e pudemos verificar como os níveis de expressão destes RNAs variam durante o desenvolvimento. A partir deste repertório, selecionamos lincRNAs que são preferencialmente expressos no epitélio olfatório quando comparados a outros tecidos de camundongo. Dez destes lincRNAs foram selecionados para validação utilizando-se RT-PCR. Cinco lincRNAs foram validados e analisados quanto à sua expressão em diferentes tecidos. Nosso trabalho estabelece uma plataforma de dados que permitirá o estudo do papel desempenhado por ncRNAs no epitélio olfatório. Além disto, os nossos resultados mostram que a abordagem utilizada permite a identificação de novos lincRNAs que apresentam expressão restrita ou preferencial no epitélio olfatório, e que, portanto, devem apresentar uma função relevante para o olfato.

Palavras-chave: Epitélio Olfatório, Receptores olfatórios, ncRNAs, miRNAs, lincRNAs, sequenciamento de DNA em larga escala, expressão gênica.

ABSTRACT

Nascimento, João Batista Placido do. Identification of noncoding RNAs expressed in the olfactory epithelium. 2018. 103 pages. PhD Thesis - Graduate Program in Biochemistry. Instituto de Química, Universidade de São Paulo, Sao Paulo.

Odorants are detected by hundreds of odorant receptors (ORs) which belong to the superfamily of G protein-coupled receptors. These receptors are expressed in the olfactory sensory neurons of the nose. Each olfactory sensory neuron expresses one single OR gene allele from a large family of OR genes. This characteristic pattern of OR gene expression results in the formation of a spatial olfactory map in the olfactory bulb, which is required for odorant discrimination by the olfactory system. The mechanisms involved in this regulation are unknown. OR genomic DNA in olfactory neurons is covered with repressive histone methylation marks, indicating that the chromatin structure should play an important role in the regulation of OR gene expression. Previous studies suggest that noncoding RNAs (ncRNAs) are involved in the deposition of histone marks in certain genes. However, the ncRNAs expressed in the olfactory epithelium are completely unknown. In this work, we used RNA-seq to identify and catalogue the complete repertoire of ncRNAs, including miRNAs, expressed in the olfactory epithelium from newborn and adult mice. In this way, we were able to identify 1161 miRNAs and 295 lincRNAs and analyze how their levels of expression varies during development. Out of these repertoire, we selected lincRNAs that are preferentially expressed in the olfactory epithelium when compared to other mouse tissues. Ten out of these lincRNAs were selected for validation by using RT-PCR, and five of them could be validated and further analyzed. Our work establishes a data platform which will enable the study of the role played by ncRNAs in the olfactory epithelium. In addition, our results show that our approach can be successfully used to identify ncRNAs that are restrictedly or preferentially expressed in the olfactory epithelium, and which therefore must be relevant for olfaction.

Keywords: Olfactory epithelium, odorant receptors, ncRNAs, miRNAs, lincRNAs, RNA-seq.

LISTA DE ABREVIATURAS E SIGLAS

CNGC: Cyclic-nucleotide gated channel CPM: Counts per million DAPI: 4 ',6-diamino-2-fenilindol Ezh2: Enhancer of zeste homolog 2 FPKM: Fragments Per Kilobase of exon model per Million mapped reads Gaolf: Guanine nucleotide-binding proteins /olfactory type KDM6A: Lysine Demethylase 6A KDM6B: Lysine Demethylase 6B lincRNA: Long Intergenic Non-coding RNAs miRNA: microRNA MOE: Sistema olfatório principal mRNA: RNA mensageiro NC: Cavidade nasal ncRNA: RNA não codificador OMP: Olfactory Marker Protein **OR:** Odorant receptors piRNA: Piwi-interacting RNA PRC2: Polycomb repressive complex 2 **RIN:** RNA Integrity Number rRNA: RNA ribossomal snoRNA: Small nucleolar RNAs snRNA: Small nuclear RNA SOLiD: Sequencing by Oligonucleotide Ligation and Detection tRNA: RNA transportador V1Rs: Type-1 vomeronasal receptors V2Rs: Type-2 vomeronasal receptors VNO: Órgão vomeronasal Xist: X-inactivation specific transcript

SUMÁRIO

1	INT	ITRODUÇÃO1				
	1.1	O sistema olfatório	12			
	1.2	Regulação da expressão dos genes OR	17			
	1.3	RNAs não codificadores (ncRNAs)	21			
	1.4	Importância dos ncRNAs na regulação gênica	22			
	1.5	RNAs não codificadores presentes no sistema olfatório	25			
2	OB.	JETIVOS	28			
	2.1	Objetivo geral:	28			
	2.1.	1 Objetivos específicos:	28			
3	MA	TERIAL E MÉTODOS	29			
	3.1	Extração de RNA	29			
	3.2	Produção de cDNA	29			
	3.3	Construção das bibliotecas de cDNA e sequenciamento por ligação	30			
	3.4	Programas utilizados nas análises de bioinformática	31			
	3.4.	1 Cuffdiff	31			
3.4 3.4 3.4		2 DESeq	32			
		3 EdgeR	32			
		4 TargetScan	32			
	3.5	RT-PCR	33			
	3.6	Clonagem dos moldes para as sondas	35			
	3.7	Sequenciamento automatizado de DNA	35			
	3.8	Síntese das sondas de RNA sense e anti-sense marcadas com				
digoxigenina						

	3.9	Preparação das lâminas com cortes de camundongos C57BL/6						
	3.10	Hib	ridização <i>in situ</i>	36				
4	RES	SULT	TADOS E DISCUSSÃO	38				
	4.1	Dis	secção do epitélio olfatório	38				
	4.2	Qua	Quantificação do RNA total39					
	4.3	Aná	Análises do sequenciamento41					
	4.4	Aná	álise dos RNAs longos	44				
	4.5	Aná	álise dos miRNAs	47				
	4.6	Pre	dição dos alvos para os miRNAs	51				
	4.7	Loc	alização da expressão de Ezh2 no epitélio olfatório	54				
	4.7.	1	Hibridização <i>in situ</i>	54				
	4.8	Aná	alise dos RNAs não codificadores longos	56				
	4.8.	1	Análise de ncRNAs que mapeiam na região de genes OR	58				
	4.8.	2	Validação dos RNAs não codificadores longos	59				
	4.8.	3	Amplificação dos lincRNAs	62				
	4.8.	4	Sequenciamento dos lincRNAs	65				
	4.8.	5	Distribuição tecidual da expressão dos lincRNAs	71				
	4.8.	6	Localização da expressão dos lincRNAs no epitélio olfatório	74				
	4.9	Exp	pressão dos lincRNAs no VNO	82				
5	DIS	CUS	SÃO	84				
6	CO	CONCLUSÕES						
7	REF	REFERÊNCIAS						
8	ANE	EXO	S	98				
	exo A - SÚMULA CURRICULAR	98						

8.2	Anexo B - Artigo em preparação, e submetio	do para a disciplina QBQ57499
(Red	ação Científica)	
8.3	Arquivos suplementares	

1 INTRODUÇÃO

1.1 O sistema olfatório

O início da olfação ocorre no epitélio olfatório, localizado no recesso posterior da cavidade nasal. O epitélio olfatório é composto de três tipos celulares principais, neurônios sensoriais olfatórios, células de suporte e células basais, que dão origem a novos neurônios olfatórios. Um odorante disperso no ar entra em contato com receptores presentes nos neurônios sensoriais olfatórios levando à despolarização desse neurônio. O axônio do neurônio olfatório faz sinapse com neurônios do bulbo olfatório (denominados de células mitrais), que por sua vez projetam para regiões corticais e sub-corticais, onde ocorrerá o processamento dessa informação pelo cérebro (Figura 1).

O início da detecção dos odorantes ocorre nos cílios dos neurônios olfatórios maduros através da ligação do odorante a uma família de receptores acoplados à proteína G, identificada e denominada como a família dos receptores olfatórios (ORs) (BUCK, L.; AXEL, 1991). Estudos sugerem a existência de aproximadamente 1.000 genes que codificam para diferentes ORs em camundongos (GODFREY; MALNIC; BUCK, 2004) e cerca de 350 genes no homem (MALNIC; GODFREY; BUCK, 2004), correspondendo de 1 a 3 % do genoma e formando assim a maior família de genes identificada até o momento no genoma de mamíferos.



Figura 1: Organização esquemática do sistema olfatório de humanos. Adaptado de Karolinska Institutet and Nobel Foundation, Stockholm, Sweden, 2004. Os axônios dos neurônios olfatórios do epitélio olfatório atravessam a placa cribriforme e fazem sinapses com os dendritos das células mitrais em estruturas denominadas de glomérulos. Os axônios das células mitrais enviam a informação para outras regiões do cérebro, desencadeando a percepção dos odorantes.

Os ORs são bastante diversificados em relação às suas sequências de aminoácidos (apresentam identidade de aminoácidos entre 40%-80%) o que possibilita o reconhecimento e a discriminação de uma grande variedade de odorantes. Regiões de grande variabilidade são encontradas nos domínios transmembrânicos III, IV e V da estrutura dos ORs, o que sugere serem estas as regiões importantes para a ligação com os odorantes (BUCK, L.; AXEL, 1991).

Estudos demonstraram que um mesmo OR pode ser ativado por diferentes odorantes e que um mesmo odorante pode interagir com diferentes ORs (MALNIC et al., 1999). Tais resultados demonstraram que os odorantes ativam certo número de ORs específicos sendo que este conjunto de receptores ativados possibilitará tal identificação, formando o início do "código" para um dado odorante. A possibilidade de combinação de diferentes ORs para a detecção de um dado odorante aumenta muito a capacidade de discriminação do sistema olfatório. Por exemplo, se considerarmos um repertório de 300 ORs e que em média um odorante ative uma combinação de 10 ORs, teríamos mais de 10¹⁴ combinações diferentes de receptores (FIRESTEIN, 2004). Por fim, ORs ativados por um dado odorante irão promover a geração de potenciais de ação nos neurônios olfatórios maduros que se propagarão até o bulbo olfatório onde ocorrerá o estímulo de um conjunto específico de glomérulos, sendo este o segundo passo para a formação do código para um dado odorante.

Este código é então transmitido do bulbo para o córtex olfatório onde será distribuído para outras áreas do cérebro podendo alcançar regiões superiores envolvidas com consciência, emoção e motivação (BUCK, L. B., 1996).

Como o sistema olfatório é capaz de organizar tal código proveniente da combinação de alguns membros do repertório de 1.000 ORs? As primeiras respostas para entender tal organização foram obtidas através de ensaios de hibridização *in situ* em epitélio olfatório de camundongo e de rato. Estes estudos indicaram que cada um dos 1.000 ORs é expresso em apenas 0,1% dos neurônios olfatórios, sugerindo que cada neurônio olfatório deveria expressar um único OR

(RESSLER; SULLIVAN; BUCK, 1993; VASSAR; NGAI; AXEL, 1993). Essa hipótese foi posteriormente confirmada pela técnica de RT-PCR utilizando cDNA proveniente de neurônios olfatórios únicos (MALNIC et al., 1999).

Estudos utilizando camundongos transgênicos demonstraram que os axônios de todos os neurônios olfatórios que expressam um dado OR convergem para um mesmo ponto no bulbo olfatório (MOMBAERTS et al., 1996; WANG, F. et al., 1998). Esse ponto é o glomérulo, onde os axônios dos neurônios olfatórios fazem sinapse com os dendritos das células mitrais do bulbo, estas por sua vez, projetam seus axônios para várias regiões do cérebro (FIRESTEIN, 2001). Dessa forma, o bulbo olfatório possui um mapa organizado, representado pelo conjunto de aproximados 2.000 glomérulos, que mantém o código proveniente de um conjunto de ORs ativados por um dado odorante. Foi também demonstrado que os ORs não estão apenas envolvidos na ligação a odorantes, mas também no direcionamento dos axônios para glomérulos específicos no bulbo olfatório (MOMBAERTS et al., 1996; WANG, F. et al., 1998).

Em mamíferos, os neurônios sensoriais localizados no MOE detectam principalmente odorantes voláteis, enquanto os que estão no sistema olfatório VNO (órgão vomeronasal), é o principal responsável pela detectção de feromônios (MATSUNAMI; BUCK, 1997). No entanto, os neurônios sensoriais presentes no VNO são tipicamente segregados (Figura 2) daqueles que estão presente no MOE (HILDEBRAND, 1997; WYSOCKI C. J, 1987).



Figura 2: Segregação espacial do órgão vomeronasal (VNO) e do sistema olfatório principal (MOE). (A) Esquema de uma seção parasagital do crânio de um camundongo. O sistema olfatório principal (MOE) reside dentro do recesso posterior da cavidade nasal (NC), enquanto o VNO reside mais anteriormente em uma bolsa dentro do septo nasal. (B) Esquema de uma secção coronal mostrando a anatomia distinta do VNO, MOE, NC e palato (P). (Figura retirada de Dulac and Axel, 1995).

A organização molecular do neuroepitélio apical e basal do VNO em roedores é constituída por um padrão de expressão específico de duas proteínas G (Gαi2 e Gαo) e duas distintas superfamílias de receptores de feromônios, os *type-1 vomeronasal receptors* (V1Rs) (DULAC; AXEL, 1995) e os *type-2 vomeronasal receptors* (V2Rs) (HERRADA; DULAC; MATSUNAMI; BUCK, 1997; RYBA; TIRINDELLI). Os V1Rs representam a primeira superfamília de receptores para feromônios descoberta e apresentam características semelhantes aos ORs, como expressão monogênica onde cada neurônio do VNO expressa apenas um tipo de receptor (RODRIGUEZ et al., 2002).

1.2 Regulação da expressão dos genes OR

Os genes que codificam para os receptores olfatórios apresentam um padrão de expressão muito peculiar. Em primeiro lugar, os genes OR são preferencialmente expressos nos neurônios olfatórios, em relação a outros tecidos. Em segundo lugar, sabe-se que cada um dos milhões de neurônios olfatórios presentes na cavidade nasal expressa apenas um dos 1000 (no caso dos camundongos) genes OR existentes no seu genoma (CHESS et al., 1994; MALNIC et al., 1999). Além disto, em cada neurônio olfatório, apenas um dos dois alelos de um dado gene OR é expresso, ou o materno ou o paterno, e nunca ambos (CHESS et al., 1994; SARAIVA et al., 2015; SERIZAWA et al., 2000). Finalmente, cada gene OR é expresso em neurônios olfatórios que são localizados em uma dentre quatro possíveis diferentes zonas de expressão de ORs no epitélio olfatório, Z1 a Z4, sendo que a Z1 é localizada na região apical e a Z4 na região mais ventral (MIYAMICHI et al., 2005; RESSLER et al., 1993; VASSAR et al., 1993). Estudos de hibridização in situ e de microarray indicam que apesar de os primeiros genes OR começarem a ser expressos no estágio embrionário E12 em camundongos, uma grande fração dos genes OR passa a ser expressa apenas após o nascimento (RODRIGUEZ-GIL et al.; SULLIVAN; BUCK, 1995; ZHANG, X. et al., 2004).

Pouco se sabe a respeito dos mecanismos através dos quais este padrão restrito de expressão gênico é estabelecido (para uma revisão recente ver (NAGAI; ARMELIN-CORREA; MALNIC, 2016). Trabalhos do grupo de Hitoshi Sakano (da

Universidade de Tokyo, Japão) e Randal Reed (da Johns Hopkins University School of Medicine, EUA) demonstraram a existência de um mecanismo de *feedback* negativo, onde a expressão de um receptor olfatório funcional inibe a expressão dos outros ORs, desta forma garantindo que cada neurônio olfatório expresse um único tipo de OR (SERIZAWA et al., 2003; WANG, S. S. et al., 2004). Se um pseudogene é expresso resultando na expressão de um OR não funcional, não ocorre a inibição e outro OR é escolhido para expressão. Portanto o mecanismo de *feedback* ocorre apenas se uma proteína OR funcional é expressa no neurônio olfatório.

Normalmente, a expressão de genes é regulada através da interação de fatores específicos de transcrição com elementos de DNA localizados em cis, a 5' dos genes. No caso de alguns genes OR, experimentos utilizando animais transgênicos indicaram que regiões curtas de DNA localizadas próximas à região 5' do gene são capazes de regular de maneira correta a sua expressão (QASBA; REED, 1998; ROTHMAN et al., 2005; VASSALLI et al., 2002; ZHANG; BREER; STROTMANN, 2007). Para outros ORs foi demonstrado que um segmento de DNA de aproximadamente 2 kb (região H), localizado 75 kb a 5' de um cluster de genes OR, é capaz de agir como um enhancer para ativar genes OR localizados a 3', em cis, a este segmento (SERIZAWA et al., 2000). O grupo de Richard Axel, da Universidade de Columbia, NY, sugeriu que a região H pode ativar genes OR localizados em outros cromossomos, indicando que este enhancer pode agir também em trans (LOMVARDAS et al., 2006). No entanto, camundongos que tiveram a sua região H deletada apresentam um padrão normal de expressão de seus genes OR, a não ser dos genes que estão localizados em cis em relação à região H, o que mostrou que este enhancer não é capaz de atuar em trans (FUSS; OMURA; MOMBAERTS, 2007; NISHIZUMI et al., 2007).

Realizamos em nosso laboratório uma análise extensiva onde as regiões promotoras de 198 genes ORs foram comparadas entre si (MICHALOSKI; GALANTE; MALNIC, 2006). Nossos resultados demonstraram que a maioria destes genes OR compartilha sítios de ligação para fatores de transcrição do tipo olf1 (O/E-1), localizados próximos ao início de transcrição destes genes. O fator de transcrição olf1 (O/E-1) é expresso em neurônios olfatórios e se liga a regiões regulatórias de vários genes que também são especificamente expressos nestes neurônios, como Gaolf, OMP e CNGC (WANG, M. M.; REED, 1993). No entanto, animais que são knockout para olf1 (O/E-1) não apresentam alterações na expressão de genes OR (WANG, S. S. et al., 2004). É possível que a ausência de um fenótipo neste caso se deva ao fato de que os neurônios olfatórios expressam três outros genes O/E like relacionados à olf1 (O/E-2, O/E-3 e O/E-4) (WANG, S. S.; BETZ; REED, 2002; WANG, S. S.; TSAI; REED, 1997). Além disto, como estes genes são expressos em todos os neurônios olfatórios, a mera presença destes sítios para olf1 (O/E-1) nos promotores dos genes ORs não explica o seu padrão mosaico de expressão gênica. Apesar disto, nossos resultados indicam que os fatores de transcrição do tipo olf1 (O/E like) devem ser importantes para a expressão dos genes OR. Sabe-se que a organização da cromatina no núcleo das células está relacionada com a expressão gênica (LANCTÔT et al., 2007; MISTELI, 2007; RAJAPAKSE; GROUDINE, 2011). Trabalhos recentes indicam que o núcleo dos neurônios olfatórios apresenta uma organização peculiar. Diferentemente do que se observa na maior parte das células eucarióticas, onde a heterocromatina constitutiva se encontra na periferia nuclear e a eucromatina está localizada principalmente no centro do núcleo (RAGOCZY; GROUDINE, 2014), nos neurônios olfatórios observamos uma arquitetura nuclear singular caracterizada por um grande bloco de heterocromatina constitutiva localizada centralmente e pela presença proeminente de domínios de heterocromatina facultativos que estão localizadas em torno deste bloco de heterocromatina constitutiva (ARMELIN-CORREA et al., 2014; CLOWNEY et al., 2012).

Nós também descobrimos que os dois alelos homólogos de um determinado gene OR são frequentemente segregados em compartimentos no núcleo, com um dos alelos localizados no bloco de heterocromatina constitutiva e a outra localizada na heterocromatina facultativa, ou próxima a ela como observado na Figura 3 (ARMELIN-CORREA et al., 2014). Diferentemente da heterocromatina constitutiva, a heterocromatina facultativa, por possuir maior plasticidade, está envolvida diretamente na regulação transcricional, podendo ter um papel importante na inativação da expressão gênica (como a observada na inativação do cromossomo X), bem como ser revertida a promover a expressão de genes durante o desenvolvimento. Os nossos resultados sugerem que a heterocromatina facultativa deve desempenhar um papel crucial na expressão de genes ORs (ARMELIN-CORREA et al., 2014).



Figura 3: Localização dos loci de um gene OR no núcleo de neurônios olfatórios. (A e B) Imagens de imuno-DNA FISH dos núcleos olfatórios com marcação para heterocromatina facultativa (H3K27me3 em verde), gene OR P2 (em vermelho) e heterocromatina constitutiva (fortemente corada com DAPI, em azul). No núcleo mostrado em A, os alelos podem ser visualizados em diferentes *stacks*; um alelo está dentro do bloco de

heterocromatina constitutiva (*stack* 19), e o outro se colocaliza com H3K27me3 (*stack* 12). No núcleo mostrado em B, os dois alelos são visualizados em um mesmo *stack* (retirado de ARMELIN-CORREA *et al.*, 2014). Os limites dos núcleos estão indicados pela linha branca.

1.3 RNAs não codificadores (ncRNAs)

Até recentemente, pensava-se que a maioria das moléculas de ácidos ribonucleicos (RNAs) estavam relacionados ao envio de informação genética para tradução de proteínas, com exceção apenas do RNA transportador (tRNA) e do RNA ribossômico (rRNA), que também desempenham funções relacionadas diretamente à tradução de proteínas. Porém, desde a década de 90 descobriu-se outros tipos de moléculas de RNA, que não são traduzidos e estão presentes em muitos organismos diferentes, afetando uma grande variedade de processos (LIU et al., 2005). Essas moléculas são chamadas de RNAs não-codificadores (ncRNAs).

Os RNAs não codificadores controlam uma gama notável de reações biológicas e processos, como iniciação da tradução, controle da abundância de RNA mensageiro (mRNA), arquitetura do cromossomo, manutenção de células-tronco, desenvolvimento do cérebro e músculos, secreção de insulina, dentre outras (MICHALAK, 2006).

Por volta de 98% do que é transcrito pelo genoma humano é constituído de ncRNA e a diferença na complexidade de um organismo pode ocorrer principalmente devido à vasta diferença na quantidade de ncRNAs presentes nos organismos eucarióticos e nos organismos mais simples (BURENINA, 2017; MATTICK, 2001).

Os ncRNAs geralmente são classificados pelo seu tamanho: (i) **RNAs não codificadores pequenos,** que possuem menos de 200 nucleotídeos, incluindo RNAs de infra-estrutura como tRNAs, rRNAs e pequenos RNAs nucleares / spliceossoma (snRNAs), bem como vários tipos de RNAs reguladores, incluindo micro RNAs (miRNAs), pequenos RNAs de interferência (siRNAs), piwi-RNAs (piRNAs) e pequenos RNAs nucleolares (snoRNAs) (MATTICK; MAKUNIN, 2005), e (ii) **RNAs não codificadores longos (IncRNAs)** que podem variar de algumas centenas bases até bem mais de 100 kilobases de comprimento (FURUNO et al., 2006; MERCER; DINGER; MATTICK, 2009).

1.4 Importância dos ncRNAs na regulação gênica

Como já mencionado, os RNAs não codificadores controlam uma gama notável de reações biológicas e processos incluindo a regulação da expressão gênica. A classe de ncRNAs mais bem caracterizada e estudada com essa função são os miRNAs. Os miRNAs possuem um papel importantíssimo na regulação da expressão gênica e o mecanismo de como isso ocorre está bem descrito na literatura. Foi relatado também que os miRNAs se ligam a transcritos pseudogenes e IncRNAs através da interação com os response elements, que competem pela ligação destes miRNAs com os seus mRNAs alvos. Esses RNAs atuam como esponjas moleculares ou chamarizes e reprimem a degradação dos mRNAs alvo e, portanto, são chamados de RNAs endógenos competitivos. Para poder agir como um RNA endógeno competitivo, um response element de um IncRNA requer interações incompletas com o miRNA. Assim, as interações de IncRNAs com miRNAs não desencadeiam a decomposição dos IncRNAs, ou apenas desencadeiam lentamente a decomposição. A atividade de um RNA endógeno competitivo regula diversos processos celulares de desenvolvimento e doenças (YAMAMURA et al., 2018).

Outro tipo de ncRNA bem conhecido por estar envolvido na regulação da expressão gênica é conhecido como *Long Intergenic Non-coding RNAs* (lincRNA).

lincRNAs são RNAs não codificadores longos intergênicos, definidos como transcritos em uma região distante em pelo menos 1kb de algum gene codificador de proteína, podendo ou não sofrer *splicing*. Essa classe de ncRNAs apresenta características semelhantes aos dos RNAm codificadores de proteínas tais como a presença da estrutura cap 5' do mRNA, transcrição a partir de uma região promotora independente (identificada por meio de marca de histona H3K4me3) e presença de RNA pol II no sitio de início de transcrição, além de poderem ser poli-adenilados ou não (CONLEY; KING JORDAN, 2012; GUTTMAN et al., 2009). Há fortes evidências de que o mecanismo de ação desses RNAs não codificadores longos envolve a interação entre IncRNA e proteínas que compõem complexos remodeladores de cromatina com funções tanto ativadoras como repressoras (FLYNN; CHANG, 2012; KHALIL et al., 2009).

Um exemplo bem descrito de lincRNA que interage fisicamente com proteínas durante a modulação da expressão genica é o do lincRNA HOTAIR. O lincRNA HOTAIR é transcrito a partir do *locus* HOXC, e silencia os genes HOXD, bem como outros alvos. Inicialmente, foi relatado que o HOTAIR se liga ao *polycomb repressive complex* 2 (PCR2) e direciona esse complexo para as regiões do promotor HOXD (RINN et al., 2007). Este complexo medeia o silenciamento de genes através da reorganização da cromatina por metilação da lisina 27 da histona H3 (H3K27). O lincRNA HOTAIR serve como um andaime para pelo menos dois complexos distintos de modificação de histonas. Um domínio 5 'de HOTAIR liga PRC2 enquanto um domínio 3' de HOTAIR liga o complexo LSD1/CoREST/REST que remove a dimetilação da lisina 4 na histona 3 (H3K4me2), que também está relacionado ao silenciamento de genes HOXD (TSAI et al., 2010). Outro exemplo bem estudado é o lincRNA *Xist* (*X-inactivation specific transcript*). *Xist* é um ncRNA conhecido por

promover a compensação de dose em células humanas por meio da sinalização da inativação do cromossomo X (BROWN et al.; SAMUEL C. CHANG, 2006). A inativação do cromossomo X por *Xist* é iniciada pelo espalhamento em *cis* a partir do futuro cromossomo X inativo (AVNER; HEARD, 2001; PLATH et al., 2003), recrutamento PRC2 (PLATH et al., 2003; SILVA et al.; ZHAO, J. et al., 2008), formação de um compartimento de silenciamento transcricional nuclear (CHAUMEIL et al., 2006) seguido por modificações repressivas da cromatina incluindo a trimetilação da lisina 27 da histona 3 (H3K27) (PLATH et al., 2003; SILVA et al., 2003; SILVA et al.). Essa função de *Xist* de compactar a cromatina e silenciar a expressão gênica é mediada por distintos domínios de RNA.

Sabe-se que as interações entre *Xist* e proteínas são mediadas principalmente pelas estruturadas regiões do RNA *Xist* (CHEN et al., 2016; CHU et al., 2015; FANG et al., 2015) ou motivos *Xist* (SMOLA et al., 2016). Em camundongos, *Xist* possui seis regiões repetitivas conservadas em tandem, nomeadas de A a F, que são essenciais para sua função (BROCKDORFF, 2002; WUTZ; RASMUSSEN; JAENISCH, 2002). As regiões repetitivas em tandem de *Xist* são conservadas em mamíferos vertebrados, entretanto fora dessas regiões repetitivas a conservação é relativamente pobre (NESTEROVA et al., 2001). O silenciamento transcricional necessita desses domínios repetitivos de A que interagem com o complexo regulador de cromatina PRC2 (ZHAO, J. et al., 2008), visto que a localização da cromatina necessita de vários domínios distintos e interações com proteínas associadas com a matriz nuclear (BELETSKII et al., 2001; CHAUMEIL et al., 2006).

1.5 RNAs não codificadores presentes no sistema olfatório

Magklara e colaboradores (MAGKLARA et al., 2011) demonstraram recentemente que os genes OR apresentam metilações em histonas que representam marcas para heterocromatina constitutiva (H3K9me3 e H4K20me3) nos núcleos de neurônios olfatórios, mas não nos núcleos de células do fígado. Este trabalho mostra que a regulação da estrutura da cromatina deve desempenhar um papel muito importante na regulação da expressão dos genes OR. No entanto, pouco se sabe a respeito de como estas alterações na estrutura da cromatina são estabelecidas e reguladas nos neurônios olfatórios. Praticamente nada se sabe também a respeito de RNAs não codificadores expressos no epitélio olfatório. Até hoje, um único trabalho desenvolvido pelo grupo da Dra. Catherine Dulac, da Harvard University, analisou a expressão de microRNAs (miRNAs) neste tecido (CHOI et al., 2008). Utilizando um microarray de miRNAs, foi possível identificar miRNAs que são expressos preferencialmente no epitélio olfatório de camundongos. Além disto, através do sequenciamento de miRNAs purificados a partir do epitélio olfatório, os autores identificaram miRNAs adicionais que não estavam presentes no microarray. Dentre estes miRNAs, o grupo analisou uma família de miRNAs (família miRNA-200) que é necessária para a diferenciação de neurônios olfatórios. O epitélio olfatório é organizado em basicamente três camadas: uma camada apical que contém as células de suporte, uma camada central onde estão os neurônios olfatórios maduros e uma camada basal onde estão as células basais, que são as células-tronco olfatórias, que podem se diferenciar em neurônios olfatórios, regenerando assim o tecido (Figura 4).

Interessantemente, Choi e colaboradores (2008) observaram que o desenvolvimento do epitélio olfatório é normal, com expressão normal dos genes OR, em um camundongo *knockout* condicional onde o gene que codifica para *Dicer*, uma enzima necessária para a produção dos miRNAs funcionais, é inativado apenas nos neurônios olfatórios maduros.



Figura 4: O epitélio olfatório. Hibridização *in situ* em corte coronal da região da cavidade nasal com sonda para OMP (*Olfactory Marker Protein*) que marca (em cor púrpura) apenas os neurônios olfatórios maduros do epitélio olfatório que reveste as turbinadas na cavidade nasal. Uma região ampliada do epitélio olfatório está apresentada à direita, mostrando as três camadas do epitélio olfatório: a camada que contém as células suporte, localizada na região apical do epitélio, a camada central contendo os neurônios olfatórios maduros (marcados), e a camada basal, contendo as células basais que são capazes de se diferenciar em neurônios olfatórios.

Por outro lado, o *knock out* condicional de *Dicer* nas células olfatórias precursoras, que estão localizadas na região basal do epitélio olfatório (Figura 7), resultou na degeneração do epitélio olfatório. Estes resultados indicam que os miRNAs são dispensáveis em neurônios olfatórios já diferenciados, mas são necessários para o desenvolvimento das células precursoras em neurônios olfatórios maduros. Recentemente Wang e colaboradores demonstraram que vários IncRNAs são diferencialmente expressos no bulbo olfatório de camundongos mais velhos (camundongos com mais de 20 semanas de vida) em comparação com

camundongos mais jovens. Além disso, os autores associaram o declínio da função olfatória de camundongos velhos à expressão diferencial dos lincRNAs NONMMUT004524 e NONMMUT000384 por ambos estarem associados à *neuroactive ligand-receptor interaction pathway* no bulbo olfatório de camundongos velhos (WANG, M. et al., 2017). Esta via consiste em uma variedade de sinais moleculares incluindo vários tipos de neuroreceptores que estão localizados na membrana plasmática e envolvidos na transdução de sinal do ambiente extracelular para as células (SU et al., 2009).

2 OBJETIVOS

2.1 Objetivo geral:

Determinar o repertório de RNAs não codificadores anotados expressos no epitélio olfatório ao longo do desenvolvimento, com enfoque principal em lincRNAs e miRNAs.

2.1.1 Objetivos específicos:

- Sequenciar os RNAs longos e pequenos de epitélio olfatório de camundongos recém-nascidos (com 1 a 3 dias de idade) e camundongos jovens (com 4 semanas de idade).
- 2. Comparar os transcritomas das duas idades.
- Analisar a expressão dos receptores olfatórios ao longo do desenvolvimento do epitélio olfatório.
- 4. Determinar o repertório de miRNAs no epitélio olfatório ao longo do desenvolvimento do epitélio olfatório.
- 5. Determinar o repertório de lincRNAs anotados no epitélio olfatório ao longo do desenvolvimento do epitélio olfatório.
- Selecionar e validar lincRNAs que sejam preferencialmente expressos no epitélio olfatório.

3 MATERIAL E MÉTODOS

3.1 Extração de RNA

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 avalizados pela CEUA. A extração de RNA a partir de epitélio olfatório foi feita com Trizol (*Invitrogen*) seguindo as especificações do fabricante. A quantificação do RNA obtida foi feita através da leitura em espectrofotômetro *Nanodrop*. Para eliminar a contaminação com DNA genômico, todo RNA utilizado na construção das bibliotecas foi tratado com DNAse (*Promega*), seguindo-se as recomendações do fabricante.

A qualidade dos RNAs foi verificada utilizando o equipamento *Bioanalyzer* 2100 (*Agilent Technologies*), que realiza uma eletroforese capilar de alta tensão. A integridade das amostras foi avaliada com o programa 2100 *Expert (Agilent)*. Este programa atribui um valor de integridade de RNA (*RNA Integrity Number*, RIN), que permite uma estimativa da integridade através de todo traçado eletroforético da amostra, e não apenas da razão entre os RNA ribossomais. Apenas amostras que tinham uma boa integridade (RIN > 7) foram utilizadas em todos os experimentos.

3.2 Produção de cDNA

Usando o sistema RQ1 RNAse free DNAse (*Promega*), tratamos 1 µg de RNA extraído com Trizol. Foram adicionados 100 ng de oligo(dT) ou 100 ng de *random* primers (Invitrogen) e completou-se o volume para 13,5 µL com água livre de DNase/RNase. A reação foi incubada por 2 min a 70° C e esfriada rapidamente no gelo. Adicionou-se 4 µL de tampão 5x, 1µL de dNTPs (10 µM cada), 0,5 µL de RNAse OUT (Invitrogen) e 1,0 µL de *Superscript* II 200U/µL (*Invitrogen*) e incubou-se a reação por 1h a 42° C. Em seguida a reação foi aquecida a 94° C por 5 min, diluída para 100 µL com água livre de DNase/RNase e a armazenada em alíquotas a - 80° C.

3.3 Construção das bibliotecas de cDNA e sequenciamento por ligação

O sequenciamento do RNA foi realizado em colaboração com Pedro Alexandre Favoretto Galante e Anamaria Aranha Camargo do Instituto Sírio-Libanês de Ensino e Pesquisa no Hospital Sírio-Libanês. O SOLiD™ Total RNA-Seq Kit (Applied Biosystems) foi utilizado para a construção das bibliotecas de cDNA, seguindo as especificações do fornecedor. Resumidamente, o RNA total, livre de contaminação com DNA genômico, foi depletado de RNA ribossômico através da utilização do Ribominus kit (Invitrogen). Em seguida o RNA purificado foi fragmentado por digestão com RNAse III. Ao RNA fragmentado foram ligados adaptadores, e esse material foi utilizado em reação de transcrição reversa. O cDNA obtido foi fracionado em gel denaturante de poliacrilamida, para a purificação de fragmentos com tamanho entre 100 e 200 pb. O material purificado foi utilizado como molde em reação de PCR, utilizando iniciadores específicos para os adaptadores empregados na construção das bibliotecas. Após amplificação, o produto de PCR foi purificado e quantificado com o auxílio do fluorômetro Qubit (Invitrogen) e utilizado no PCR em emulsão. Para tanto, a solução contendo os fragmentos de PCR, beads recobertas com sequências complementares aos adaptadores utilizados na construção da biblioteca de cDNA e reagentes para

30

amplificação dos fragmentos por PCR foi homogeneizada na presença de óleo sob agitação constante. Ao final, foi obtida uma emulsão com microbolhas (microrreatores) contendo uma *bead*, uma molécula molde de cDNA e reagentes para a amplificação desta molécula por PCR. A emulsão então foi submetida à amplificação em um termociclador convencional, de maneira que em cada microrreator ocorre a amplificação clonal de uma única molécula. Após a amplificação, os *beads* recobertos por milhares de cópias de uma mesma molécula foram recuperados e submetidos ao sequenciamento por ligação na plataforma SOLiD V4.0 gerando sequências de 50-75 pb.

3.4 Programas utilizados nas análises de bioinformática

As análises de bioinformática foram realizadas em colaboração com o pesquisador Dr. Pedro Alexandre Favoretto Galante do Instituto Sírio-Libanês de Ensino e Pesquisa no Hospital Sírio-Libanês.

3.4.1 Cuffdiff

Cuffdiff é um programa separado incluído no Cufflinks utilizado nas análises diferencias que calcula a expressão em duas ou mais amostras e testa significâncias estatísticas de cada mudança observada na expressão entre elas. O modelo estatístico utilizado para avaliar as mudanças assume que o número de leituras produzido por cada transcrito é proporcional à sua abundância, mas flutua devido à variabilidade técnica durante a preparação da biblioteca e sequenciamento e por causa variabilidade biológica entre repetições do mesmo experimento (TRAPNELL et al., 2012).

3.4.2 DESeq

DESeq é um pacote de R para analisar dados de contagem de ensaios de sequenciamento de última geração, como RNA-Seq e teste para a expressão diferencial. DESeq utiliza um modelo baseado na distribuição binomial negativa (ANDERS; HUBER, 2010).

3.4.3 EdgeR

Edger é um pacote de software Bioconductor para a análise de expressão diferencial de dados de contagem em replicatas. Um modelo de sobre dispersão de Poisson é usada para explicar tanto a variabilidade biológica como a técnica (ROBINSON; MCCARTHY; SMYTH, 2010).

3.4.4 TargetScan

TargetScan é um programa que prevê alvos biológicos de miRNAs, procurando a presença de sítios 8mer e 7mer conservados que correspondem à região de *seed* de cada miRNA (LEWIS, B. P.; BURGE; BARTEL, 2005). A região denominada *seed* corresponde à sequência de 6 a 8 nucleotídeos de comprimento localizado na porção 5' do miRNA (LEWIS, B. P. et al., 2003), sendo esta região a responsável pelo reconhecimento dos alvos dos miRNAs (BARTEL, 2009; NIELSEN et al., 2007). Como opção, os sítios não conservados também são previstos. Também são identificados locais com inadequações na região da *seed* que são compensados por emparelhamento da porção 3' conservada (FRIEDMAN et al., 2009).

3.5 RT-PCR

Para cada reação de PCR foram utilizados 2 µL de cDNA, 2,5 µL de solução tampão, 1 µL de Mg⁺²,1 µL de dNTP, 0,5 µL de Taq polimerase, 1 µL de cada iniciador (primers *forward* e *reverse*) específicos para GAPDH (gliceraldeído fosfato dehidrogenase) e para os lincRNAs de 1 a 10. Para cada reação foi incluído um controle negativo, com os primers, porém sem cDNA. As reações de amplificação foram realizadas em um termociclador utilizando-se o seguinte programa: 2 min a 95°C, 35 ciclos 45 seg a 95 °C, 45 seg a 58 °C e 1 min a 72°C, 1 ciclo 10 min a 72°C e 4°C por tempo indeterminado. Os produtos de PCR foram então submetidos à eletroforese em gel de agarose 1,5%. Para o RT-PCR dos lincRNAs desenhamos os seguintes primers:

LincRNA 1:

NC1 *Forward* 5'- CTC AGC GAT CAA CTG GTT GAC- 3' NC1 *Reverse* 5'- GGT GGA GCT GAT GTG ATA AGC- 3'

LincRNA 2:

NC2 Forward 5'- TGC CTT CTG GTC ATC TCA GG- 3' NC2 Reverse 5'- TAC TTC AGC TCC TGC CCT TG- 3'

LincRNA 3:

NC3 Forward 5'- CCT CTG CCT CTC AAG TGT TG- 3' NC3 Reverse 5'- CAC CAC ATC TCC CTT GGA TC- 3' LincRNA 4:

NC4 Forward 5'- TTT CAC ACC AGA CTC TGG CC- 3' NC4 Reverse 5'- AAG CTG GAA GTG TGG ACA GTG- 3'

LincRNA 5:

NC5 Forward 5'- GCT GCA GAA GTC ACC AAA GG 3' NC5 Reverse 5'- TAA CGG ACT GCA GGC TGT AG- 3'

LincRNA 6:

NC6 Forward 5'- CTG GCT CAT TCA ACT CTG CAC- 3' NC6 Reverse 5'- AGC CAT AGT TCA TCA CTG AGT G- 3'

LincRNA 7:

NC7 Forward 5'- ATG GTG CCA GAC AAG ACT GC- 3' NC7 Reverse 5'- TTC ACA CCC AGA AGA GGC AG- 3'

LincRNA 8:

NC8 Forward 5'- TTG AGT TCC TGC CTC AGC TAC- 3' NC8 Reverse 5'- GCT TCT GAG TCT ACT GTG ATG- 3'

LincRNA 9:

NC9 *Forward* 5'- GGT GTC TCA CCG AAA GGT AG- 3' NC9 *Reverse* 5'- TGG CAG TGG TAG CAC ATG CC- 3'

LincRNA 10:
NC10 *Forward* 5'- GAT GGA CCA TGA GCA CAT GG- 3' NC10 *Reverse* 5'- CTG CAG AGA CAT CTG CTT GC- 3'

Para o RT-PCR do Ezh2 desenhamos os seguintes primers: Ezh2_F *forward* 5'- GTATGAGTGCTTCTTACATC-3' Ezh2_R *reverse* 5'- GGTGCTATGATAGTGGACTC-3'

3.6 Clonagem dos moldes para as sondas

Regiões das sequências dos genes dos lincRNAs foram amplificadas e clonadas no vetor pGEM (*pGEM®-T Vector Sysytem I Kit, Promega*). Os plasmídeos foram linearizados com as enzimas de restrição adequadas e, em seguida, purificados com o *kit Illustra GFX PCR DNA and Gel Band Purification (GE Healthcare*). Regiões das sequências codificadoras dos genes Ezh2 foram amplificadas e clonadas no vetor pCRII (*TA Cloning® Kit, Invitrogen*).

3.7 Sequenciamento automatizado de DNA

Sequenciamos os amplicos obtidos no RT_PCR para confirmamar a identidade dos mesmos. O sequenciamento de DNA foi realizado no aparelho *ABI PRISM-TM 3100 (Applied Biosystems*), conforme recomendações do fabricante. Em cada reação de sequenciamento foram utilizados 2,0 µL de *Big Dye terminator mix (Perkin Elmer*), 100-200 ng de DNA, 9,6 pmol de cada oligonucleotídeo, Tris-HCI 200 mM pH 9,0 e MgCl₂ 5 mM em um volume final de 15 µL. A reação foi incubada por 2 minutos a 95°C, seguida de 35 ciclos de 45 segundos a 96°C, 30 segundos a 50°C e 4 minutos a 60°C. Terminada a reação, as amostras foram precipitadas com etanol 100% e glicogênio 1 mg/mL, lavadas com etanol 70% (v/v), suspensas em 20 µL de

formamida, desnaturadas a 94°C por 5 minutos e submetidas ao sequenciamento automatizado.

3.8 Síntese das sondas de RNA *sense* e *anti-sense* marcadas com digoxigenina

Uma reação de transcrição foi preparada em solução contendo 1 µg de plasmídeo linearizado, tampão de transcrição 1x (*Ambion*), DIG *RNA labeling mix* 1x (*Roche*), 40 U de RNA polimerase T7 ou Sp6 (*Ambion*), 1 µL de inibidor de RNase (*RNase OUT, Invitrogen*). Essa solução foi incubada por 2 horas a 37 °C e, em seguida, filtrada em colunas de Sephadex-50 (*Quick Spin G-50, GE Healthcare*) e misturadas com 20 µL de formamida. Em seguida, 2 µL foram aplicados em gel de agarose 1,5% para a quantificação, o restante foi armazenado a -80 °C.

3.9 Preparação das lâminas com cortes de camundongos C57BL6

Narizes de camundongos da linhagem C57BL/6 com 4 semanas de vida foram dissecados e embebidos em TissueTek OCT (*Sakura*) e congelados em gelo seco. Com ajuda de um criostato (*MICRON*), o embrião emblocado em OCT foi seccionado na espessura de 14 ou 16 µm e os cortes foram coletados em lâminas de vidro silanizadas (*StarFrost*). Até o momento do uso, tanto o bloco quanto as lâminas contendo as secções foram mantidos a -80 °C.

3.10 Hibridização *in situ*

Primeiramente, as lâminas foram secas por 10 minutos a 50°C para ajudar na aderência dos cortes à lâmina. Os cortes foram então fixados por 10 minutos em formaldeído 4% pH 7,4 recém preparado e lavados três vezes em PBS por 5

minutos. A acetilação foi realizada com solução contendo água DEPC e trietanolamina onde as lâminas ficaram por 10 minutos sobre agitação e então foi realizada novamente a lavagem com PBS três vezes por 5 minutos. A préhibridização foi realizada durante duas horas a 72 °C, incubando as lâminas em 500 µL de solução contendo formamida 50%, SSC 5X, Denhardt's solution 5x (Invitrogen) e 125 µg de RNA de levedura (yeast tRNA, Invitrogen). Nessa mesma solução, 200 ng de sonda foram diluídos e incubados com as lâminas durante pelo menos 16 horas a 65 °C. As lâminas foram, então, lavadas a 72 °C uma vez em SSC 5X por 10 minutos e duas vezes em SSC 0,2X por 30 minutos. Seguiu-se mais uma lavagem em SSC 0,2X, porém durante 5 min à temperatura ambiente. Os tecidos foram digeridos durante 30 minutos a 37 °C em solução de Tris pH 7,5 10 nM, NaCl 400 nM e RNase A 20 µg/mL. A detecção foi realizada com o kit DIG Wash and Block Buffer Set (Roche), iniciando com uma lavagem em Wash Buffer 1x, seguida de bloqueio de uma a duas horas em ácido maleico 1x e reagente de bloqueio 1x. O anticorpo anti-digoxigenina conjugado a fosfatase alcalina (Roche, 1:500) foi diluído em solução de bloqueio e incubado por uma hora à temperatura ambiente. Em seguida, foram realizadas 3 lavagens, de 20 minutos cada, em Wash Buffer 1x (ácido maléico 1x contendo 3% de Tween-20). Então, as lâminas foram incubadas durante 5 minutos em Detection Buffer 1x (Tris pH 9,5 100 mM, NaCl 100 mM). A solução de detecção foi preparada adicionando-se 2,4 mg de levamisole e 200 µL de NBT/BCIP a 10 mL de detection buffer 1x, depois foi adicionada às lâminas que foram incubadas overnight à temperatura ambiente. A reação da fosfatase alcalina foi interrompida com uma lavagem de 2 minutos em Tris 10 mM pH 8 e EDTA 1 mM, seguida de uma lavagem em água. Finalmente, as lâminas foram montadas com gelvatol.

4 RESULTADOS E DISCUSSÃO

4.1 Dissecção do epitélio olfatório

Para o sequenciamento de RNAs foram feitas quatro extrações de epitélio olfatório de camundongos (C57BL/6). Na Figura 5 pode-se visualizar o epitélio olfatório durante a remoção deste tecido.

Amostra A: quatro camundongos com 1 mês de vida.

Amostra B: três camundongos com 1 mês de vida.

Amostra C: nove camundongos neonatos com 3 dias de vida.

Amostra D: oito neonatos com 3 dias de vida.

O tecido recolhido foi congelado e descongelado na presença de Trizol. O RNA total foi purificado e então quantificado



Figura 5: Demonstração do epitélio olfatório de camundongo. Foto tirada durante a remoção do epitélio olfatório do camundongo A (neonato) e B (4 semanas de vida).

4.2 Quantificação do RNA total

As quantificações feitas no espectrofotômetro Nanodrop resultaram nos seguintes valores:

Amostra A = 1,02 μg/μL RNA total	Razão 260/280: 1.86
(4 semanas de vida)	Razão 260/230: 1.24
Amostra B = 2,73 μ g/ μ L RNA total	Razão 260/280: 1.93
(4 semanas de vida)	Razão 260/230: 0.69
Amostra C = 1,86 μ g/ μ L RNA total	Razão 260/280: 1.88
(Neonatos)	Razão 260/230: 1.83
Amostra D = 1,42 μ g/ μ L RNA total	Razão 260/280: 1.96
(Neonatos)	Razão 260/230: 1.90

Para verificar a integridade do RNA, as amostras foram submetidas à eletroforese em gel de agarose 1,5% como mostram as Figuras 6 A e B.



Figura 6: Avaliação da integridade do RNA. A. RNA total de camundongos neonatos, canaleta 1 (500 ng amostra A), canaleta 2 (500 ng amostra B). B. RNA total de camundongos com 4 semanas de vida, canaleta 1 (500 ng amostra C), canaleta 2 (500 ng amostra D). A e B figura mostra foto de gel de agarose 1,5 % corado com brometo de etídeo. C e D Avaliação da integridade do RNA utilizando o Bioanalyzer. (C) RNA extraído de camundongos neonatos. (D) RNA extraído de camundongos com 4 semanas de vida.

De acordo com os resultados obtidos através da corrida em gel, podemos avaliar que a integridade do RNA estava de acordo com o esperado na amostra A de camundongos com um mês e nas amostras de neonatos, sendo possível visualizar com clareza duas bandas mais fortes correspondentes às subunidades ribossomais 28S, 18S e uma mais fraca correspondendo à subunidade 5S. A amostra B foi desprezada já que a primeira banda está mais fraca indicando degradação do RNA.

Estas amostras de RNA foram também analisadas utilizando-se *Bioanalyzer* 2100 (*Agilent Technologies*). Utilizando esse método obtivemos um RIN (*RNA Integrity Number*) de 7,7 para os RNAs extraídos de camundongos neonatos (Figura 9 C). Para os RNAs extraídos de camundongos com 4 semanas de vidas obtivemos o valor de RIN de 8.1 (Figura 9 D). Por meio dessas analises avaliamos que as amostras de RNAs extraídos do epitélio olfatório possuíam boa qualidade para serem sequenciados.

4.3 Análises do sequenciamento

Construímos duas bibliotecas de pequenos RNAs e duas bibliotecas do transcritoma total como esquematizado no fluxograma (Figura 7).



Figura 7: Fluxograma da estratégia utilizada para sequenciar os transcritos do epitélio olfatório. O epitélio olfatório de recém-nascidos (P3) e 4 semanas de idade (4W) foi dissecado para extração de RNA. O RNA ribossomico foi removido usando o kit *RiboMinus* ™ (*Invitrogen*). Bibliotecas para pequenos RNAs e para o transcritoma total foram preparadas e sequenciadas usando a plataforma SOLiD.

O sequenciamento do epitélio olfatório de camundongos neonatos e com quatro semanas de vida foi realizado utilizando-se a plataforma SOLiD. O número total de *reads* obtidos nas bibliotecas de Transcritoma e pequenos RNAs estão listados na Tabela 1. Tabela 1: Reads obtidos no sequenciamento SOLiD. Os números totais de reads obtidos das bibliotecas do transcritoma e de pequenos RNAs em camundongos neonatos e camundongos com 4 semanas estão apresentados.

Amostras	Transcritoma total	Pequenos RNAs
Neonatos	220,873,385	65,550,415
4 semanas	231,348,303	78,952,593
Total	452,221,688	144,503,008

O número de *reads* que mapeiam no genoma de camundongos (GRCm38/mm10) e correspondem a RNAs presentes nas bibliotecas estão listados

na Tabela 2.

Tabela 2: Reads mapeados. Os números de reads (% dos reads totais) das bibliotecas do transcritoma e de pequenos RNAs em camundongos neonatos e camundongos com 4 semanas que puderam ser mapeados no genoma estão apresentados.

Amostras	Transcritoma total	Pequenos RNAs
Neonatos	114,894,462 (52%)	45,412,275 (69%)
4 semanas	113,886,582 (49%)	50,772,004 (64%)
Total	228,781,044 (51%)	96,184,279 (66%)

O número de *reads*, que apresentam FPKM \geq 2 (FPKM = *Fragments Per* Kilobase of exon model per Million mapped reads = n° de reads que mapeiam em exons do transcrito / [nº total de sequências (milhão) x tamanho do exon (kb)]), que mapeiam no genoma para genes codificadores, genes não codificadores e aqueles que não possuem função conhecida seguem na Tabela 3.

Amostras	Genes codificadores	Genes não codificadores	Não anotados	Total
Neonatos	12,242	1,246	3,295	16,783
4 semanas	11,799	1,207	4,352	17,358

Tabela 3: Total de *reads* **para RNAs longos.** Número de *reads* que correspondem a genes codificadores, não codificadores e não anotados Expressed genes: FPKM >= 2

4.4 Análise dos RNAs longos

Nos bancos de RNAs longos identificamos, como esperado, transcritos de genes que são especificamente expressos em epitélio olfatório (Tabela 4). Um dado importante é a presença do transcrito correspondente ao OMP (*Olfactory Marker Protein*). O gene OMP é expresso apenas em neurônios olfatórios maduros, e podemos observar que sua expressão é mais abundante em OE de 4 semanas do que em neonatos, como seria esperado, já que em camundongos de 4 semanas o número de neurônios olfatórios maduros é maior do que em neonatos. RIC8B também é um gene específico de neurônios olfatórios maduros. Além disso, os genes OR que também são expressos apenas no epitélio olfatório estão presentes em nossos dados. Cada OR apresentou baixos níveis de expressão como esperado, já que cada um dos 1000 ORs é expresso em apenas 0,1% dos neurônios olfatórios (RESSLER et al., 1993; VASSAR et al., 1993).

Gene	Neonatos (FPKM)	4 semanas (FPKM)
OMP	79.4255	545.743
RIC8B	29.6544	153.336
Olfr 60	14.7869	26.9414
Olfr 15	10.3516	23.5755
Olfr 533	6.77719	23.7291
Olfr 1507	5.87564	21.2753
Olfr 1321	4.69354	8.762
Olfr 539	4.56816	6.82214
Olfr 166	4.54571	11.5149
Olfr 309	4.09144	13.7442

Tabela 4: Transcritos obtidos de genes que são especificamente expressos no epitélio olfatório, com seus respectivos FPKMs, nos dois estágios de desenvolvimento.

Na Figura 13 mostramos o padrão de expressão dos genes OR nos dois diferentes estágios de desenvolvimento (neonatos x 4 semanas). Como esperado, observamos que uma grande fração dos genes OR não é expressa ou pouca expressa em camundongos neonatos e que com o desenvolvimento os mesmos começam a ser expressos, como já demonstrado por diversos autores (RODRIGUEZ-GIL et al.; SULLIVAN; BUCK, 1995; ZHANG, X. et al., 2004). Além disso, em camundongos com 4 semanas de vida a expressão destes genes é muito mais abundante, como mostrado no *heat map* dos genes OR diferencialmente expressos (Figura 8).



Figura 8. *Heat map* dos genes OR expressos no epitélio olfatório. a) Todos os genes OR expressos presentes em nossas amostras (Coluna esquerda = 6 réplicas técnicas do sequenciamento para camundongos com 4 semanas de vida / Coluna direita = 6 réplicas técnicas para camundongos neonatos).

Ao comparar o perfil de expressão dos genes de RNAs em nosso transriptoma de neonatos e de 4 semanas (Arquivos suplementares 1 e 2), observamos que 15.036 genes estão expressos em ambas as amostras, 2.322 e 1.747 são expressos exclusivamente em camundongos com 4 semanas de vida e neonatos e respectivamente (Figura 9).



Figura 9: Expressão dos genes de RNAs longos entre camundongos neonatos e com 4 semanas. Número de genes expressos em camundongos neonatos (azul) e com 4 semanas (cinza). A intersecção entre os círculos indica quantos genes são expressos em ambas as amostras.

A análise de expressão diferencial de transcritos entre as amostras de neonato e 4 semanas para RNAs codificadores, não codificadores e RNAs desconhecidos foi realizada utilizando-se os programas Cuffdiff e DESeq. Como mostrado na Tabela 5.

Tabela 5: Número de genes com expressão diferencial. (log2 fold change > 1 || log2 fold change < -1) * Cuffdiff / ** DESeq / *** Cuffdiff \cap DESeq).

Amostras	Codificador	Não codificador	Não anotados
Neonatato x 4 semanas*	2,809	451	2,315
Neonatato x 4 semanas **	4,250	1,213	2,736
Neonatato x 4 semanas ***	2,236	369	2,035

4.5 Análise dos miRNAs

Os *reads* obtidos para miRNAs mapeados no genoma foram sobrepostos a miRNAs maduros conhecidos (Tabela 6). Do total de *reads* para miRNAs obtidos em ambas as amostras, 28,30% correspondem a miRNAs maduros conhecidos.

Amostras	Número de <i>reads</i> para miRNAs (% do total # de <i>reads</i>)		
Neonatos	17,504,932 (26,70%)		
4 semanas	23,395,692 (29.63%)		

Tabela 6: *Reads* que correspondem a miRNAs maduros conhecidos.

O comprimento dos miRNAs sequenciados podem variar entre 16 a 35 nucleotídeos, porém a maioria das sequências identificadas descritas na literatura possuem em torno de 22 nucleotídeos (EBHARDT; FEDYNAK; FAHLMAN, 2010). O comprimento dos miRNAs sequenciados em nossas bibliotecas apresenta-se em maior frequência na faixa de 22 a 23 nucleotídeos (Figura 10), como observado também em (EBHARDT et al., 2010).



Figura 10: Mapeamento do perfil dos miRNAs para cada tamanho. Eixo Y: Frequência dos transcritos. Eixo X: Tamanhos dos transcritos.

Dentre 957 miRNAs expressos em neonatos e 827 miRNA expressos em camundongos com 4 semanas de vida, foi observado que há 477 miRNAs que possuem expressão diferencial entre as duas amostras (Arquivo suplementar 3).

Nas Tabelas 7 e 8 estão mostrados os 20 miRNAs que são mais diferencialmente expressos entre os dois estágios selecionados por Edge-R e DESeq. Na Tabela 7 estão os que são mais expressos em neonatos do que em 4 semanas, e na Tabela 8 estão mostrados os que são mais expressos em 4 semanas que em neonatos.

Tabela 7: miRNAs mais diferencialmente expressos em neonatos do que em 4 semanas. Os miRNAs estão ordenados de acordo com seus valores de fold change (Fold positivo = *upregulated* em Neonato).

miRNA	Log 2 Fold (CPM+1) Change
mmu-miR-1a-2-5p	8,514
mmu-miR-679-3p	8,514
mmu-miR-299a-3p	8,462
mmu-miR-206-5p	8,445
mmu-miR-410-5p	7,811
mmu-miR-133b-5p	7,597
mmu-miR-483-3p	7,573
mmu-miR-496a-5p	7,491
mmu-miR-133b-3p	7,317
mmu-miR-544-3p	7,184
mmu-miR-1197-3p	7,174

O miRNA mais diferencialmente expresso em camundongos neonatos é o mmu-miR-1a-2-5p. Este miRNA faz parte de uma grande família de miRNAs conhecida como miR-1. Essa família de miRNAs é fundamental no desenvolvimento e fisiologia dos tecidos musculares, incluindo o coração (TOWNLEY-TILSON; CALLIS; WANG, 2010; ZHAO, Y. et al., 2007). É também conhecida por desempenhar um importante papel em doenças cardíacas, como hipertrofia, infarto do miocárdio e arritmias (CAI; PAN; LU, 2010; ZORIO et al., 2009). Outro miRNA da mesma família que esta abundantemente expresso é o mmu-miR-206-5p (Tabela 7). Cabe também ressaltar que a maior parte dos miRNAs mais expressos em neonatos ainda não possuem um papel conhecido.

Tabela 8: miRNAs mais diferencialmente expressos em 4 semanas que em neonatos. Os miRNAs estão ordenados de acordo com os seus valores de fold change (Fold Negativo *= downregulated* em Neonato).

miRNA	Log 2 Fold (CPM+1) Change
mmu-miR-96-5p	-3,626
mmu-miR-670-5p	-3,652
mmu-miR-200b-3p	-3,672
mmu-miR-874-5p	-3,774
mmu-miR-182-5p	-3,796
mmu-miR-204-3p	-3,839
mmu-miR-150-5p	-4,347
mmu-miR-22-3p	-4,527
mmu-miR-1968-5p	-4,628
mmu-miR-324-3p	-4,747
mmu-miR-219a-2-3p	-5,202

Um dos miRNA mais diferencialmente expressos em camundongos com 4 semanas de vida é o mmu-miR-96-5p. A variação dentro da região *seed* do miR-96 maduro tem sido associada a progressiva perda auditiva em humanos e camundongos (LEWIS, M. A. et al., 2009). Cinco genes (Aqp5, Celsr2, Myrip, Odf2 e Ryk) de 132 alvos preditos já foram validados para este miRNA (LEWIS, M. A. et al., 2009). Outro miRNA mais abundante em camundongos com 4 semanas de vida que já possui função conhecida é o mmu-miR-150-5p. Este se associa ao complexo RISC afetando o mecanismo de interferência de RNA (GREGORY et al., 2005). O mir-150 atua também na hematopoese regulando genes cujos produtos incentivam a diferenciação de células-tronco para se tornarem megacariócitos em vez de eritrócitos (EDELSTEIN; BRAY, 2011; LU et al., 2008). Outro microRNA mais expresso em camundongos com 4 semanas é o miR-219. Este miRNA está associado à sinalização do receptor NMDA (*N-methyl-D-aspartate receptor*) em seres humanos e sugeriu-se que a desregulação deste miRNA pode levar à expressão de transtornos mentais como a esquizofrenia (KOCERHA et al., 2009).

Cabe ressaltar também que em nossas amostras foram encontrados membros da família miR-200 como o presente na Tabela 8 (mmu-miR-200b-3p). Esta família de miRNAs foi descrita como expressa em epitélio olfatório de camundongos e possui um importante papel na diferenciação de neurônios (CHOI et al., 2008).

4.6 Predição dos alvos para os miRNAs

Como já mencionado, a heterocromatina facultativa deve desempenhar um papel crucial na expressão de genes ORs (ARMELIN-CORREA et al., 2014). A

partir destes dados achamos interessante investigar se havia em nosso banco sequências de miRNAs que modulam a expressão de enzimas envolvidas na formação ou remoção de heterocromatina facultativa bem como investigar a localização dessas enzimas no epitélio olfatório. Um destes alvos é a enzima *Enhancer of zeste homolog* 2 (Ezh2). Ezh2 é um membro do grupo Policombo (PcG) da família de proteínas envolvidas em suprimir a expressão do gene através de remodelação da cromatina (SCHUETTENGRUBER et al., 2007). Por ser uma histona metiltransferase, Ezh2 catalisa a trimetilação da lisina 27 da histona H3 (CAO, R. et al., 2002), que é uma característica do silenciamento gênico (SMITH, 2008). EZH2 é um componente importante do complexo repressor Policombo 2 (PRC2) e é necessária para manter o silenciamento gênico. Outros alvos investigados são o KDM6B (ou Jmjd3) e KDM6A (ou Utx), que possuem função de histona demetilase específica para a lisina 27 da histona H3, ou seja, possuem um papel inverso ao da Ezh2.

Através da utilização do programa TargetScan identificamos possíveis miRNAs que apresentam estas enzimas como alvo (Tabela 9).

O miRNA-101b-5p chamou nossa atenção por ser mais abundantemente expresso em neonatos em relação a amostra de 4 semanas (Tabela 9 em vermelho). Além do miRNA-101b-5p estar presente em nossas análises de predição tendo como alvo Ezh2, artigos relatam a importância do papel que este miRNA exerce sobre a inibição da expressão de Ezh2 (CAO, P. et al., 2010), portanto o miRNA-101b-5p é um bom candidato para futuros ensaios de validação. **Tabela 9: miRNAs que possuem como possível alvo Ezh2.** miRNAs expressos em CPM que possuem como alvo Ezh2 segundo o programa TagetScan.

miRNA ID	4 semanas (CPM)	Neonatos (CPM)
mmu-miR-138-1-3p	0	10,000
mmu-miR-138-5p	14660,075	6318,474
mmu-miR-138-2-3p	194,193	80,01
mmu-miR-101b-5p	3,445	16,842
mmu-miR-26b-5p	6040,225	11508,518
mmu-miR-98-3p	34,447	47,895
mmu-miR-98-5p	5526,108	5089,517
mmu-miR-101b-3p	183,429	845,797
mmu-miR-26a-1-3p	147,690	795,796
mmu-miR-26b-3p	360,398	573,163
mmu-miR-101a-5p	68,463	65,264
mmu-miR-26a-5p	74039,254	53616,241
mmu-miR-26a-2-3p	296,672	459,478
mmu-miR-101a-3p	40809,838	101504,539

É importante destacar que todos os demais miRNAs que possuem como alvo Ezh2 não são expressos na mesma abundância em camundongos neonatos versus camundongos com 4 semanas de vida, sugerindo, portanto, que estes possam desempenhar algum papel nos diferentes estágios de desenvolvimento. Encontramos também 6 miRNAs preditos como alvo de KDM6A e 41 como alvo de KDM6B. É importante ressaltar que estas enzimas são expressas no OE como mostram nossos dados de sequenciamento (Tabela 10). **Tabela 10: Ezh2, Kdm6a e Kdm6b presentes no transcritoma de OE.** Os níveis de expressão de Ezh2, Kdm6a e Kdm6b no OE de camundongos neonatos e com 4 semanas estão expressos em FPKM.

Gene	Neonatos	4 semanas
Ezh2	35,98	33,52
Kdm6a	14,55	15,07
Kdm6b	36,06	23,27

4.7 Localização da expressão de Ezh2 no epitélio olfatório

A seguir, investigamos a localização da expressão de Ezh2 por hibridização *in situ* no epitélio olfatório de camundongo.

4.7.1 Hibridização in situ

Realizando a técnica de hibridização *in situ* com cortes da região do nariz de camundongos C57BL/6 com 4 semanas de vida, conseguimos identificar a localização dos transcritos de Ezh2 no epitélio olfatório (Figura 11). Como controle positivo foi utilizado a sonda para OMP que está presente apenas em neurônios olfatórios maduros e, como esperado marcou apenas este tipo celular. Com a sonda *sense* para Ezh2 não houve hibridização, no entanto na lâmina com a sonda *antisense* para Ezh2 podemos observar que a marcação para Ezh2 em roxo se restringe a uma região específica no epitélio olfatório.





Figura 11: Hibridização *in situ.* A (marcado com sonda *anti-sense* para Ezh2). B (marcado com sonda *sense* para Ezh2). C (marcado com *sonda anti-sense* para OMP).

A Figura 12 mostra que a marcação com a sonda *anti-sense* para Ezh2 se encontra especificamente nas células basais do epitélio olfatório. Então podemos supor que a metilação H3K27 pelo complexo PRC2 ocorre nas células basais que irão se diferenciar em neurônios olfatórios, ou seja, possivelmente a formação da heterocromatina facultativa ocorra nesse estágio.



Figura 12: Imagem ampliada da hibridização *in situ* mostrada na Figura 14. Zoom da região marcada pela sonda Ezh2.

4.8 Análise dos RNAs não codificadores longos

Os lincRNAs expressos no epitélio olfatório em camundongos neonatos e com quatro semanas foram catalogados e analisados (Arquivo suplementar 4). Em nossos arquivos possuímos dados referentes a expressão de 295 lincRNAs. Dentre todos os lincRNAs expressos decidimos investigar se há lincRNAs preferencialmente expressos no epitélio olfatório quando comparado a outros tecidos. Para isto, utilizamos dados de RNA-seq público de cérebro, cerebelo, coração, fígado, rins e testículos e selecionamos os lincRNAs expressos apenas no epitélio olfatório, e não nestes outros tecidos. Dos 295 lincRNAs sequenciados em nossas amostras, 104 foram encontrados apenas no epitélio olfatório. Além das análises dos lincRNAs, achamos interessante verificar quais os dez ncRNAs mais longos presentes em nossas amostras (Tabela 11). Com essa abordagem pretendíamos investigar se há RNAs longos como *Xist,* que possui 17 kb, e desempenha um papel crucial na remodelação da cromatina durante a inativação do cromossomo X.

Tabela 11: Os dez ncRNAs mais longos expressos no epitélio olfatório de camundongos. Os FPKMs para neonatos e 4 semanas estão indicados na terceira e quarta colunas. O tamanho de cada um dos ncRNAs está indicado na última coluna.

Tipo ncRNA	Gene	Neonatos	4 semanas	Tamanho pb
lincRNA	RP24-538B17.1	7,35003	7,66885	4837
Pseudogene	Sowahc	3,84001	3,93706	4478
sense_intronic	B930095G15Rik	10,8047	8,8443	2791
processed_pseudogene	Gm20469	4,8849	4,11681	3838
processed_pseudogen 0	Gm13722	32,827	24,899	3065
lincRNA	2610528B01Rik	4,15821	4,60078	2006
Pseudogene	B930041F14Rik	4,64066	2,50671	2139
lincRNA	Gm11210	0,852789	2,63208	4562
processed_pseudogene	Gm5939	2,22399	1,03861	2873
processed_pseudogene	Gm14738	1,36458	3,04731	2479

O maior ncRNA encontrado em nossas amostras possui aproximadamente 5 kb, ou seja, não há nenhum RNA não codificador tão grande como *Xist*. Cabe ressaltar que foram avaliados apenas ncRNAs já anotados, não excluindo a possibilidade da existência de ncRNA maiores do que estão apresentados na Tabela 11. Dentre os dez ncRNAs mais longos em nossas amostras observamos também a presença 3 lincRNAs que poderão ser analisados juntamente com outros lincRNAs em estudos posteriores.

4.8.1 Análise de ncRNAs que mapeiam na região de genes OR

Verificamos se há RNAs não codificadores que mapeiam em genes OR ou próximos a *loci* de genes OR (Tabela 12).

Gene OR	Posição OR	Tipo ncRNA	Neonatos	4 semana	
OR10J8P upstream		snoRNA	924.556	1407.08	
OR52B1P	overlap	processed	3.33173	2.26812	
OR11H7	upstream	snRNA	0	58.2573	
OR7E12P	downstream	sense intronic	2.39777	0.989985	
OR2H1	downstream	miRNA	13.7286	7.04888	
OR2AE1	upstream	snoRNA	131234	114385	

Tabela 12: RNAs não codificadores que mapeiam próximos a loci de genes OR.

Podemos observar na Tabela 12, que em nossas bibliotecas não há praticamente ncRNAs que mapeiam próximos a *loci* de genes de receptores olfatórios, havendo apenas 6 ncRNAs presentes nessas regiões sendo que dentre estes 2 são snoRNAs, 1 snRNA e 1 pseudogene. Há um gene para miRNA localizado a *downstream* ao receptor olfatório 2H1 e um ncRNA intrônico a *downstream* do receptor olfatório 7E12P, o que a princípio não aparenta ser muito relevante uma vez que *loci* dos genes que codificam os receptores olfatórios estão distribuídos em praticamente todos os cromossomos. Se houvessem ncRNAs que agissem em *cis* nos *loci* OR, esperaríamos encontrar ncRNAs mais amplamente distribuídos nesses *loci*.

4.8.2 Validação dos RNAs não codificadores longos

Para validação desses RNAs não codificadores longos, decidimos escolher 10 dos 104 lincRNAs que são preferencialmente expressos no epitélio olfatório quando comparado a dados de RNA-seq público de diversos tecidos (cérebro, cerebelo, coração, fígado, rins e testículos) (Arquivo suplementar 4).

Para a seleção adotamos os seguintes critérios:

- I. LincRNAs expressos mais abundantemente em neonatos quando comparado com camundongos com 4 semanas de vida.
- II. LincRNAs que possuem valores de FPKM maior ou igual a 6.
- III. Somente lincRNAs anotados presentes no banco de dados ENSEMBL.

Os lincRNAs que apresentaram essas características estão representados na Tabela 13.

Dentre os candidatos escolhidos para validação, apenas o lincRNA Gm13499 não atende os critérios estabelecidos. Porém este lincRNA apresentou um padrão de expressão interessante, possuindo valor de FPKM muito mais alto em camundongos com 4 semanas de vida que em neonatos.

Tabela 13: Lista dos 20 lincRNAs preferencialmente expressos no epitélio olfatório,que obedecem aos critérios estabelecidos.Os lincRNAs estão separados segundovalores de FPKM em camundongos neonatos com exceção do último presente na tabela.

Gene	Locus	FPKM Neonatos	FPKM 4 semanas
AC093350.1	Chr3:96328424-96422897	48601,8	36489,3
AC121792.1	Chr3:96328424-96422897	5403,4	8814,1
Gm14074	Chr2:135648906-136014604	42,4609	35,3581
AC131744.1	Chr3:126921606-127409008	38,1353	45,1907
AC124108.1	Chr1:132191435-132257381	34,7526	6,71802
Gm12336	Chr11:75409768-75423340	29,9472	24,468
AC160538.1	Chr15:89211558-89246394	15,9735	18,8969
Gm14168	Chr2:156547583-156577457	15,1637	11,5465
AC127595.1	Chr7:66109514-66173789	13,9235	9,55044
AC102317.1	Chr18:58210100-58221885	12,622	0,269009
AC140071.1	Chr3:97203581-97297851	11,4152	12,4764
AL732403.1	Chr2:5137748-5230878	11,0353	6,2341
ltih5l-ps	ChrX:150826971-150857136	10,9521	0,356144
AC148089.1	Chr8:19789797-19798548	9,53059	10,2826
Gm15713	Chr16:42875880-43642606	7,91438	5,21452
AC127314.1	Chr18:54888044-55079414	6,67094	5,88383
8430419K02Rik	Chr11:18845850-18874309	6,6455	1,39929
E330012B07Rik	Chr6:147180243-147208285	6,48943	6,9646
2700069118Rik	Chr3:5177585-5415855	6,01844	1,41534
AC141473.1	Chr3:83766320-83789956	5,97785	1,85623
Gm13499	Chr2:51363304-51480597	1,92126	17,7975

Utilizando o banco de dados *Ensembl*, analisamos a anotação destes lincRNAs. Verificamos também utilizando esse banco de dados em quais tecidos estes lincRNAs já foram descritos. A partir dessas análises selecionamos 10 lincRNAs para nossas validações (Tabela 14).

Tabela 14: Seleção dos 10 lincRNAs utilizados na validação.Os 10 lincRNAsselecionados de acordo com anotação no Ensembl.

Nomenclatura	Gene	Ensembl: nome / anotação	Tecido
LincRNA_1	Gm14074	ENSMUSG00000146980 / novel lincRNA	Baço
LincRNA_2	AC124108.1	ENSMUSG0000097540 / known lincRNA.	Testículo
LincRNA_3	Gm12336	ENSMUSG0000085359 / putative lincRNA	Fibroblastos
LincRNA_4	Gm14168	ENSMUSG000000144818 / novel lincRNA	Tumor de Fígado, Cérebro
LincRNA_5	AC102317.1	ENSMUSG00000097618 / known lincRNA	Epitélio Olfatório
LincRNA_6	AC140071.1	ENSMUSG00000097143 / known lincRNA	Glândula Mamária, Timo
LincRNA_7	Gm15713	ENSMUSG0000085499 / Novel lincRNA,	Timo, Rim
LincRNA_8	E330012B07Rik	ENSMUSG00000151979 / Known lincRNA	Ovários, Embrião
LincRNA_9	2700069118Rik	ENSMUSG00000128815/ Novel antisense	Retina, Medula Espinal, Embrião
LincRNA_10	Gm13499	ENSMUSG0000086523 / putative lincRNA	Testículo

Obs: Decidimos nomear de forma crescente os lincRNAs segundo seus respectivos FPKM para facilitar a citação dos mesmos no texto (lincRNA_1 a lincRNA_10).

4.8.3 Amplificação dos lincRNAs

Para confirmar a expressão dos lincRNAs no epitélio olfatório, realizamos reações de PCR utilizando cDNA de epitélio olfatório de camundongos neonatos e camundongos com 4 semanas (Figura 16).



Figura 12: Amplificação dos lincRNAs a partir de cDNA de epitélio olfatório. Amplificação a partir de cDNA da região correspondente aos lincRNAs. O marcador de peso molecular (*Low DNA Mass™ Ladder – Invitrogen*) pode ser observado na primeira canaleta da figura. Na última canaleta temos o controle GAPDH amplificado a partir de cDNA de camundongos com 4 semanas.

Na Figura 12 observamos que os *primers* para os lincRNAs 2 e 3 são aparentemente inespecíficos, gerando diversas bandas. Porém a amplificação com os primers para os lincRNAs 1, 4 e 5 geraram bandas únicas do tamanho esperado (lincRNA 1: 311 pb, lincRNA 4: 578 pb e lincRNA 5: 404 pb).



Figura 13: Amplificação dos lincRNAs a partir de cDNA. Amplificação a partir de cDNA da região correspondente aos lincRNAs. O marcador de peso molecular (*Low DNA Mass™ Ladder - Invitrogen*) pode ser observado na primeira canaleta da figura. Na última canaleta temos o controle GAPDH amplificado a partir de cDNA de camundongos com 4 semanas.

A amplificação dos lincRNAs 8 e 9 apresentaram bandas com o tamanho esperado (lincRNA 8: 422 pb e lincRNA 9: 566 pb) (Figura 13). Decidimos então realizar novamente um PCR apenas com os primers que geraram bandas específicas e aumentamos o número de ciclos no PCR para os lincRNAs que apresentaram banda fracas na amplificação anterior (Figura 14 B). Como se pode observar na Figura 14 A, os primers *forward* e *reverse* utilizados em cada caso anelam-se a exons que são separados por introns muito grandes. Apesar disto, foi possível amplificar estes 5 lincRNAs, validando, portanto, a sua expressão no epitélio olfatório.





4.8.4 Sequenciamento dos lincRNAs

Para confirmar a identidade dos 5 lincRNAs amplificados em nossos ensaios, os produtos de PCR (Figura 14 B) foram purificados e sequenciados (sequenciamento automatizado método Sanger). Os resultados do sequenciamento foram analisados utilizando a ferramenta BLAT disponível no site *UCSC Genome Browser.*

Na Figura 15 temos o alinhamento da sequência do lincRNA 1. Como é possível visualizar no alinhamento sobre o DNA genômico, os primers *forward* e reverso de todos os lincRNAs utilizados em nossas reações de PCR foram construídos para se parear com éxons diferentes. Deste modo garantimos que os produtos de PCR sejam sintetizados exclusivamente a partir do cDNA, e não sejam fruto de contaminação por DNA genômico. Como esperado, na Figura 16 é possível observar que nossa sequência se alinha perfeitamente com o gene correspondente ao lincRNA 1 (ENSMUSG000000146980). Interessantemente os exons 3 e 4 do lincRNA 1 são conservados em diferentes espécies, incluído a humana.

mment of YourSeq and chr2:135648939-13565199	Genomic chr2 (reverse strand):
A YourSeq caagat gaaatgotoc cggccgccat gccgrgggs ttouT00440C 59 stuara AGLACTCT CTAGAGAGT GAACACTRO GATCCCCGT 100 TCTGC ATGCCACAAA ATCATTGCAT GCTTTCTTG GOSCAGOGAC 150 GTTCAT CTOCTTGCAA ATCATTGCAT GCTTCTTG GOSCAGOGAC 250 AAAATC TUCCUTCAA TTTGCTGCC CACCCCCCA AATGCTOGAA 250 AAAATC TUCCUTCAA TTTGCTGCC CACCCCCCA AATGCTOGAA 250 AAAATC TUCCUTCAA CCUBOCTTGC AAAGTCTCACA ATACGOTGA 250 AAAATC GOCTGAAAA CCUBOCTTGC AAAGTCTCACA ATACGOTGA 250 AAAATC GOCTGAAAA CCUBOCTTGC AAAGTCTCACA ATACGOTGA 250 AAAATC GOCTGAAAAA CCUBOCTTGC AAAGTCTCACA ATACGOTGA 250 AAAATC GOCTGAAAAA CCUBOCTTGC AAAGTCTCACA ATACGOTGA 250 AAAATC TUCCUTCAA CCUBOCTTGC AAAGTCTCACA ATACGOTGA 250 AAAATC GOCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	<pre>aagcaagaca atgattttgg gctcagagtt taagtgttgg ccaaagtaaa 135652046 aatataacct ttttattta ttaattatct atgtgtaata tcccatcaa 135651996 GGTGGAGCTG ATGTGATAAG CACTCTCT AGAGAGTGA ACACTTAGGA 135651946 TCCCCGTTGC TCTCTGCCAT GCCACAAAAT CATTGCATGC TTTCTTGAGG 135651846 ttaggacaaa gtggggggctt aatgaatatt tgctgaacca acaaacaaac 135651796 aaacaacaaa aaccccattg gctcggaat ctactatgtg gctgaaacca agaaccaacaa gtcattcca cattatatt gggtcacta ttagtagta 135651646 aatacagttt cgctaacct tctaggtgga attaagcctc ctatgttc 13565196 ttacagcctg gaaccattc ctattttata gcacttg gactagatca 13565196 ttacagcctg gaacaattc ctattttata gcacttg gatagattag 135651446 tcaagatta agaatagt caccttatt tgacattg 13565146 gacaaagtaa aagaatagc aacctatatt ttacagtcg tagaagtg 135651446 tcaagatta attatacc cttatttat gcacttgc taggattg 135651446 tcaagatta attatacc ttatttt ttacagtcg taggatcat 135651496 gacaaagtaa aagaatagc aaccttatt ttacagtcg taggatgg 135651446 tcaaggtatt attatacc ttatttt ttacgtcg taggatcg 135651446 gactcatag cactagagg ctttttt ttttt ttatcagtcg taggattg 135651446 gactattg cactagagg tatttttt tctctctt tctcttgtta 135651346 gactatag cactagagg cttttttt ttttt ttatcagtcg taggattg 135651446 aaggttttt accttgtta gctattacc gaggtaact aactttattg 135651296 / // gaggctgttc tgatagtttt tgccttaaag atgattaatt tttcaaaatc 135649496 aggtatttta cactagatg ctctttttt tcttcctct tcctctgtta 135649496 tttattttt tcttaatta aggttagt ttattttg tgttggcacc 135649496 ttgacatttt agccaggg cagtatcaa tttagcgt tgggtacca cactcggtc 135649496 tgacatttt agccaggat ctattttg tctgggact tacGTCAAT 135649496 dGCGGCTGCAA AGTctgcatt taaaatgtg tattattat taactaga 135649496 dtttattg tcaacagta aacctagga atattaac tttatatag 135649496 dtttatag tcaacagta aacctagga atattaac ttagatca 135649496 dCTGGAA AGTctgcatt taaaatgtg tattattat taactaga 135649496 dCTGAAAga aacacagcc tgcaaaaca ccgttcatt ttatattt 135648946 dCTGAAtga aacacagcc tgcaaaaca ccgttactt tttatttt 135648946 dCTGAAtga aacacagcc tgcaaaaca ccgttactt tttatatttt 135648946 dCTGAAtga aacacagcc tgcaaaaca ccgttactt tttatttt 135648946 dCTGAAtga aacacagcc tgcaaaacac ccgttactt tttt</pre>

Figura 15: Alinhamento da sequência do lincRNA 1. Canto superior esquerdo: coordenada genômica do sequenciamento. Lado esquerdo: alinhamento com o cDNA. Lado direito: alinhamento com o genoma. Azul escuro: bases pareadas. Azul claro: *gaps*.



Figura 16: Coordenada genômica do lincRNA 1. A nossa sequência depositada no Blat está nomeada como *YourSeq.* Em vermelho: Gene predito no *Ensembl.* Parte inferior: Conservação em diferentes espécies.

Na Figura 17 temos o alinhamento da sequência do lincRNA 4 com o cDNA e

DNA genômico na coordenada esperada. É também possível observar que nossa

sequência se alinha com o gene correspondente ao lincRNA 4 (ENSMUSG000000144818). Além disso, o lincRNA 4 possui o exon 2 conservado em outras espécies, incluído a humana (Figura 18).

Alignment of YourSeq and chr2:156571793-156575839		Genomic chr2 (reverse strand):		
cDNA YourSeq		tccaggetge acteccagas getgetetgt gteettagge tgtgaacaat atgaagggae acaggettg tgtgaagtee ttgstggetg aggaaaggae MADCTGGAAG TGTGGACAGT GTUATTGGAE TTAAAATGTA AGAAGAAA AAAAAAATGA AGATCASACE TGAATGTGAG AGGCAGCEC GOOGETGOOG	156575890 156575840 156575790 156575740	
asggccgase atgastgctc cggccgccat ggccgcgggs tt/A0CT00	50	ADDECADDSC TETDSAAACA AATQEABAAD DECACTADSE AAAQOCADAD	156575690	
ACTOROACA CTOTUATTOD ACTIAAAATO TAAGAAGAAG AAAAAAAAA	100	CAAAAAISTAD OTTAATOTOC AFCTLATCLT AADDIGACAD TODATOTOCA	156575940	
CAMERICAGA CETUADIUTO AGAGGEAGEE CIGEDELIO, ULADOLAR	100	SULATIINAN ULANIUUUAA ANLUNUMILA HUHUUUUAA CICTIBIUUL	155515590	
AGOTTAATOT CCATCTCATC CTAADGTGAC AGTBGATGTC CABBCATTT	258	ASSAUCTORY ATTENTATION ACAUTATION SATATTICAS STREETS	186676400	
AAGCAATEGE AAAACCAGAT CAAGAGGGGA ATCTCTTETE ECATETEAC	300	ETGATOGATO COTTUATOS CARCACUTOS SASACADOCA GABACAGOTO	156575440	
GAAGGGAAGA AAACATGCAT TTCCAGTGTC ACAAAGCTGC AAAGGAACTI	358	BATCTCTOTE AGTTTCABEC CAECCTEDTC TACATABCAA OTTTEABACC	156575390	
OCATTCATAT TTACADTATT DAGATATITC ADDITUDECT TODIOATOG	400	ACCTTODOCT TCTTACTIGAE AT/ctstage teastagets estagetage	156575340	
TECCTITAAT CCCRECACET BEEAGACAGE CAGAGACAGE TEGATCTCT	458	tanotanato autoaccean concepted energenene energenene	156575290	
TOAGTTTCAG GCCAGCCTGG TCTACATAGC AAGTTTGAGA CCAGCTTGG	500	Estattants tatatatats tatatutats tatatasatt autiones	156575240 //	
CTTCTTAGEG AGATICTOTC COSTITUOTE TECTIGATET CTTECCTC	558			
OTCTOGADEE AAGAATTEEC TEETEETEETE TETOOTECAO TOTEECCAC	000	/ saggicagge aggatgetge astasseste tetregeste teagtateet	156572240	
COCCAMMOTE TOBTOTOWAN ANTEACEMENT property progetting	200	ccagggtgas tettgtcagg getgeettgs etteattggt gtetsaacat	156572190	
catatigging agriccomo pretogente catagortes staticiati	958	staactitics caggigatat acatticigi actaacatat giaiggaata	156572140	
maattetta torgetenen attoracaca acatargage comanget	800	tgungtgaac ecasticagt tanggatgat tototgotot otgitosage	156572090	
andtetanne cotempter constrants anotancica cattantte	858	angtangget trigittite itiaianatt agittggact gialaaccaa	156572040	
gttgogotes otgocogott tocagtoggs associtgtog tgoagotgo	900	accasegtas atgittatet atgecattig agitgacagi gigagicetg	156571998	
ttaatggast cggcccacgc cgcggggas gasggcgggs tttggtccg	950	agggttitgt agortoacat agorgaggto tggotgaaga cattitoato	156571948	
tsetttigtt g		TINGENELSE COURSENS STRACTORE SATURATE ACTIVITY	156571648	
		TOTOTATET OSTECASTET COTCACASE CASASTETAS TOTALITAT	156571798	
		parchagage compositet ecclectare teasartese causigases	156571740	
		cterateect ecteecctra esacctritet critacetas caestre		

Figura 17: Alinhamento da sequência do lincRNA 4. Canto superior esquerdo: coordenada genômica do sequenciamento. Lado esquerdo: alinhamento com o cDNA. Lado direito: alinhamento com o genoma. Azul escuro: bases pareadas. Azul claro: *gaps*.



Figura 18: Coordenada genômica do lincRNA 4. A nossa sequência depositada no Blat está nomeada como *YourSeq.* Em vermelho: Gene predito no *Ensembl.* Parte inferior: Conservação em diferentes espécies.

Na Figura 19 temos o alinhamento da sequência e a coordenada genômica do lincRNA 5 que está de acordo com o esperado. É também possível observar que nossa sequência se alinha com o gene correspondente ao lincRNA 5 (Gm26507) Além disso o lincRNA 5 possui seus 2 exons conservados em outras espécies, incluído a humana (Figura 20).

Figura 19: Alinhamento da sequência do lincRNA 5. Canto superior esquerdo: coordenada genômica do sequenciamento. Lado esquerdo: alinhamento com o cDNA. Lado direito: alinhamento com o genoma. Azul escuro: bases pareadas. Azul claro: *gaps*.



Figura 20: Coordenadas genômica do lincRNA 5. A sequência nucleotídica do lincRNA 5 validado foi submetida a ferramenta Blat no *UCSC Genome Browser*.

O sequenciamento do lincRNA 8 esta de acordo com o esperado. A cordenada genômica e gene Ensambl (ENSMUSG000000151979) estão alinhados corretamente (Figura 21 e 22).

Alignment of YourSeq and chr6:147180274-147206389		Genomic chr6 (reverse strand):
cDNA YourSeq tcttggtacg agctcggatc cactagtaac ggccgccagt gtgctggaat tcggctiGCT TCTGAGTCac aaGgGAgGTG AGGAGTCCCA GGCATCAGAA GCCAATGTTC ATTCttsATG CCCTGGGATA AACTCTCTTT GTTGGTGACA TCTGTTCTGT AAACTACATT AATAGAAATA TGCATCTTGA GTCCTGACGA CAGCTTAGTG GTTACATTG CTGAATTATC CTCCAAGGAG CTCAAGGA CCACTGCTG GTTACATTG CTGAATTATC CTCCAAGGAG CTCAAGGA CCACTGCTG GTTACATTG CTGAATATC CTCCAAGGAC CCAAAGGATC CCATGCCCTT TCTGGCTCC CCCGAGCCTC TGCATACCTC TTTGGTTTC TAAAAATGTC TATTATAAGA AATGACAAG AATGCTTCTG TCCACTTCTG GAACTCCTTA ATGAACCATG GCAGCTTTCT TGTGACAGCA GCCCCAGCG TAGCTGAGGC AGGAACTCAA aggccgaatt ctgcaggat ccctcacact ggcggccgt cgagcaggca ctctagaggc cccattcgcc ctatagtgag tcgtattaca attcactgc gtcgtttta caacgtcgtg actgggaaaa ccctggcgta tagcgaaaga ggcccgcac gtagcgcctt cccaacagt gcggccgccg atggcagaa ggccgccc gtagcgcct gataggcc ctagcggctg atggcagaa ggccgccc gtagcggca cgttaggcc ctagcgccg ctttccccg gtcaaggac gctacattt tccgacaacg tatccgccg ctttccccg gtcaaggtc taaatcgggg ggggcccttc cccgtttagg gggtttcc gccgagaa gtcacattt tccgacaag tatccgccg ctttcccc gtcaaggtc taaatcgggg ggggccctt cccgtttagg gggtttacc gccgagaac gctacattt	50 100 250 200 350 400 550 600 550 600 700 750 800 850 900	actgttttaa ggagctgtgg gtacaaagga gtttggggat gtagaacta 147206440 actggaggaa gtgggctggt aagagctcat tagggatgag cgctcagctc 147206390 TGCTTCTGAG TCtactGGA tGTGAGGAGT CCCAGGCATC AGAAGCCAAT 147206390 TGCTATTCgG ATGCCCTGGG ATAAACTCT TTGTTGGTG ACATCTGTC 147206290 TGTAAACTAC ATTAATAGAA ATATGCATC TTGATGGTG ACATCTGTC 147206190 TGTAAACTAC ATTAATAGAA ATATGCATCT TTGATGGTG ACATCGTC GCGTGTACAG TTGCTAGATT ATCCTCCAAA GGACTCAGTC CCAGCACCCA 147206190 TGTTGGGCAG CTCAGTTACC TGAAGCCCCA GCACCAAAAG ATCCCCATGCC 47206190 CTCTTCTGGT CTCCCCGAGC CTCTGCATAC ctgtgagata aacacagta 147206090 cacaatattt tttaattaag atcgatcaga gtcaccttaa agcatgaaac 147206090 ttatcatccc cacagctgtt aaactatta cctgctgtga ggagatgg 147205940 ttgaaggcgag aaattcttc ctgccttga atcagacac acagagtaca 147205940 tgaaggtgg cattcaata gaaccaggag gtccaactca acagagtaca 147205840 // aagttggttg ctgggggctg ctggctctgt actagaatca actactctgt 147185540 taaatattct cagcactata ctgctccaa gtcggcttca gtagataca 147205840 // aaaatgtgt atcagtagt aattattac CTCTTGGTT TCAAAAAAT 147185540 TTAATGAACC ATGGCAGCTT TTGTGACA GCCtgaagga acaggtca 147185540 TTAATGAACC ATGGCAGCTT TTGTGACA GCctgaagga acaggtca 147185540 TTAATGAACC ATGGCAGCTT TTGTGACA GCctgaagga caaggtca 147185540 TTAATGAACC ATGGCAGCTT TTGTGACA GCctgaagga acaggtca 147185340 taattcagt aagaagagg acaaatttgg aaaccaagta tagtgggtt 147185340 tacctgtaat ctcaeggag acaaatttgg aaaccaagta tagtggtt 147185340 tacctgtaat ctcaegagt tagegga acaagtta gttgggcat 147185340 tacctgtaat ctcaegagt tagegga acaagtta geccaggaga 147185340 tacctgtaat ctcaegagt tageggga acaagttag gaacaagta gttggggg 147185340 tecctgtaat ctcaegagt tagegga acaagttag aaccaagta tagtggttt 147185340 tecctgtaat ctcaegagt tagegga acaagttag agccaggaga 147185340
	,	/ aggacggaag aaaatacttt gttgcttgtg atctcttagg ggcctgacct 147180390 ggtttgcttc tctgtaactg tgataaaccc tgaccaaaaa gcacattagg 147180340 cagaaaaagg ataattttca cttacAGGTT TCAGTCTATC CCAGGGTAGC 147180290 TGAGGCAGGA ACTCAAggca ggagaacagt tgcttactgc cttcctttcc 147180240 ctgacttgct cagctacctt tcttatacaa cccaggttac ctgcccagag 147180190 atgactccgc ctaccg

Figura 21: Alinhamento da sequência do lincRNA 8. Canto superior esquerdo: coordenada genômica do sequenciamento. Lado esquerdo: alinhamento com o cDNA. Lado direito: alinhamento com o genoma. Azul escuro: bases pareadas. Azul claro: *gaps*.



Figura 22: Coordenada genômica do lincRNA 8. A nossa sequência depositada no Blat está nomeada como *YourSeq.* Em vermelho: Gene predito no *Ensembl.* Parte inferior: Conservação em diferentes espécies.

Alignment of YourSeq and chr3:5178026-5220811		Genomic chr3 (reverse strand):
CDNA YourSeq	50 100 150 200 250 350 500 550 550 550 550 550 550 550 5	<pre>tgsctgtage tgtptgtgt tgtgtgtgt tgtgtgtgt tgtgtgtttg 5220862 gegggggggt gtccetgtt gtgstgtgt tgtgtgtgt tgtgtgtttg 5220812 stccetgest gtccetgtt gtgstgtgt gtgtgtgt tgtgtgtttg 5220812 ctrasacca colored colo</pre>

Figura 23: Alinhamento da sequência do lincRNA 9. Canto superior esquerdo: coordenada genômica do sequenciamento. Lado esquerdo: alinhamento com o cDNA. Lado direito: alinhamento com o genoma. Azul escuro: bases pareadas. Azul claro: *gaps*.



Figura 24: Coordenada genômica do lincRNA 9. A nossa sequência depositada no Blat está nomeada como *YourSeq.* Em vermelho: Gene predito no *Ensembl.* Parte inferior: Conservação em diferentes espécies.
Nas Figuras 23 e 24 observamos o alinhamento e a coordenada genômica do sequenciamento do lincRNA 9. O gene ENSMUST00000128815 que corresponde ao lincRNA 9 está alinhado corretamente ao sequenciado. Além disso, observamos conservação em um modelo mais distante dos mamíferos e roedores, o Zebra Fish, indicando que este deva desempenhar um papel crucial em diferentes espécies, mantendo-se conservado ao longo da evolução.

4.8.5 Distribuição tecidual da expressão dos lincRNAs

Para validar se estes lincRNAs são de fato preferencialmente expressos no epitélio olfatório, realizamos ensaios de PCR com cDNA proveniente de 10 tecidos de camundongo diferentes: cérebro, testículo, baço, olhos, língua, pulmão, coração, tecido adiposo e rim. Para certificar-se sobre a qualidade dos cDNA oriundos dos 10 tecidos diferentes utilizados em nossos ensaios, amplificamos o gene GAPDH como controle. A amplificação de GAPDH por PCR foi realizada com sucesso em todos os tecidos indicando a boa qualidade do cDNA utilizado em nossos ensaios.

Garantida a qualidade do cDNA dos diversos tecidos, realizamos então o mesmo ensaio separadamente para cada lincRNA utilizando cDNAs oriundos de camundongos neonatos (Figura 25 A) e camundongos com 4 semanas de vida (Figura 25 B). Observamos que a amplificação do lincRNA 1 ocorre apenas no epitélio olfatório em camundongos com 4 semanas de vida sugerindo que este lincRNA está sendo expresso exclusivamente nesse tecido quando comparado aos demais nesse estágio de desenvolvimento. Porém em camundongos neonatos o padrão de expressão tecidual do lincRNA 1 não se restringe apenas ao epitélio olfatório. O mesmo padrão de expressão foi observado para o lincRNA 4. Em ambos

os casos é possível visualizar também bandas fracas em alguns dos diversos tecidos na amplificação em camundongos com 4 semanas de vida. Outra observação importante em camundongos com 4 semanas de vida é que a expressão dos lincRNAs 5 e 9 não se restringem apenas ao epitélio olfatório já que houve também amplificação desses lincRNAs utilizando cDNA oriundo do cérebro. Vale destacar também que o lincRNA 9 apresenta uma forte banda quando amplificado a partir de cDNA de olhos em camundongos com 4 semanas de vida. Ao contrário do padrão de expressão observado em camundongos com 4 semanas de vida onde os lincRNAs estudados são expressos preferencialmente no epitélio olfatório, em camundongos neonatos a expressão desses mesmos lincRNAs não se restringe apenas ao epitélio olfatório com exceção do lincRNA 8 que é expresso apenas no OE em ambos os estágios de desenvolvimento.





Figura 25: Distribuição tecidual da expressão dos lincRNAs no MOE. (A) Reações de RT-PCR usando cDNA de dez tecidos diferentes de camundongos neonatos. (B) Reações de RT-PCR usando cDNA de dez tecidos diferentes de camundongos de 4 semanas de idade. Para cada reação foi incluída uma reação de controle negativo (sem adição de cDNA) e um controle positivo (GAPDH).

Juntos esses dados sugerem que os lincRNAs selecionados para validação são expressos em diferentes tecidos em camundongos neonatos indicando que estes ncRNAs possam ter um papel em estágios iniciais de desenvolvimento, porém a expressão destes é mantida apenas no epitélio olfatório em estágios mais tardios de desenvolvimento.

A partir desses resultados decidimos então realizar ensaios de hibridização *in situ* dos lincRNAs validados para verificar em quais células do epitélio olfatório estes são expressos.

4.8.6 Localização da expressão dos lincRNAs no epitélio olfatório

Após confirmar a clonagem por sequenciamento, linearizamos os plasmídeos contendo os lincRNAs desejados com enzimas de restrição específicas e as utilizamos para síntese das sondas de RNA *sense* e *antisense*. Preparamos lâminas com cortes da região do nariz de camundongos C57BL/6 com 3 dias de vida (neonatos) e com 4 semanas de vida. Realizamos então a técnica de hibridização *in situ* do epitélio olfatório. Como controle positivo utilizamos a sonda para OMP que está presente apenas em neurônios olfatórios maduros e, como esperado marcou apenas este tipo celular (Figura 30 C).



Figura 26: Detecção do lincRNA 1 expresso no epitélio olfatório por hibridização *in situ* para o lincRNA 1 em camundongos com 4 semanas de vida (4W). (B) Hibridização *in situ* para o lincRNA 1 em camundongos neonatos (P3). (C) Controle negativo.



Figura 27: Detecção do lincRNA 4 expresso no epitélio olfatório por hibridização *in situ*. (A) Hibridização *in situ* para o lincRNA 4 em camundongos com 4 semanas de vida (4W). (B) Hibridização *in situ* para o lincRNA 4 em camundongos neonatos (P3). (C) Controle negativo.



Figura 28: Detecção do lincRNA 5 expresso no epitélio olfatório por hibridização *in situ*. (A) Hibridização *in situ* para o lincRNA 5 em camundongos com 4 semanas de vida (4W). (B) Hibridização *in situ* para o lincRNA 5 em camundongos neonatos (P3). (C) Controle negativo.



Figura 29: Detecção do lincRNA 8 expresso no epitélio olfatório por hibridização *in situ*. (A) Hibridização *in situ* para o lincRNA 8 em camundongos com 4 semanas de vida (4W). (B) Hibridização *in situ* para o lincRNA 8 em camundongos neonatos (P3). (C) Controle negativo.



Figura 30: Detecção do lincRNA 9 expresso no epitélio olfatório por hibridização *in situ*. (A) Hibridização *in situ* para o lincRNA 9 em camundongos com 4 semanas de vida (4W). (B) Controle negativo (4W). (C) Controle positivo (OMP).



Figura 31: Detecção do lincRNA 4 expresso no VNO e bulbo olfatório. (A) Hibridização *in situ* do VNO para o lincRNA 4 em camundongos com 4 semanas de vida (4W). (B) Hibridização *in situ* do VNO para o lincRNA 4 em camundongos neonatos (P3). (C) Hibridização *in situ* do bulbo olfatório para o lincRNA 4 em camundongos neonatos (P3).

Como descrito na Figura 4, o epitélio olfatório apresenta três camadas de células principais: uma camada que contém as células de suporte, localizada na região apical do epitélio, a camada central contendo os neurônios olfatórios maduros e a camada basal, contendo as células basais que são capazes de se diferenciar em neurônios olfatórios. Não observamos em nossas lâminas a marcação em nenhuma dessas camadas na hibridização *in situ* feita para os lincRNAs 1 e 8 (Figuras 26 e 29). Como esperado também não houve marcação da sonda *sense* que utilizamos como controle negativo nesse ensaio (Figuras 26 e 29 C). Para os lincRNAs 5 e 9 observamos uma forte marcação por todo o OE em todas as lâminas incluindo o controle negativo (Figuras 29 e 30), indicando que estas sondas não são boas.

É importante ressaltar que os IncRNAs podem estar distribuídos em vários compartimentos celulares. Alguns IncRNAs se acumulam em diferentes níveis no núcleo versus citoplasma enquanto outros são igualmente distribuídos em ambos compartimentos (CABILI et al., 2015; ULITSKY; BARTEL, 2013). É possível que a hibridização *in situ* empregada para determinar a localização dos lincRNAs não seja suficientemente sensível para detectá-los em compartimentos subnucleares, e diferentes abordagens devem ser utilizadas para analisar esses lincRNAs. Esses dados poderiam explicar a não visualização de sinal quando utilizamos a sonda *antisense* para os lincRNAs 1 e 8.

Outra observação importante é a respeito do lincRNA 9 que é um transcrito *antisense*. Na fita *sense* desta mesma coordenada existe a presença de 2 transcritos anotados como codificadores de proteína (*Zinc Finger Homeobox* 4 (Zfhx4) e Gm10748 como pode-se observar na Figura 24. É possível observar também que o exon 3 do lincRNA 9 está alinhado com o intron 1 de Zfhx4. Isso poderia ocasionar a hibridização da sonda *sense* do lincRNA 9 ao transcrito Zfhx4 ainda não processado ou somente ao intron já excisado deste transcrito codificador de proteína. Isso explicaria a forte marcação no OE pela sonda *sense* do lincRNA 9 (Figura 30 B).

Na hibridização *in situ* utilizando a sonda *antisense* para o lincRNA 4, obtivemos uma forte marcação em todas as camadas do OE, VNO e bulbo olfatório de camundongos neonatos e com 4 semanas de vida (Figura 27 A e B) e nenhuma marcação utilizando a sonda *sense* (Figura 27 C). Na Figura 31 podemos observar também que a sonda *antisense* do lincRNA 4 hibridizou fortemente no VNO e no bulbo olfatório.

4.9 Expressão dos lincRNAs no VNO

Como o VNO apresenta uma organização e mecanismos moleculares similares aos do OE, achamos interessante avaliar por RT-PCR se os lincRNAs estudados neste trabalho são expressos também nesse tecido (Figura 32). Para nossa surpresa os resultados mostram que todos os lincRNAs, exceto o lincRNA *9*, são expressos também no VNO, indicando mais uma vez que estes lincRNAs devam desempenhar um papel fundamental no desenvolvimento e ou manutenção de ambos os tecidos olfatórios sensoriais, MOE e VNO.



Figura 32: Amplificação dos lincRNAs a partir de cDNA de VNO. Amplificação a partir de cDNA de VNO da região correspondente aos lincRNAs. O marcador de peso molecular (*Low DNA Mass™ Ladder - Invitrogen*) pode ser observado na primeira canaleta de cada figura. Para cada reação adicionamos um controle negativo (mix de reação sem a presença de cDNA).

5 DISCUSSÃO

Até hoje, o repertório de ncRNAs presentes no epitélio olfatório, assim como suas possíveis funções, é pouco conhecido. Neste trabalho, identificamos e catalogamos os ncRNAs já anotados presentes no genoma de camundongos que são expressos no epitélio olfatório de camundongos neonatos e com 4 semanas de vida. Muitos destes, apesar de já anotados como ncRNAs, ainda não foram descritos na literatura como expressos no MOE. A análise realizada nestes dois estágios de desenvolvimento nos permitiu identificar ncRNAs que possam estar envolvidos no desenvolvimento do OE, e na regulação da expressão dos genes OR. Estes genes começam a ser expressos no período embrionário E12/E13, porém é apenas após o nascimento que um grande número de ORs é aparentemente detectado, e em menos de 2 semanas após o nascimento o pico de expressão continua alto até o animal ter aproximadamente dois meses de idade (ZHANG, XINMIN et al., 2004). Em nossa abordagem só analisamos os ncRNAs anotados, porém possivelmente existem muitos outros ncRNAs ainda não anotados em nossos dados de sequenciamento, já que existem poucos sequenciamentos em larga escala de epitélio olfatório disponíveis nos bancos de dados de maneira geral. Análises futuras dos nossos transcritomas, que não se limitem aos genes já anotados, poderão identificar ncRNAs adicionais, que apresentem papéis específicos para este tecido.

Encontramos também um grande número de miRNAs expressos em nossas amostras, muitos deles ainda sem função conhecida. Na amostra de neonatos, foram encontrados 957 miRNAs, e na de 4 semanas 827 miRNAs. Considerando-se que atualmente o banco de miRNAs miRBase contém 1915 miRNAs maduros para *Mus musculus* (http://www.mirbase.org/cgi-bin/browse.pl?org=mmu), podemos dizer que por volta de 43 a 49% dos miRNAs conhecidos em camundongos foram identificados nas nossas amostras. Além disto, dentre estes miRNAs identificamos vários que apresentaram expressão diferencial entre os dois estágios de desenvolvimento. A predição de alvo para alguns desses miRNAs indicam que alguns poderiam estar associados a regulação de enzimas envolvidas na remodelação da cromatina. Uma dessas enzimas é a Ezh2 que identificamos como sendo expressa preferencialmente nas células basais (portanto progenitoras) do epitélio olfatório. A identificação do repertório de miRNAs presente nesse trabalho estabelece um fundamento para o estudo funcional de miRNAs no epitélio olfatório no futuro.

Identificamos também um grande número de lincRNAs expressos no epitélio olfatório, sendo que a maioria destes ainda também não possui função conhecida. Dentre estes, 104 são preferencialmente expressos no epitélio olfatório quando comparados a dados de RNA-seq públicos de diversos tecidos (cérebro, cerebelo, coração, fígado, rins e testículos). Também validamos por RT-PCR a expressão de 5 lincRNAs preferencialmente expressos no OE utilizando cDNAs provenientes de outros 10 tecidos nos diferentes estágios de desenvolvimento (neonatos e 4 semanas de vida): cérebro, testículo, baço, olhos, língua, pulmão, coração, tecido adiposo e rim. Observamos que estes lincRNAs são expressos em diferentes tecidos de camundongos neonatos, indicando que estes ncRNAs possam apresentar funções em estágios iniciais de desenvolvimento, porém a expressão destes é mantida apenas no epitélio olfatório em estágios mais tardios de desenvolvimento corroborando com a nossa abordagem de bioinformática na seleção dos ncRNAs preferencialmente expressos no epitélio olfatório. Além disto, mostramos que os lincRNAs validados neste trabalho também são expressos no VNO como mostrado em nossos ensaios de RT-PCR e hibridização *in situ*. O fato de que os lincRNAs validados são expressos tanto nos neurônios vomeronasais como nos neurônios olfatórios, que são células extremamente especializadas que apresentam características muito semelhantes entre si, indica que eles apresentam funções relacionadas a estes tipos celulares. Cabe ressaltar que os cDNAs dos genes de lincRNAs validados por RT-PCR possuem introns enormes, como o do lincRNA *8* que possui um gene de 28 kb e introns de 5,14 e 20,5 kb. Outro exemplo se trata do gene do lincRNA *9* que possui 43 kb com um intron de 42 kb.

Como além de neurônios olfatórios outros tipos celulares também estão presentes no epitélio olfatório, nós realizamos experimentos de hibridização in situ para verificar em que células do epitélio estes lincRNAs são expressos. Contudo nossos resultados destes experimentos foram bastante inconclusivos, já que para dois dos lincRNAs a sonda sense também apresentou um sinal (lincRNAs 5 e 9), para outros dois nenhum sinal foi visualizado (lincRNA 1 e 8), e para um deles, (lincRNA 4) apesar de que obtivemos apenas um sinal utilizando a sonda antisense, como deveria ser correto, a sonda hibridizou com todos os tipos celulares do epitélio, e também do bulbo olfatório, de uma maneira pouco tipo celular-específica. Estes resultados podem se dever a uma inespecificidade da sonda, ou a uma baixa sensibilidade da técnica utilizada. Pode ser também que estes lincRNAs estejam localizados no núcleo da célula, e devem ser visualizados em maior aumento. Poderemos futuramente avaliar se os lincRNAs presentes nesse trabalho estão localizados no núcleo ou no citoplasma utilizando a técnica de fracionamento celular do epitélio olfatório. Após obtenção da porção citoplasmática e nuclear separadamente, conseguiremos avaliar em qual destas frações estes lincRNAs estão localizados.

86

6 CONCLUSÕES

O epitélio olfatório expressa um vasto número de RNAs não codificadores, sendo que a grande maioria ainda não é conhecida. Vários destes ncRNAs apresentam expressão diferencial entre os estágios de desenvolvimento do epitélio olfatório que foram analisados neste estudo, sendo que alguns são mais abundantes no estágio neonatal que no de 4 semanas, ou vice-versa, indicando que eles apresentam papéis relacionados a estas respectivas fases.

Identificamos 103 miRNAs que são preferencialmente expressos no epitélio olfatório, e não nos outros tecidos analisados, o que os torna candidatos interessantes para estudo no futuro. Por exemplo, alguns dos miRNAs identificados possuem como alvo enzimas envolvidas na remodelação da cromatina, como a Ezh2 que mostramos estar presente em células basais progenitoras do epitélio olfatório.

Ao todo, identificamos 295 lincRNAs expressos no epitélio olfatório, sendo que destes 104 apresentaram expressão preferencial no epitélio olfatório quando comparados a outros tecidos. Deste *pool* de lincRNAs 'OE específicos', conseguimos validar experimentalmente a expressão de 5 lincRNAs em epitélio olfatório. Interesantemente, os lincRNAs validados são conservados em diversas espécies incluindo mamíferos, aves e até mesmo peixes como observado para o LincRNA 9. De maneira geral, estes 5 lincRNAs apresentaram uma expressão menos tecido específica nos tecidos analisados de neonatos, e passaram a ser mais tecido específicas nos tecidos de camundongos de 4 semanas. Neste estágio, a expressão dos lincRNAs mostrou-se bastante tecido específica, com expressão apenas no epitélio olfatório, ou com expressão no epitélio olfatório e cérebro. Além disso, mostramos que quatro destes lincRNAs validados são também expressos no VNO, que possui funções semelhantes ao epitélio olfatório.

Como já mencionado, o foco deste trabalho foi verificar quais ncRNAs já anotados em camundongos são expressos no epitélio olfatório. Deste modo não podemos excluir que novos ncRNAs ainda não anotados são expressos no MOE. Além disso, os ncRNAs descritos neste trabalho em sua grande maioria não estão descritos na literatura. Em conjunto, os dados aqui presentes indicam que a abordagem adotada permite a identificação de lincRNAs potencialmente relevantes para a função olfatória.

7 REFERÊNCIAS

ANDERS, S.; HUBER, W. Differential expression analysis for sequence count data. **Genome Biol,** v. 11, n. 10, p. R106, 2010.

ARMELIN-CORREA, L. M. et al. Nuclear compartmentalization of odorant receptor genes. **Proceedings of the National Academy of Sciences of the United States of America,** v. 111, n. 7, p. 2782-2787, 2014.

AVNER, P.; HEARD, E. X-chromosome inactivation: counting, choice and initiation. **Nat Rev Genet**, v. 2, n. 1, p. 59-67, 01//print 2001.

BARTEL, D. P. MicroRNAs: target recognition and regulatory functions. **Cell**, v. 136, n. 2, p. 215-33, Jan 23 2009.

BELETSKII, A. et al. PNA interference mapping demonstrates functional domains in the noncoding RNA Xist. **Proceedings of the National Academy of Sciences of the United States of America,** v. 98, n. 16, p. 9215-9220, 04/08/received 2001.

BROCKDORFF, N. X-chromosome inactivation: closing in on proteins that bind Xist RNA. **Trends in Genetics**, v. 18, n. 7, p. 352-358, 2002.

BROWN, C. J. et al. The human XIST gene: Analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. **Cell**, v. 71, n. 3, p. 527-542,

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.

BURENINA, O. T. S., AND KUBAREVA E.A. Non-Coding RNAs As Transcriptional Regulators In Eukaryotes. **Acta Naturae**, v. v.9(4);, 2017.

CABILI, M. N. et al. Localization and abundance analysis of human IncRNAs at single-cell and single-molecule resolution. **Genome Biology,** v. 16, n. 1, p. 20, 2015.

CAI, B. Z.; PAN, Z. W.; LU, Y. J. The Roles of MicroRNAs in Heart Diseases: A Novel Important Regulator. **Current Medicinal Chemistry,** v. 17, n. 5, p. 407-411, 2010.

CAO, P. et al. MicroRNA-101 negatively regulates Ezh2 and its expression is modulated by androgen receptor and HIF-1alpha/HIF-1beta. **Mol Cancer,** v. 9, p. 108, 2010.

CAO, R. et al. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. **Science,** v. 298, n. 5595, p. 1039-43, Nov 1 2002.

CHAUMEIL, J. et al. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. **Genes & Development,** v. 20, n. 16, p. 2223-2237, 01/22/received

06/22/accepted 2006.

CHEN, C.-K. et al. Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. **Science**, v. 354, n. 6311, p. 468, 2016.

CHESS, A. et al. Allelic inactivation regulates olfactory receptor gene expression. **Cell**, v. 78, n. 5, p. 823-34, Sep 9 1994.

CHOI, P. S. et al. Members of the miRNA-200 family regulate olfactory neurogenesis. **Neuron**, v. 57, n. 1, p. 41-55, Jan 10 2008.

CHU, C. et al. Systematic discovery of Xist RNA binding proteins. **Cell**, v. 161, n. 2, p. 404-416, 04/02 2015.

CLOWNEY, E. J. et al. Nuclear Aggregation of Olfactory Receptor Genes Governs Their Monogenic Expression. **Cell**, v. 151, n. 4, p. 724-737, 2012.

CONLEY, A. B.; KING JORDAN, I. Epigenetic regulation of human cis -natural antisense transcripts. **Nucleic Acids Research**, v. 40, n. 4, p. 1438-1445, 2012.

DULAC, C.; AXEL, R. A novel family of genes encoding putative pheromone receptors in mammals. **Cell,** v. 83, n. 2, p. 195-206, 1995/10/20/ 1995.

EBHARDT, H. A.; FEDYNAK, A.; FAHLMAN, R. P. Naturally occurring variations in sequence length creates microRNA isoforms that differ in argonaute effector complex specificity. **Silence**, v. 1, n. 1, p. 12, 2010.

EDELSTEIN, L. C.; BRAY, P. F. MicroRNAs in platelet production and activation. **Blood**, v. 117, n. 20, p. 5289-5296, 01/05/received

02/15/accepted 2011.

FANG, R. et al. Probing Xist RNA Structure in Cells Using Targeted Structure-Seq. **PLoS Genetics**, v. 11, n. 12, p. e1005668, 12/08

05/26/received

10/24/accepted 2015.

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

_____. A code in the nose. **Sci STKE,** v. 2004, n. 227, p. pe15, Apr 6 2004.

FLYNN, R. A.; CHANG, H. Y. Active chromatin and noncoding RNAs: an intimate relationship. **Current opinion in genetics & development,** v. 22, n. 2, p. 172-178, 12/07 2012.

FRIEDMAN, R. C. et al. Most mammalian mRNAs are conserved targets of microRNAs. **Genome Res**, v. 19, n. 1, p. 92-105, Jan 2009.

FURUNO, M. et al. Clusters of internally primed transcripts reveal novel long noncoding RNAs. **PLoS Genet,** v. 2, n. 4, p. e37, Apr 2006.

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.

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.

GREGORY, R. I. et al. Human RISC Couples MicroRNA Biogenesis and Posttranscriptional Gene Silencing. **Cell**, v. 123, n. 4, p. 631-640, 2005.

GUTTMAN, M. et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. **Nature,** v. 458, n. 7235, p. 223-227, 03/12/print 2009.

HERRADA, G.; DULAC, C. A Novel Family of Putative Pheromone Receptors in Mammals with a Topographically Organized and Sexually Dimorphic Distribution. **Cell**, v. 90, n. 4, p. 763-773,

HILDEBRAND, J. G., AND SHEPHERD, G.M. (MECHANISMS OF OLFACTORY DISCRIMINATION:Converging Evidence for Common Principles Across Phyla. **Annual Review of Neuroscience,** v. 20, n. 1, p. 595-631, 1997/03/01 1997.

KHALIL, A. M. et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. **Proceedings of the National Academy of Sciences,** v. 106, n. 28, p. 11667-11672, 2009.

KOCERHA, J. et al. MicroRNA-219 modulates NMDA receptor-mediated neurobehavioral dysfunction. **Proceedings of the National Academy of Sciences of the United States of America,** v. 106, n. 9, p. 3507-3512, 02/05

06/17/received 2009.

LANCTÔT, C. et al. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. **Nature Reviews Genetics**, v. 8, p. 104, 02/01/online 2007.

LEWIS, B. P.; BURGE, C. B.; BARTEL, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. **Cell**, v. 120, n. 1, p. 15-20, Jan 14 2005.

LEWIS, B. P. et al. Prediction of mammalian microRNA targets. **Cell,** v. 115, n. 7, p. 787-98, Dec 26 2003.

LEWIS, M. A. et al. An ENU-induced mutation of miR-96 associated with progressive hearing loss in mice. **Nature genetics**, v. 41, n. 5, p. 614-618, 04/12 2009.

LIU, C. et al. NONCODE: an integrated knowledge database of non-coding RNAs. **Nucleic Acids Res,** v. 33, n. Database issue, p. D112-5, Jan 1 2005.

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

LU, J. et al. microRNA-mediated control of cell fate in megakaryocyte-erythrocyte progenitors. **Developmental cell**, v. 14, n. 6, p. 843-853, 2008.

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

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. 713-23, Mar 5 1999.

MATSUNAMI, H.; BUCK, L. B. A Multigene Family Encoding a Diverse Array of Putative Pheromone Receptors in Mammals. **Cell**, v. 90, n. 4, p. 775-784, 1997.

MATTICK, J. S. Non-coding RNAs: the architects of eukaryotic complexity. **EMBO Rep**, v. 2, n. 11, p. 986-91, Nov 2001.

MATTICK, J. S.; MAKUNIN, I. V. Small regulatory RNAs in mammals. Hum Mol Genet, v. 14 Spec No 1, p. R121-32, Apr 15 2005.

MERCER, T. R.; DINGER, M. E.; MATTICK, J. S. Long non-coding RNAs: insights into functions. **Nat Rev Genet,** v. 10, n. 3, p. 155-9, Mar 2009.

MICHALAK, P. RNA world - the dark matter of evolutionary genomics. **J Evol Biol**, v. 19, n. 6, p. 1768-74, Nov 2006.

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.

MISTELI, T. Beyond the Sequence: Cellular Organization of Genome Function. **Cell**, v. 128, n. 4, p. 787-800, 2007.

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. **The Journal of Neuroscience**, v. 25, n. 14, p. 3586, 2005.

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

NAGAI, M. H.; ARMELIN-CORREA, L. M.; MALNIC, B. Monogenic and Monoallelic Expression of Odorant Receptors. **Molecular Pharmacology**, v. 90, n. 5, p. 633, 2016.

NESTEROVA, T. B. et al. Characterization of the Genomic Xist Locus in Rodents Reveals Conservation of Overall Gene Structure and Tandem Repeats but Rapid Evolution of Unique Sequence. **Genome Research,** v. 11, n. 5, p. 833-849, 12/11/received

02/27/accepted 2001.

NIELSEN, C. B. et al. Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. **RNA**, v. 13, n. 11, p. 1894-910, Nov 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.

PLATH, K. et al. Role of Histone H3 Lysine 27 Methylation in X Inactivation. **Science,** v. 300, n. 5616, p. 131-135, 2003.

QASBA, P.; REED, R. R. Tissue and zonal-specific expression of an olfactory receptor transgene. **J Neurosci**, v. 18, n. 1, p. 227-36, Jan 1 1998.

RAGOCZY, T.; GROUDINE, M. The Nucleus Inside Out—Through a Rod Darkly. **Cell**, v. 137, n. 2, p. 205-207, 2014.

RAJAPAKSE, I.; GROUDINE, M. On emerging nuclear order. **The Journal of Cell Biology,** v. 192, n. 5, p. 711-721, 10/26/received

02/01/accepted 2011.

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.

RINN, J. L. et al. Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. **Cell**, v. 129, n. 7, p. 1311-1323, 2007.

ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics**, v. 26, n. 1, p. 139-40, Jan 1 2010.

RODRIGUEZ-GIL, D. J. et al. Chromosomal location-dependent nonstochastic onset of odor receptor expression. **J Neurosci**, v. 30, n. 30, p. 10067-75, Jul 28

_____. Chromosomal location-dependent nonstochastic onset of odor receptor expression. **J Neurosci**, v. 30, n. 30, p. 10067-75, Jul 28 2010.

RODRIGUEZ, I. et al. Multiple new and isolated families within the mouse superfamily of V1r vomeronasal receptors. **Nat Neurosci,** v. 5, n. 2, p. 134-140, 02//print 2002.

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.

RYBA, N. J. P.; TIRINDELLI, R. A New Multigene Family of Putative Pheromone Receptors. **Neuron**, v. 19, n. 2, p. 371-379,

SAMUEL C. CHANG, T. T., NANCY P. THOROGOOD, AND CAROLYN J. BROWN. Mechanisms of X-chromosome inactivation. **Frontiers in Bioscience**, v. 11, 852-866, 2006.

SARAIVA, L. R. et al. Molecular and neuronal homology between the olfactory systems of zebrafish and mouse. **Scientific Reports**, v. 5, p. 11487, 06/25

03/17/received

05/27/accepted 2015.

SCHUETTENGRUBER, B. et al. Genome regulation by polycomb and trithorax proteins. **Cell**, v. 128, n. 4, p. 735-45, Feb 23 2007.

SERIZAWA, S. et al. Mutually exclusive expression of odorant receptor transgenes. **Nat Neurosci,** v. 3, n. 7, p. 687-93, 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.

SILVA, J. et al. Establishment of Histone H3 Methylation on the Inactive X Chromosome Requires Transient Recruitment of Eed-Enx1 Polycomb Group Complexes. **Developmental Cell**, v. 4, n. 4, p. 481-495,

SMITH, A. Whisperings of ovarian cancer: acknowledging women's voices. **Clin J Oncol Nurs,** v. 12, n. 6, p. 913-20, Dec 2008.

SMOLA, M. J. et al. SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist IncRNA in living cells. **Proceedings** of the National Academy of Sciences of the United States of America, v. 113, n. 37, p. 10322-10327, 08/30 2016.

SU, S.-Y. et al. Transcriptomic analysis of EGb 761-regulated neuroactive receptor pathway in vivo. **Journal of Ethnopharmacology,** v. 123, n. 1, p. 68-73, 2009/05/04/ 2009.

SULLIVAN, S.; BUCK, L. Target-independent pattern specification in the olfactory epithelium. **Neuron**, v. 15, p. 779-789, 1995.

TOWNLEY-TILSON, W. H. D.; CALLIS, T. E.; WANG, D.-Z. MicroRNAs 1, 133, and 206: Critical factors of skeletal and cardiac muscle development, function, and disease. **The international journal of biochemistry & cell biology**, v. 42, n. 8, p. 1252-1255, 03/14 2010.

TRAPNELL, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. **Nat Protoc,** v. 7, n. 3, p. 562-78, Mar 2012.

TSAI, M.-C. et al. Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. **Science**, v. 329, n. 5992, p. 689-693, 2010.

ULITSKY, I.; BARTEL, D. P. lincRNAs: Genomics, Evolution, and Mechanisms. **Cell**, v. 154, n. 1, p. 26-46, 2013.

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.; 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, F. et al. Odorant receptors govern the formation of a precise topographic map. **Cell**, v. 93, n. 1, p. 47-60, Apr 3 1998.

WANG, M. et al. Expression Profiling of mRNAs and Long Non-Coding RNAs in Aged Mouse Olfactory Bulb. **Scientific Reports,** v. 7, n. 1, p. 2079, 2017/05/18 2017.

WANG, M. M.; REED, R. R. Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. **Nature**, v. 364, n. 6433, p. 121-6, Jul 8 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.

WUTZ; RASMUSSEN; JAENISCH. Chromosomal silencing and localization are mediated by different domains of Xist RNA. **Nature Genetics**, v. 30, p. 167, 01/07/online 2002.

WYSOCKI C. J, M. M. The vomeronasal system. In: Finger TE, Silver WL, editors. In: Neurobiology of Taste and Smell. New York, p. pp. 125–150., 1987.

YAMAMURA, S. et al. Interaction and cross-talk between non-coding RNAs. **Cellular and Molecular Life Sciences,** v. 75, n. 3, p. 467-484, 2018/02/01 2018.

ZHANG, X. et al. High-throughput microarray detection of olfactory receptor gene expression in the mouse. **Proc Natl Acad Sci U S A,** v. 101, n. 39, p. 14168-14173, 2004.

ZHANG, X. et al. 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,** v. 101, n. 39, p. 14168-14173, 2004.

ZHANG, Y. Q.; BREER, H.; STROTMANN, J. Promotor elements governing the clustered expression pattern of odorant receptor genes. **Mol Cell Neurosci**, v. 36, n. 1, p. 95-107, Sep 2007.

ZHAO, J. et al. Polycomb Proteins Targeted by a Short Repeat RNA to the Mouse X Chromosome. **Science**, v. 322, n. 5902, p. 750-756, 2008.

ZHAO, Y. et al. Dysregulation of Cardiogenesis, Cardiac Conduction, and Cell Cycle in Mice Lacking miRNA-1-2. **Cell**, v. 129, n. 2, p. 303-317, 2007.

ZORIO, E. et al. Insights Into the Role of microRNAs in Cardiac Diseases: From Biological Signalling to Therapeutic Targets. **Cardiovascular & Hematological Agents in Medicinal Chemistry,** v. 7, n. 1, p. 82-90, 2009.

8 ANEXOS

8.1 Anexo A - SÚMULA CURRICULAR

João Batista Placido do Nascimento

Nascimento: 24/06/1985 - Santos, SP, Brasil.

Formação Acadêmica

- 2007 2011 Graduação em Farmácia
 Universidade Católica de Santos, Santos, São Paulo, Brasil
 Orientadora: Rosângela Ballego Campanhã
 Título: Determinação de ligantes para receptores olfatórios humanos
- 2012 2018 Curso de Doutorado Direto, Programa de Ciências Biológicas (Bioquímica)
 Universidade de São Paulo, USP, São Paulo, Brasil
 Orientadora: Profa. Dra. Bettina Malnic
 Projeto/Tese: Identificação de RNAs não codificadores expressos no epitélio olfatório
 Bolsista da Fundação de Amparo à Pesquisa do Estado de São Paulo

Formação Complementar

Fev/2014 - Jul/2014	Monitor na disciplina Biologia Molecular, sob a supervisão da Profa. Dra. Daniella Basseres. Universidade de São Paulo, USP, São Paulo, Brasil Bolsista no Programa de Aperfeiçoamento de Ensino – PAE da Universidade de São Paulo
Jul/2013 - Nov/2013	Monitor na disciplina Biologia Molecular, sob a supervisão da Profa. Dra. Bettina. Universidade de São Paulo, USP, São Paulo, Brasil Bolsista no Programa de Aperfeiçoamento de Ensino – PAE da Universidade de São Paulo
Set/2012	Curso "Células-tronco aplicadas ao estudo do cérebro" 58° Congresso Brasileiro de Genética, Foz do Iguaçu, PR
Set/2012	Curso "RNAs não codificadores e a dinâmica genômica" 58° Congresso Brasileiro de Genética, Foz do Iguaçu, PR

Jul/2012	Monitor do VII Curso de Inverno Temas Avançados de Bioquímica e Biologia Molecular, Instituto de Química USP. Universidade de São Paulo, USP, São Paulo, Brasil
Jul/2010 - Dez/2011	Iniciação Científica Universidade de São Paulo, USP, São Paulo, Brasil Pesquisadora responsável: Dra. Daniela Carvalho Gonzalez Kristeller Projeto: Determinação de ligantes para receptores olfatórios humano e chimpanzé-específicos Bolsista da Fundação de Amparo à Pesquisa do Estado de São Paulo
Mai/2010	Oficina Interpretação de Exames Laboratoriais (4h) IV Jornada Científica da Saúde, Universidade Católica de Santos, Santos, São Paulo
Mai/2010	Oficina Toxicologia forense (4h) IV Jornada Científica da Saúde, Universidade Católica de Santos, Santos, São Paulo
Jan/2010 - Jul/2010	Estagiário no setor de oncologia Santa Casa da Misericórdia de Santos, Santos, São Paulo, Brasil
Set/2008	Seminário de Atualização em Técnicas de Aplicação de Injetáveis (12 h) BD, São Paulo, SP.
Ocupação	
Maio/2017 até o pres	ente Farmacêutico Clínico
	Hospital Frei Galvão., Santos, São Paulo
Mai/2017 – Abril/2018	B Farmacêutico Raia Drogasil S. A., Santos, São Paulo

Palestras, Publicações e Comunicações em Congressos

NASCIMENTO, J. B. P. "Pós-Graduação: Desafios e Perspectivas". In: Universidade Católica de Santos, UNISANTOS, Santos, São Paulo

GONZALEZ-KRISTELLER, DANIELA C.; **DO NASCIMENTO, JOÃO B. P.**; GALANTE, PEDRO A. F.; MALNIC, BETTINA. Identification of agonists for a group of human odorant receptors. *Frontiers in Pharmacology*, v. 6, p. 1, 2015 (https://www.frontiersin.org/articles/10.3389/fphar.2015.00035/full).

NASCIMENTO, J. B. P.; GONZALEZ-KRISTELLER, D. C.; GALANTE, P. A.; MALNIC, B. New ligands for human-specific odorant receptors. XLI Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular. Foz do Iguaçu, PR (2012).

NASCIMENTO, J. B. P.; GONZALEZ-KRISTELLER, D. C.; MALNIC, B. Analysis of gene expression in the olfactory epithelium. 58° Congresso Brasileiro de Genética. Foz do Iguaçu, PR (2012).

GONZALEZ-KRISTELLER, D. C.; GALANTE, P. A.; **NASCIMENTO, J. B. P.**; MALNIC, B. Deorphanizing human odorant receptors. 36^o Federation of European Biochemical Societies Congress. Torino, Itália (2011).

NASCIMENTO, J. B. P.; GONZALEZ-KRISTELLER, D. C. Determinação de ligantes para receptores olfatórios humano e chimpanzé-específicos. I Mostra de Trabalhos Acadêmicos-Científicos, Universidade Católica de Santos, UNISANTOS. Santos, SP (2010).

CORRÊA, C. F.; CAMPANHÃ, R. B.; MONTANARI, P. M.; **NASCIMENTO, J. B. P.**; FEITOSA JUNIOR, O.; LIRA, F. B. C.; ARRUDA, G. F. F. Levantamento de espécies vegetais utilizadas com finalidade terapêutica pela comunidade Vila de Itatinga Bertioga / SP. IX Congresso de Saúde Coletiva. Recife, PE (2009).

8.2 Anexo B - Artigo em preparação, e submetido para a disciplina QBQ57499 (Redação Científica)

The noncoding RNA repertoire of the mouse olfactory epithelium

¹ João Batista P.Nascimento, ² Pedro A.F. Galante, ¹ Bettina Malnic.

¹Departamento de Bioquímica, IQ, USP, SP, Brazil, ²Laboratório de Bioinformática, Instituto Sírio-Libanês de Ensino e Pesquisa, SP, Brazil.

Corresponding author: Bettina Malnic, bmalnic@iq.usp.br Telephone: +55 11 3091-1201 Instituto de Química, Universidade de São Paulo Av. Prof. Lineu Prestes, 748, São Paulo – SP, Brazil. CEP: 05508-000

ABSTRACT

Odorants are detected by hundreds of odorant receptors (ORs) which belong to the superfamily of G protein-coupled receptors. These receptors are expressed in the olfactory sensory neurons of the nose. Each olfactory sensory neuron expresses one single OR gene allele from a large family of OR genes. This characteristic pattern of OR gene expression results in the formation of a spatial olfactory map in the olfactory bulb, which is required for odorant discrimination by the olfactory system. The mechanisms involved in this regulation are unknown. OR genomic DNA in olfactory neurons is covered with repressive histone methylation marks, indicating that the regulation of chromatin structure should play an important role in the regulation of OR gene expression. Previous studies suggest that ncRNAs are involved in the deposition of histone marks. In this work, we identified and cataloged the complete repertoire of ncRNAs expressed in the olfactory epithelium from newborn and adult mice. Out of these repertoires, we selected lincRNAs that are preferentially expressed in the olfactory epithelium when compared to other tissues by using

bioinformatics approaches. Ten out of these lincRNAs were validated by using RT-PCR. We also performed *in situ* hybridization experiments to determine the location of the lincRNAs expression within the olfactory epithelium.

Keywords: Olfactory epithelium, ncRNAs, lincRNAs, deep sequencing.

INTRODUCTION

The olfactory system has the ability to detect and distinguish a large number of olfactory stimuli [1]. This vast receptive field is afforded by the large repertoire of olfactory receptors (OR), which, in most mammals, are encoded by more than a thousand genes located in numerous genomic clusters throughout the genome [2-4]. In the mouse, each olfactory sensory neuron expresses one out of ~1,000 different OR genes, which are dispersed throughout the genome [5,6]. Moreover, OR genes are monoallelically expressed, that is, only one allele of the gene (maternal or paternal) is transcribed per neuron [5,7,8]. The mechanism that involves both monoallelic and monogenic expression of OR genes are unclear.

Magklara and collaborators [9] demonstrated that the OR genes have constitutive heterochromatin histone methylation marks (H3K9me3 and H4K20me3) in the nuclei of olfactory neurons but not in the nuclei of liver cells. In addition, immuno DNA-FISH experiments showed that even though the OR genes are spread throughout the genome, in the three-dimensional organization of the olfactory nucleus they are clustered together, and commonly associated with centrally located repressive constitutive and facultative (H3K27me3) heterochromatin compartments [10,11]. These results indicate that the regulation of chromatin structure is likely to play an important role in the regulation of OR gene expression. However, it is not clear yet how these changes in chromatin structure are established and regulated in the olfactory nucleus.

It has been demonstrated that lincRNAs are involved in the regulation of chromatin structure, such as the formation of facultative heterochromatin. These lincRNAs interact with proteins that compose remodeling chromatin complexes with repressor and activator functions [12-14]. However, very little is known about non-coding RNAs (ncRNAs) expressed in the olfactory epithelium (OE). Several microRNAs (miRNAs) expressed in the mouse olfactory epithelium were previously

identified by using a miRNA microarray chip and by sequencing of miRNAs purified from the OE [15]. Among these miRNAs, a family of miRNA (denominated miRNA-200 family) that is required for the differentiation of olfactory neurons was identified. Here we used RNA seq to determine the complete repertoires of non-coding RNAs expressed in the olfactory epithelium from newborn mice (P3) and from 4 weeks old mice. We used bioinformatics to screen for Long Intergenic Non-coding RNAs (lincRNAs) that are preferentially expressed in the OE. Some of the identified lincRNAs were selected for validation through RT-PCR and in situ hybridization.

RESULTS

RNA-seq to identify noncoding RNAs expressed in the olfactory epithelium

Previous *in situ* hybridization and microarray studies showed that although OR genes are first expressed during the embryonic stage E12, the majority of the OR genes are expressed only after birth [16-18]. These results indicate that in the olfactory epithelium from newborn mice, most of the olfactory neurons are selecting one OR gene to express, while in the olfactory epithelium from 4 weeks old mice, most neurons have already done so. Therefore, we decided to compare the olfactory epithelium transcriptomes from newborn and 4 weeks old mice.

Total RNA was purified from the olfactory epithelium of newborn and 4 weeks old C57BL/6 mice. The integrity of the RNA samples was analyzed using an Agilent Bioanalyzer. The complete transcriptomes (including coding and ncRNAs) and the small RNAs were sequenced by using the sequencing by ligation method (SOLiD). A flow diagram describing our sequencing strategy is shown in Figure 1.

The total number of transcripts corresponding to coding genes obtained in the two samples are shown in Table 1. As expected, transcripts for OE specific genes such as odorant receptors (ORs), olfactory marker protein (OMP), and Ric-8B were identified in both samples. The number of noncoding RNAs and miRNAs identified in both samples are also shown in Table 1. Altogether, considering fpkm \geq 2, we found respectively 12,242 and 1,246 coding and noncoding transcripts in the newborn OE transcriptome library, and 11,799 and 1,207 different coding and noncoding transcripts in the 4 weeks old OE transcriptome library. In the small RNAs libraries, we found transcripts for 630 miRNAs in the newborn OE and transcripts for 548

miRNAs in the 4 weeks old OE. The small RNAs libraries included members of the miR-200 family, a family of miRNAs previously described as expressed in the olfactory epithelium of mice [15]. The transcriptome sequencing data from our four different libraries (two transcriptome libraries and two libraries containing small RNAs) is shown in Supplementary Table 1.

Analysis of differential gene expression between samples from newborns and 4 weeks old olfactory epithelia, showed that while 15,036 genes (including coding and noncoding) are expressed in both samples, 2,322 genes are expressed only in 4 weeks old OE and 1,747 genes are expressed only in P3 OE (Figure 2).

Noncoding RNAs preferentially expressed in the olfactory epithelium

We next selected the ncRNAs that are preferentially expressed in the olfactory epithelium, and not in other tissues. This was done by browsing our lincRNAs against public RNA-seq (containing ncRNAs) data from 6 different tissues (brain, cerebellum, heart, liver, kidneys and testicles), and selecting the ncRNAs that are expressed in our libraries but not in these different tissues. We first analyzed a particular group of ncRNAs, the lincRNAs. Among all ncRNAs identified in our samples, we found a total number of 295 lincRNAs. By comparing these lincRNAs with the public RNA-seq from the different tissues, we found 104 lincRNAs that are exclusively expressed in our samples, that is, in the OE (Figure 2 c). We used the same approach to select the miRNAs that are expressed in the OE, and found that out of a total number of 200 miRNAs identified in our samples, 103 miRNAs are exclusively expressed in the OE (Figure 2 b). It is important to note that in this work we considered only annotated ncRNAs, so that we cannot exclude the possibility that additional unannotated ncRNAs which we were not able to identify are expressed in the OE. We also searched for ncRNAs transcripts which genes map in or near OR gene loci, however none were found.

Validation of lincRNA transcripts

We selected 10 out of the 104 lincRNAs preferentially expressed in the OE for validation. Our selection was based on the following criteria: I. LincRNAs should be expressed more abundantly in neonate OE when compared to OE from 4 weeks old

mice; II. LincRNAs should have FPKM values higher than 6. The selected lincRNAs were named lincRNA 1 to lincRNA 10 and are shown in Table 2. Next, we designed specific primers for each one of the selected lincRNA so that each primer matched to different exons to ensure that lincRNAs were amplified only from cDNA and avoiding amplicons generated from contaminating genomic DNA (Figure 3 a). These primers were used to perform PCR reactions using cDNA synthetized from total RNA prepared from OE from newborn and 4 weeks old mice. We were able to amplify PCR products with the expected sizes for 5 out of the 10 selected lincRNAs (lincRNAs 1, 4, 5, 8, 9 and 10) (Figure 3 b).

Tissue distribution of lincRNA expression

After validating the expression of 5 lincRNAs in OE, we verified by using RT-PCR the expression of these lincRNAs in 10 additional tissues (brain, testis, spleen, eyes, tongue, lung, heart, kidney, liver andadipose tissue) (Figure 4). Tissue distribution of the lincRNA expression is more widespread in newborn mice (Figure 4 a), than in 4 weeks old mice (Figure 4 b). Interestingly, the expression of the majority of the lincRNAs in 4 weeks old mice is preferential in the OE when compared to the other tissues analyzed in this assay (Figure 4 b). Particularly noticeable, is lincRNA 8, which is preferentially expressed in the OE from all tissues in both P3 and 4 weeks old mice.

Spatial localization of lincRNA expression in the in olfactory epithelium

The lincRNAs amplified by RT-PCR were sequenced and used to synthesize probes for in situ hybridization that was performed using newborn and 4 weeks old mice. As a positive control we used an olfactory marker protein (OMP) antisense probe, because this gene is expressed only in mature olfactory neurons, and as expected, only this cell type was stained (Figure 5 a). This assay was performed for all validated lincRNAs, however we obtained positive staining only for lincRNA 4. The antisense probe to lincRNA 4 clearly stained the OE in both newborns and in 4 weeks old OE (Figure 5 b and c). Besides the main olfactory epithelium (MOE), detection of olfactory stimuli is also accomplished by a second sensory structure in

the nose, the vomeronasal organ (VNO)[19,20]. Curiously, we also observed the expression of lincRNA 4 in the VNO (Figure 5 d).

DISCUSSION

In this work, we identified and cataloged the repertoire of ncRNAs expressed in the OE. Using a bioinformatics approach, we compared the expression of miRNAs and lincRNAs expressed in OE and in several tissues and identified ncRNAs that are preferentially expressed in the OE. It is important emphasize that all ncRNAs identified in our study were previously annotated, not excluding the possible existence of additional unannotated ncRNAs expressed in OE which we could not identify.

Because lincRNAs can regulate gene transcription through different mechanisms, we decided to focus our analysis on this group of ncRNAs. A well characterized example for interaction between lincRNAs with proteins is the lincRNA HOTAIR (Transcript Antisense RNA) [21], which is transcribed from the HOXC locus, and silences HOXD genes as well as other targets. Initially it was reported that HOTAIR binds to the Polycomb repressive complex 2 (PCR2) and directs this complex to HOXD promoter regions [21]. This complex mediates gene silencing through chromatin reorganization by methylation of histone H3 lysine 27 (H3K27). The lincRNA HOTAIR serves as a scaffold for at least two distinct histone modification complexes. A 5' domain of HOTAIR binds PRC2 while a 3' domain of HOTAIR binds the LSD1/CoREST/REST complex that removes dimethyl residues from lysine 4 on histone 3 (H3K4me2), which is also related to silencing of HOXD genes [22]. Previous chromatin immunoprecipitation (ChIP) experiments showed that OR genes are marked with the repressive constitutive heterochromatin marks in the nuclei of olfactory neurons [9]. In addition, experiments using a complex DNA-FISH probe that recognizes the complete OR gene repertoire have shown that the OR genes are aggregated in large constitutive heterochromatic foci in the nucleus of olfactory sensory neurons [11]. We have shown that compartments containing facultative (H3K27me3) marks are located adjacent to these constitutive heterochromatin foci in the nuclei of olfactory neurons [10]. This aggregation would organize the OR genes for monogenic activation, so that only a single OR gene allele would escape from chromatin repression and become active [11]. Together these data provide evidence that lincRNAs may be involved in maintaining histone marks
on the olfactory genome. We identified and validated by RT-PCR 5 lincRNAs. These lincRNAs were expressed in various tissues from newborn mice, however, in older mice they were more restrictedly expressed in the OE (Figure 4 a and b), suggesting that these lincRNAs should have a significant role in mature OE.

In situ hybridization experiments show that lincRNA 4 is expressed in the olfactory neurons in the OE. Moreover, lincRNA 4 is also expressed in the vomeronasal neurons in the VNO (Figure 5). Olfactory and vomeronasal neurons are sensory neurons that share common features, such as monogenic and monoallelic expression of odorant receptor genes. It is possible that lincRNA 4 plays a role that is specific to these types of cells. LincRNA 4 has one exon (exon 2) which is conserved in several other species including humans (Supplementary figure 1). No staining was observed when we used probes corresponding to the other lincRNAs. Importantly, lncRNAs can shuttle between various cellular compartments. Some show markedly different levels of accumulation in the nucleus versus the cytoplasm while others are equally distributed between both compartments [23-25]. It is possible that the method employed may not be sensitive enough to detect lincRNAs in the nuclear compartments, and different approaches should be used to analyze these lincRNAs.

METHODS

Animal Procedures

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 19/2013.

RNA Sequencing

Ribosomal RNA was depleted from total RNA prepared from olfactory epithelia from newborn (P3) and 4 weeks old mice using the RiboMinus Eukaryote Kit for RNA-Seq (Invitrogen). RNA-Seq libraries were prepared using SOLiD Total RNA-Seq Kit, according to the manufacturers' recommendations, and were sequenced on the SOLiD sequencing platform (Life technologies, Carlsbad, CA) in the Molecular Oncology Laboratory at Sirio Libanes Hospital in collaboration with Dr. Anamaria Aranha Camargo and Pedro A.F. Galante. Sequences were then aligned to the mouse genome (version GRCm38/mm10) using Tophat [26] with default parameters and gene annotations provided by Ensembl version 71 [27]. Alignments were filtered with SAMtools [28]. The uniquely mapped reads with minimum mapping quality 20 were used to calculate gene expression, which was generated using Cufflinks [29].

RT-PCR

RNA was prepared from different mouse tissues using TRIzol reagent (Invitrogen). First, 1 μ g of total RNA plus 100 ng of oligo-dT in 13.5 μ l of free DNase/RNase water were incubated for 2 min at 70°C. The reaction was rapidly chilled on ice and used to synthesize cDNA in 20 μ l of 1 × Superscript II first-strand buffer containing 0.5 mM dNTP, 3 mM MgCl2, 20 U of RNase inhibitor (RNaseOUT; Invitrogen), and 200 U of Superscript II reverse transcriptase at 42°C for 60 min. The product was diluted to 100 μ l with DEPC-treated water. PCR reactions (25 μ l) containing 5 μ l of cDNA, 0.2 mM dNTP, 1.5 mM MgCl2, 0.5 μ M each of forward and reverse primers, and 1.25 U of Platinum TaqDNA polymerase (Invitrogen) were heated to 95°C for 2 min, followed by 35 thermal cycles of 95°C for 45 s, 58°C for 45 s, and 72°C for 1 min, with a final incubation at 72°C for 10 min. PCR products were analyzed in 1.5% agarose gels stained with ethidium bromide. Primer sequences used to amplify the lincRNAs were:

LincRNA 1F 5' CTC AGC GAT CAA CTG GTT GAC 3' LincRNA1R 5' GGT GGA GCT GAT GTG ATA AGC 3'

LincRNA 2F 5' TGC CTT CTG GTC ATC TCA GG 3' LincRNA 2R 5' TAC TTC AGC TCC TGC CCT TG 3'

LincRNA 3F 5'CCT CTG CCT CTC AAG TGT TG 3' LincRNA 3R 5'CAC CAC ATC TCC CTT GGA TC 3'

LincRNA 4F 5'TTT CAC ACC AGA CTC TGG CC 3' LincRNA 4R 5'AAG CTG GAA GTG TGG ACA GTG 3'

LincRNA 5F 5'GCT GCA GAA GTC ACC AAA GG 3' LincRNA 5R 5'TAA CGG ACT GCA GGC TGT AG 3' LincRNA 6F 5' CTG GCT CAT TCA ACT CTG CAC3' LincRNA 6R 5' AGC CAT AGT TCA TCA CTG AGT G 3'

LincRNA 7F 5'ATG GTG CCA GAC AAG ACT GC 3' LincRNA 7R 5'TTC ACA CCC AGA AGA GGC AG 3'

LincRNA 8F 5' TTG AGT TCC TGC CTC AGC TAC 3' LincRNA 8R 5' GCT TCT GAG TCT ACT GTG ATG 3'

LincRNA 9F 5'GGT GTC TCA CCG AAA GGT AG 3' LincRNA 9R 5'TGG CAG TGG TAG CAC ATG CC 3'

LincRNA 10F 5'GAT GGA CCA TGA GCA CAT GG 3' LincRNA 10R 5'CTG CAG AGA CAT CTG CTT GC 3'

The PCR products were cloned into the pGEM expression vector (Invitrogen) and sequences of the cloned lincRNAs were checked by DNA sequencing by ABI PRISM 3100.

In situ hybridization

In situ hybridization was performed as previously described [30] and according to [31]. Briefly, noses were dissected from 3 days and 4 weeks old C57BL/6 mice and freshly embedded in Tissue-Tek OCT compound (Sakura Finetek, Torrance, CA). Sequential 18 µm sections were prepared with a cryostat and hybridized with digoxigenin-labeled cRNA probes prepared from lincRNAs.

REFERENCES

1. Bushdid C, Magnasco MO, Vosshall LB, Keller A (2014) Humans Can Discriminate More than 1 Trillion Olfactory Stimuli. Science 343 (6177):1370-1372. doi:10.1126/science.1249168

2. Buck L, Axel R A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. Cell 65 (1):175-187. doi:10.1016/0092-8674(91)90418-x

3. Sullivan SL, Adamson MC, Ressler KJ, Kozak CA, Buck LB (1996) The chromosomal distribution of mouse odorant receptor genes. Proceedings of the National Academy of Sciences of the United States of America 93 (2):884-888

4. Zhang X, Rodriguez I, Mombaerts P, Firestein S (2004) Odorant and vomeronasal receptor genes in two mouse genome assemblies. Genomics 83 (5):802-811. doi:http://dx.doi.org/10.1016/j.ygeno.2003.10.009

5. Chess A, Simon I, Cedar H, Axel R Allelic inactivation regulates olfactory receptor gene expression. Cell 78 (5):823-834. doi:10.1016/s0092-8674(94)90562-2

6. Godfrey PA, Malnic B, Buck LB (2004) The mouse olfactory receptor gene family. Proceedings of the National Academy of Sciences 101 (7):2156-2161. doi:10.1073/pnas.0308051100

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

8. Ishii T, Serizawa S, Kohda A, Nakatani H, Shiroishi T, Okumura K, Iwakura Y, Nagawa F, Tsuboi A, Sakano H (2001) Monoallelic expression of the odourant receptor gene and axonal projection of olfactory sensory neurones. Genes to Cells 6 (1):71-78. doi:10.1046/j.1365-2443.2001.00398.x

9. Magklara A, Yen A, Colquitt BM, Clowney EJ, Allen W, Markenscoff-Papadimitriou E, Evans ZA, Kheradpour P, Mountoufaris G, Carey C, Barnea G, Kellis M, Lomvardas S (2011) An Epigenetic Signature for Monoallelic Olfactory Receptor Expression. Cell 145 (4):555-570. doi:10.1016/j.cell.2011.03.040

10. Armelin-Correa LM, Gutiyama LM, Brandt DYC, Malnic B (2014) Nuclear compartmentalization of odorant receptor genes. Proceedings of the National Academy of Sciences of the United States of America 111 (7):2782-2787. doi:10.1073/pnas.1317036111

11. Clowney EJ, LeGros Mark A, Mosley Colleen P, Clowney Fiona G, Markenskoff-Papadimitriou Eirene C, Myllys M, Barnea G, Larabell Carolyn A, Lomvardas S (2012) Nuclear Aggregation of Olfactory Receptor Genes Governs Their Monogenic Expression. Cell 151 (4):724-737. doi:http://dx.doi.org/10.1016/j.cell.2012.09.043

12. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D (2009) Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A 106. doi:10.1073/pnas.0904715106

13. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458. doi:10.1038/nature07672

14. Flynn RA, Chang HY (2012) Active chromatin and noncoding RNAs: an intimate relationship. Current opinion in genetics & development 22 (2):172-178. doi:10.1016/j.gde.2011.11.002

15. Choi PS, Zakhary L, Choi W-Y, Caron S, Alvarez-Saavedra E, Miska EA, McManus M, Harfe B, Giraldez AJ, Horvitz RH, Schier AF, Dulac C (2008) Members of the miRNA-200 Family Regulate Olfactory Neurogenesis. Neuron 57 (1):41-55. doi:http://dx.doi.org/10.1016/j.neuron.2007.11.018

16. Rodriguez-Gil DJ, Treloar HB, Zhang X, Miller AM, Two A, Iwema C, Firestein SJ, Greer CA (2010) Chromosomal location-dependent non-stochastic onset of odor receptor expression. The Journal of neuroscience : the official journal of the Society for Neuroscience 30 (30):10067-10075. doi:10.1523/jneurosci.1776-10.2010

17. Sullivan SL, Bohm S, Ressler KJ, Horowitz LF, Buck LB (1995) Target-independent pattern specification in the olfactory epithelium. Neuron 15 (4):779-789. doi:http://dx.doi.org/10.1016/0896-6273(95)90170-1

18. Zhang X, Rogers M, Tian H, Zhang X, Zou D-J, Liu J, Ma M, Shepherd GM, Firestein SJ (2004) 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 101 (39):14168-14173. doi:10.1073/pnas.0405350101

19. Munger SD, Leinders-Zufall T, Zufall F (2009) Subsystem Organization of the Mammalian Sense of Smell. Annual Review of Physiology 71 (1):115-140. doi:10.1146/annurev.physiol.70.113006.100608

20. Halpern M (1987) The Organization and Function of the Vomeronasal System.AnnualReviewofNeuroscience10(1):325-362.doi:doi:10.1146/annurev.ne.10.030187.001545

21. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough H, Helms JA, Farnham PJ, Segal E, Chang HY (2007) Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Non-Coding RNAs. Cell 129 (7):1311-1323. doi:10.1016/j.cell.2007.05.022

22. Tsai M-C, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY (2010) Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. Science (New York, NY) 329 (5992):689-693. doi:10.1126/science.1192002

23. Ulitsky I, Bartel David P lincRNAs: Genomics, Evolution, and Mechanisms. Cell 154 (1):26-46. doi:10.1016/j.cell.2013.06.020

24. van Heesch S, van Iterson M, Jacobi J, Boymans S, Essers PB, de Bruijn E, Hao W, MacInnes AW, Cuppen E, Simonis M (2014) Extensive localization of long noncoding RNAs to the cytosol and mono- and polyribosomal complexes. Genome Biology 15 (1):R6. doi:10.1186/gb-2014-15-1-r6

25. Cabili MN, Dunagin MC, McClanahan PD, Biaesch A, Padovan-Merhar O, Regev A, Rinn JL, Raj A (2015) Localization and abundance analysis of human IncRNAs at singlecell and single-molecule resolution. Genome Biology 16 (1):20. doi:10.1186/s13059-015-0586-4

26. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25 (9):1105-1111. doi:10.1093/bioinformatics/btp120

27. Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, Cummins C, Clapham P, Fitzgerald S, Gil L, Girón CG, Gordon L, Hourlier T, Hunt SE, Janacek SH, Johnson N, Juettemann T, Keenan S, Lavidas I, Martin FJ, Maurel T, McLaren W, Murphy DN, Nag R, Nuhn M, Parker A, Patricio M, Pignatelli M, Rahtz M, Riat HS, Sheppard D, Taylor K, Thormann A, Vullo A, Wilder SP, Zadissa A, Birney E, Harrow J, Muffato M, Perry E, Ruffier M, Spudich G, Trevanion SJ, Cunningham F, Aken BL, Zerbino DR, Flicek P (2016) Ensembl 2016. Nucleic Acids Research 44 (Database issue):D710-D716. doi:10.1093/nar/gkv1157

28. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25 (16):2078-2079. doi:10.1093/bioinformatics/btp352

29. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protocols 7 (3):562-578

30. Von Dannecker LEC, Mercadante AF, Malnic B (2005) Ric-8B, an Olfactory Putative GTP Exchange Factor, Amplifies Signal Transduction through the Olfactory-Specific G-Protein Gαolf. The Journal of Neuroscience 25 (15):3793-3800. doi:10.1523/jneurosci.4595-04.2005

31. Schaeren-Wiemers N, Gerfin-Moser A (1993) A single protocol to detect transcripts of various types and expression levels in neural tissue and cultured cells: in situ hybridization using digoxigenin-labelled cRNA probes. Histochemistry 100 (6):431-440. doi:10.1007/bf00267823



Figure 1: Flowchart of the strategy used for OE sequencing. The olfactory epithelium of newborns (P3) and 4 weeks old (4W) mice were dissected for RNA extraction. We evaluated RNA integrity by gel electrophoresis. After ribosomal RNA was removed using RiboMinus[™] kit (invitrogem). A library for small RNAs and for total transcriptome were made and sequenced using SOLiD platform.



Table 1: RNA sequencing results. Coding and non coding gene expression and lincRNAs with FPKM \geq 2. Genes showing differential expression (Cuffdiff \cap DESeq. log2 fold change > 1 || log2 fold change < -1). Expressed miRNAs: CPM (count per million) >= 1. miRNAs differential expression expressed genes: CPM >= 1. Methods: edgeR \cap DESeq.



Figure 2: Deep sequencing results of olfactory epithelium. (A) Expressed genes in newborns and 4 weeks old mice samples, (B) miRNAs expressed only in olfactoy epithelium when compared with public RNA-seq data from 6 differents tissues (brain, cerebellum, heart, liver, kidneys and testicles), (C) lincRNAs expressed only in olfactory epithelium when compared with public RNA-seq data from 6 differents tissues (brain, cerebellum, heart, liver, kidneys and testicles).

	Ensembl ID	Newborns FPKM	4 Weeks FPKM
LincRNA 1	ENSMUSG0000084762	42,4609	35,3581
LincRNA 2	ENSMUSG0000097540	34,7526	6,71802
LincRNA 3	ENSMUSG0000085359	29,9472	24,468
LincRNA 4	ENSMUSG0000055494	15,1637	11,5465
LincRNA 5	ENSMUSG0000097618	12,622	0.269009
LincRNA 6	ENSMUSG0000097143	11,4152	12,4764
LincRNA 7	ENSMUSG0000085499	7,91438	5,21452
LincRNA 8	ENSMUSG0000085600	6,48943	6,9646
LincRNA 9	ENSMUSG0000086224	6,01844	1,41534
LincRNA 10	ENSMUSG0000086523	1,92126	17,7975

Table 2: LincRNAs numbered from 1 to 10 according their respective FPKM.



Figure 3: RT-PCR of LincRNAs. (A) Schematic model showing how the primers were designed. In this figure we use lincRNA 1 as example. (B) RT-PCR reactions using cDNA synthetized from newborn (P3) and 4 weeks (4W) old mice OE. For each tested lincRNA RT-PCR was added a negative control reaction without cDNA and GAPDH as positive control.



Figure 4: RT-PCR in 10 differents tissues. (A) RT-PCR reactions using cDNA from differents tissues of newborn mice. (B) RT-PCR reactions using cDNA from differents tissue of 4 weeks old mice. For each tested lincRNA RT-PCR was added a negative control reaction without cDNA and GAPDH as positive control.



Figure 5: LincRNAs In situ hibridization in Olfactory Epithelium. (A) positive control OMP. (B) In situ hibridization for LincRNA 4 in OE of newborns mice (P3). (C) In situ hibridization for LincRNA 4 in OE of 4 weeks old mice (4W). (D) In situ hibridization for LincRNA 4 in VNO of 4 weeks old mice.



Supplementary figure 1: Conservation of validated LincRNAs. The validated lincRNAs were subjected to Blatt on UCSC Genome Browser tool.

8.3 Arquivos suplementares

Os arquivos suplementares encontram-se no CD anexo à tese.

Arquivo suplementar 1 - Transcritoma epitélio olfatório de camundongos neonatos. Arquivo suplementar 2 - Transcritoma de epitélio olfatório de camundongos de 4 semanas.

Arquivo suplementar 3 - miRNAs Arquivo suplementar 4 - lincRNAs