

**UNIVERSIDADE DE SÃO PAULO**  
**INSTITUTO DE QUÍMICA**  
**Programa de Pós-Graduação em Bioquímica**

**JUSSARA MICHALOSKI SOUZA**

**Análise das regiões promotoras dos  
genes de receptores olfatórios  
e de receptores de feromônios  
do tipo 1.**

São Paulo

Data de Depósito na SPG:  
01/09/2008

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

JUSSARA MICHALOSKI SOUZA

**Análise das regiões promotoras dos genes de receptores olfatórios e de receptores de feromônios do tipo 1.**

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

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Bettina Malnic

São Paulo  
2008

**Jussara Michaloski Souza**

**Análise das regiões promotoras dos genes de receptores olfatórios e de receptores de feromônios do tipo 1.**

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

Aprovado em: \_\_\_\_\_

Banca Examinadora:

Prof. Dr. \_\_\_\_\_

Instituição: \_\_\_\_\_

Assinatura: \_\_\_\_\_

Prof. Dr. \_\_\_\_\_

Instituição: \_\_\_\_\_

Assinatura: \_\_\_\_\_

Prof. Dr. \_\_\_\_\_

Instituição: \_\_\_\_\_

Assinatura: \_\_\_\_\_

Prof. Dr. \_\_\_\_\_

Instituição: \_\_\_\_\_

Assinatura: \_\_\_\_\_

Prof. Dr. \_\_\_\_\_

Instituição: \_\_\_\_\_

Assinatura: \_\_\_\_\_

Aos meus queridos pais, Piragibe e Leonor, pelo apoio e incentivo durante  
toda minha vida.

Ao meu querido companheiro João que com seu amor, paciência e bom  
humor tem deixado minha vida mais feliz.

## Agradeço...

...à Prof<sup>a</sup> Dra. Bettina Malnic pela excelente orientação, pelo incentivo e ensinamentos, pelas discussões e sugestões sempre pertinentes, enfim, por toda colaboração na execução deste trabalho.

...à Prof<sup>a</sup> Dra. Adriana F. Mercadante que durante seu *post-doc* no laboratório também colaborou com ensinamentos e sugestões muito proveitosas para meu doutorado.

...aos colegas de laboratório, Daniel Kerr, Daniela Gonzales, Eduardo Alves, Guilherme Louzada, Isafias Glezer, Lilian Inoue, Lucia Armelin, Luiz Dannecker, Marcela Davalos, Pedro Galante e Camila Fleury pelas discussões proveitosas, por toda a ajuda, pelos momentos alegres e convivência agradável.

...aos técnicos de laboratório Érica Bandeira e Edson Pereira pela valiosa colaboração técnica.

...à minha família, meus pais Leonor e Piragibe, minhas irmãs Adriana e Márcia, meu companheiro e melhor amigo João Paulo pelo apoio e torcida.

...à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo apoio financeiro para realização deste trabalho.

## RESUMO

Michaloski, J.S. **Análise das regiões promotoras dos genes de receptores olfatórios e de receptores de feromônios do tipo 1.** 2008. 103 p. Tese de Doutorado – Programa de Pós-Graduação em Bioquímica. Instituto de Química, Universidade de São Paulo, São Paulo.

No genoma de camundongo existem por volta de 1000 genes que codificam para receptores olfatórios (ORs) e 150 genes que codificam para receptores de feromônios do tipo 1 (V1Rs) distribuídos em vários cromossomos. Cada neurônio olfatório e vomeronasal seleciona um único alelo de um único gene de receptor OR ou de V1R, respectivamente, para expressar enquanto que o restante do repertório é mantido silenciado. Os mecanismos que regulam esse padrão de expressão não são conhecidos. As similaridades no padrão de expressão dos genes de ORs e de V1Rs sugerem que o mecanismo de regulação possa ser comum.

Até então poucas regiões promotoras de genes de ORs e de genes de V1Rs haviam sido experimentalmente determinadas e pesquisadas. Realizamos uma análise na qual regiões a montante de um grande número de diferentes genes de ORs e de genes de V1Rs foram comparadas. Primeiro, utilizando a técnica de RLM-RACE, combinada com o uso de oligonucleotídeos capazes de reconhecer regiões conservadas entre diversos membros das famílias de genes de ORs e de V1Rs, geramos centenas de cDNAs contendo a região 5'UTR completa para um total de 198 genes de ORs e 39 genes de V1Rs diferentes. Então, alinhamos as sequências desses cDNAs contra o genoma de camundongo e localizamos a posição exata dos sítios de início da transcrição (TSSs) de cada gene. Extraímos seqüências a 5' dos TSSs dos 198 genes de ORs e dos 39 genes de V1Rs e buscamos por motivos de DNA comuns, presentes em várias dessas regiões promotoras, que pudessem ser

candidatos a elementos *cis*-atuantes envolvidos na regulação geral desses genes de receptores sensoriais.

Identificamos, na grande maioria das regiões promotoras dos genes de ORs e dos genes de V1Rs analisadas, a presença de motivos semelhantes a sítios de ligação para os fatores de transcrição O/E que são fatores de transcrição já caracterizados e envolvidos com a expressão de genes específicos do sistema olfatório. Ensaio de EMSA mostraram que os motivos semelhantes aos sítios de ligação de O/E identificados interagem com proteínas nucleares de epitélio olfatório, mas não interagem com proteínas nucleares de cérebro e fígado.

Identificamos também nas regiões promotoras de genes de V1Rs a presença de um sítio de ligação que não se assemelha a nenhum sítio de ligação de fatores de transcrição conhecido. Esse motivo de DNA, além de estar presente na maioria dos promotores de genes de V1Rs analisados (77% do total de 39 genes pesquisados), também aparece, com alta frequência, em promotores de genes de ORs (52% do total de 198 genes analisados), preferencialmente próximo aos TSSs. Ensaio de interação *in vitro* indicam que este novo motivo de DNA interage com proteínas nucleares extraídas de órgão vomeronasal e também de epitélio olfatório, mas não interage com proteínas nucleares de cérebro, fígado e pulmão.

Nosso trabalho mostra que genes de ORs e de V1Rs compartilham elementos comuns em suas regiões promotoras os quais podem ser sítios de ligação de fatores de transcrição específicos do sistema olfatório envolvidos no mecanismo de regulação da expressão desses genes.

**Palavras-chave:** expressão gênica, promotor, fator de transcrição, receptor olfatório, receptor vomeronasal.



## ABSTRACT

Michaloski, J.S. **Analysis of promoter regions of the olfactory and type 1 vomeronasal receptor genes.** 2008. 103 p. PhD Thesis – Graduate Program in Biochemistry. Instituto de Química, Universidade de São Paulo, São Paulo.

In the mouse genome there are approximately 1000 genes that encode olfactory receptors (ORs) and 150 genes that encode type 1 vomeronasal receptors (V1Rs) dispersed in various chromosomes. Each olfactory or vomeronasal neuron selects one single allele from one single receptor gene (OR or V1R) for expression while the rest of the repertoire remains silenced. The mechanisms underlying OR and V1R gene expression are still unknown. The similarities of the pattern of expression in both types of olfactory sensory neurons suggest that the regulation of OR and V1R gene expression may be under the control of a common mechanism.

Until now, promoter regions of different OR and V1R genes had not been extensively analyzed. We carried out a comprehensive analysis in which the upstream regions of a large number of different OR and V1R genes were compared. First, using the RLM-RACE strategy, combined with degenerate PCR we generated hundreds of complete 5'UTR cDNAs for a total of 198 OR genes and 39 V1R genes. Then, these cDNAs were aligned against the mouse genome sequence and the transcription start sites (TSSs) were precisely determined. Sequences upstream of the TSSs were retrieved and searched for common DNA motifs that may play a role as *cis*-acting elements involved in the general regulation of OR and V1R gene expression.

The analysis revealed the presence of motifs that resemble O/E-like sites overrepresented in the OR and V1R promoter regions. These O/E-like motifs specifically interact with nuclear protein prepared from olfactory epithelium, but not from brain and liver.

We also identified a new motif that does not resemble any known transcription factor binding site. Besides, this new motif is present in 77% of the 39 V1R promoter regions and in 52% of the 198 OR promoter regions analyzed, preferentially concentrated near the TSSs. Interestingly, binding assays indicate that this new motif interacts with nuclear proteins prepared from the vomeronasal and the olfactory epithelia, but not from brain, liver and lung.

Our results indicate that OR and V1R genes share common promoter elements that may be binding sites for specific olfactory transcription factors and may play a role in a common mechanism of olfactory and vomeronasal gene regulation.

**Keywords:** gene expression, promoter, transcription factor, olfactory receptor, vomeronasal receptor.

## LISTA DE ABREVIATURAS E SIGLAS

cDNA: DNA complementar

CIP: fosfatase intestinal de bezerro

CSPD: Disodium 3-(4-methoxy-spiro{1,2-dioxetane-3,2'-(5'-chloro)tricyclo[3.3.1.1<sup>3,7</sup>]decan}-4-yl)phenyl phosphate

ddUTP: dideoxi-uridina trifosfato

DEPC: dietilpirocarbonato

dNTP: desoxirribonucleosídeo 5' trifosfato

DTT: ditioneitol

EDTA: ácido etilenodiaminotetracético

EGTA: ácido etilenoglicol-bis-(beta-aminoetiléter)N,N'-tetracético

FISH: hibridização in situ fluorescente

GPCRs: receptores acoplados à proteína G

kb: quilobase

M71: receptor olfatório 71 de camundongo

MOE: epitélio olfatório principal

MOR23: receptor olfatório 23 de camundongo

MOR28: receptor olfatório 28 de camundongo

NCBI: *National Center for Biotechnology Information*

OE: epitélio olfatório principal

O/E: *Olfactory/ Early B-cell factor*

OR: receptor olfatório

ORF: fase aberta de leitura

pb: pares de base

PCR: reação em cadeia da polimerase

PM: peso molecular

PMSF: fluoreto de fenilmetilsulfonila

RLM-RACE: *RNA ligase-mediated and oligo-capping rapid amplification of cDNA ends* - rápida amplificação de extremidades de cDNAs mediada pela ligação de oligo de RNA ao *cap*.

RNase: ribonuclease A

RT-PCR: transcrição reversa seguida de reação em cadeia da DNA polimerase

TAP: pirofosfatase ácida do vírus do tabaco

TSS: sítio de início da transcrição

uORF: *upstream ORF*

5'UTR: região 5' não traduzida

VNO: órgão vomeronasal

V1R: receptor vomeronasal do tipo 1/ receptor de feromônios do tipo 1.

V2R: receptor vomeronasal do tipo 2/ receptor de feromônios do tipo 2.

VR: receptor vomeronasal ou receptor de feromônio

## SUMÁRIO

<b>1) INTRODUÇÃO</b>	<b>13</b>
<b>1.1 O sistema olfatório</b>	<b>14</b>
<b>1.2 Os receptores olfatórios (ORs)</b>	<b>14</b>
1.2.1 Padrão de expressão dos receptores olfatórios	15
<b>1.3 Os receptores vomeronasais (VRs)</b>	<b>19</b>
<b>1.4 Semelhanças no padrão de expressão de genes de ORs e de V1Rs</b>	<b>22</b>
<b>2) OBJETIVOS</b>	<b>26</b>
<b>3) MATERIAL &amp; MÉTODOS</b>	<b>27</b>
<b>3.1 Método de RLM-RACE para gerar cDNAs contendo a região 5' não traduzida</b>	<b>27</b>
3.1.1 Dissecção de epitélio olfatório e órgão vomeronasal de camundongos	27
3.1.2 Extração de RNA total	27
3.1.3 Produção de cDNAs de OE e de VNO contendo a extremidade 5' intacta	28
<b>3.2 Obtenção de cDNAs correspondentes a genes de ORs e de V1Rs</b>	<b>30</b>
3.2.1 Amplificação de cDNAs com oligonucleotídeos específicos para genes de ORs e de V1Rs	30
3.2.2 Purificação dos produtos de PCR	34
3.2.3 Clonagem e determinação da orientação dos insertos	34
3.2.4 Purificação dos plasmídeos e seqüenciamento de DNA	35
<b>3.3 Análise das sequências de cDNA obtidas</b>	<b>37</b>
3.3.1 Alinhamento das sequências de cDNA obtidas	37
3.3.2 Agrupamento das sequências de cDNA em <i>clusters</i>	37
<b>3.4 Análise das sequências promotoras</b>	<b>38</b>
<b>3.5 Preparação de extratos protéicos nucleares</b>	<b>39</b>
<b>3.6 Ensaios de alteração de mobilidade eletroforética (EMSA)</b>	<b>40</b>
<b>3.7 Hibridização <i>in situ</i> em secções coronais de epitélio olfatório de camundongo</b>	<b>42</b>
3.7.1 Preparo das sondas de RNA usadas na hibridização <i>in situ</i>	42
3.7.2 Preparação das lâminas contendo cortes da cavidade nasal	42
3.7.3 Hibridização <i>in situ</i>	43
<b>4) RESULTADOS</b>	<b>45</b>
<b>4.1 Estratégia utilizada para determinar regiões promotoras de genes de ORs e V1Rs</b>	<b>45</b>
<b>4.2 Análise de regiões promotoras de genes de receptores olfatórios</b>	<b>47</b>
4.2.1 Amplificação de cDNAs correspondentes a transcritos de genes de ORs contendo a região 5'UTR intacta	47
4.2.2 Análise das sequências de cDNAs de genes de ORs obtidas por RLM-RACE	51
4.2.3 Análise das regiões promotoras dos genes de ORs	57
4.2.4 Ensaios de EMSA para avaliar o significado biológico dos motivos encontrados nos promotores de genes de ORs	63
4.2.5 Especificidade das interações dos motivos semelhantes a sítios O/E	67

4.2.6 Análise da escolha de expressão de um gene OR em relação aos motivos O/E presentes nos promotores	68
<b>4.3 Análise de regiões promotoras de genes de V1Rs</b>	<b>72</b>
4.3.1 Desenho de oligonucleotídeos degenerados e família-específicos para genes de V1R	72
4.3.2 Amplificação de cDNAs contendo a região 5'UTR completa de genes de V1R	74
4.3.3 Análise das sequências de cDNA 5'UTR completa para genes de V1R	76
4.3.4 Busca de elementos conservados na região promotora dos genes de V1R	82
4.3.5 Ensaio de EMSA com motivos encontrados nos promotores de V1Rs	85
<b>5) DISCUSSÃO</b>	<b>90</b>
<b>5.1 Motivos presentes nos promotores de genes de ORs e de V1Rs</b>	<b>90</b>
<b>5.2 Regulação da expressão de genes de ORs e V1Rs</b>	<b>95</b>
<b>6) CONCLUSÃO</b>	<b>98</b>
<b>7) REFERÊNCIAS</b>	<b>99</b>
<b>8) LISTA DE ANEXOS</b>	<b>104</b>
<b>Anexo-I: Súmula Curricular</b>	
<b>Anexo-II: Publicação resultante desta tese</b>	

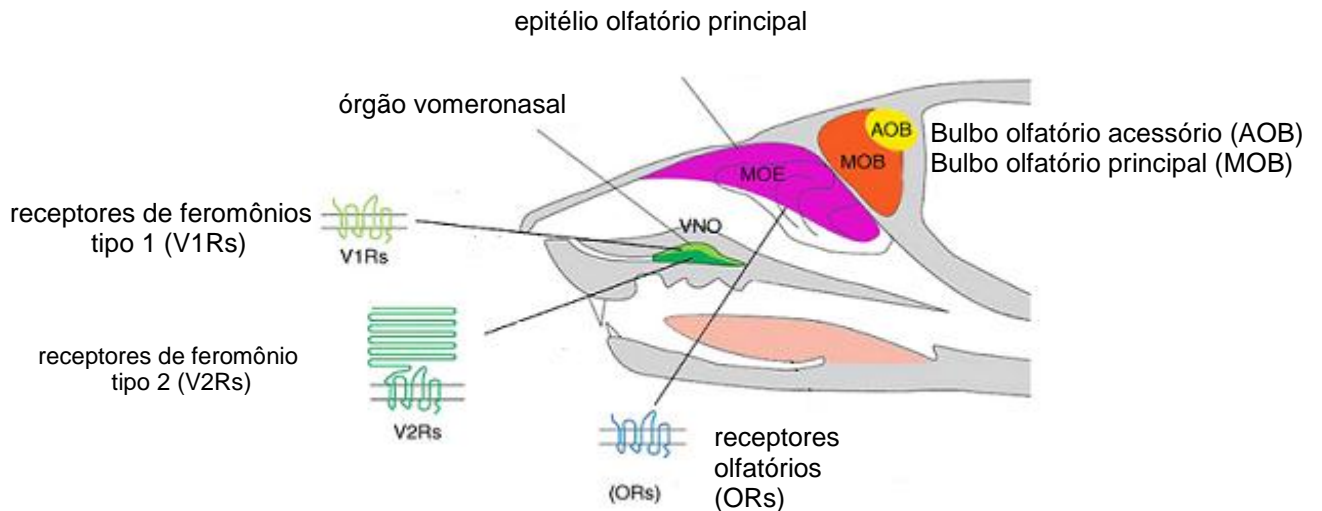
# 1) INTRODUÇÃO

## 1.1 O Sistema Olfatório

No camundongo, assim como em outros animais, o sistema olfatório desempenha um papel relevante na exploração do ambiente e no reconhecimento social e sexual de indivíduos. É através do sistema olfatório que os animais podem detectar e discriminar uma enorme quantidade de odorantes de estruturas químicas variadas e também substâncias chamadas de feromônios, secretadas por animais da mesma espécie e que desencadeiam comportamentos sociais e reprodutivos inatos, como, por exemplo, agressividade entre machos e acasalamento.

Em camundongos, o sistema olfatório é formado por dois órgãos sensoriais localizados na cavidade nasal: o epitélio olfatório principal (MOE: *Main Olfactory Epithelium*) e o órgão vomeronasal (VNO: *Vomeronasal Organ*), como está mostrado no esquema da figura 1. O MOE detecta um amplo e variado repertório de moléculas de odorantes fornecendo informações que permitem ao animal explorar o ambiente à sua volta para, por exemplo, procurar comida e evitar substâncias venenosas (Buck, 2000 e Buck, 2004). O VNO é o órgão sensorial que detecta os feromônios (Dulac & Torello, 2003; Dulac, 2006; Giorgi & Rouquier, 2007).

Os receptores que detectam os odorantes e os feromônios nos dois órgãos sensoriais (MOE e VNO) são codificados por famílias multigênicas distintas de GPCRs (receptores acoplados à proteína G) com sete domínios transmembrana putativos como apresentado no esquema da figura 1.



**Figura 1. Esquema do sistema olfatório de camundongo.** Os receptores olfatórios (ORs) são expressos nos neurônios olfatórios localizados no epitélio olfatório principal (MOE). Os receptores de feromônio do tipo 1 (V1Rs) são expressos nos neurônios localizados na camada apical do neuroepitélio que forma o órgão vomeronasal (VNO) enquanto que os receptores de feromônio do tipo 2 (V2Rs) são expressos nos neurônios localizados na camada basal do neuroepitélio. Os axônios dos neurônios olfatórios projetam-se para o bulbo olfatório principal (MOB), estabelecendo as primeiras conexões com o cérebro. Os axônios dos neurônios vomeronasais projetam-se para o bulbo olfatório acessório (AOB). Adaptado de Matsunami & Amrein, 2003.

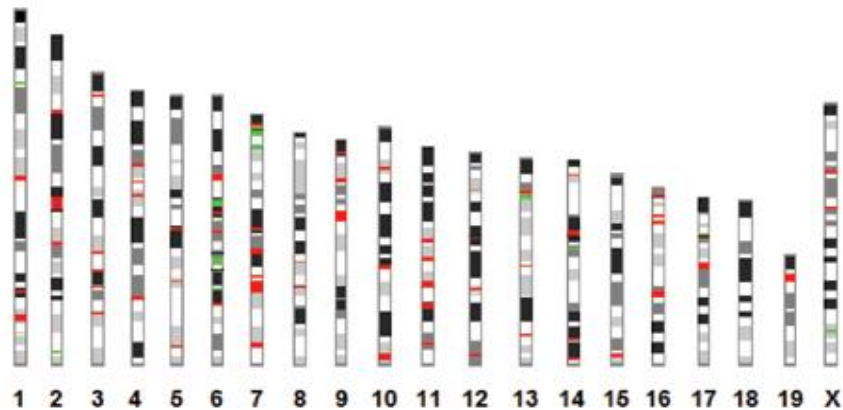
## 1.2 Os Receptores Olfatórios

Os receptores olfatórios (ORs: *olfactory receptors*) foram identificados em 1991 por Linda Buck e Richard Axel (Buck & Axel, 1991) que, por esta descoberta e por elucidar a organização do sistema olfatório, receberam o prêmio Nobel de Fisiologia ou Medicina em 2004.

Os receptores olfatórios são membros da superfamília de receptores acoplados à proteína G (GPCRs) apresentando sete domínios transmembrana putativos. Camundongos possuem por volta de 1000 genes que codificam receptores olfatórios, constituindo a maior família gênica presente no genoma, e estão organizados em *clusters* distribuídos em quase todos os cromossomos, como



mostrado na figura 2 (Young *et al.*, 2002; Zhang & Firestein, 2002; Godfrey *et al.*, 2004).

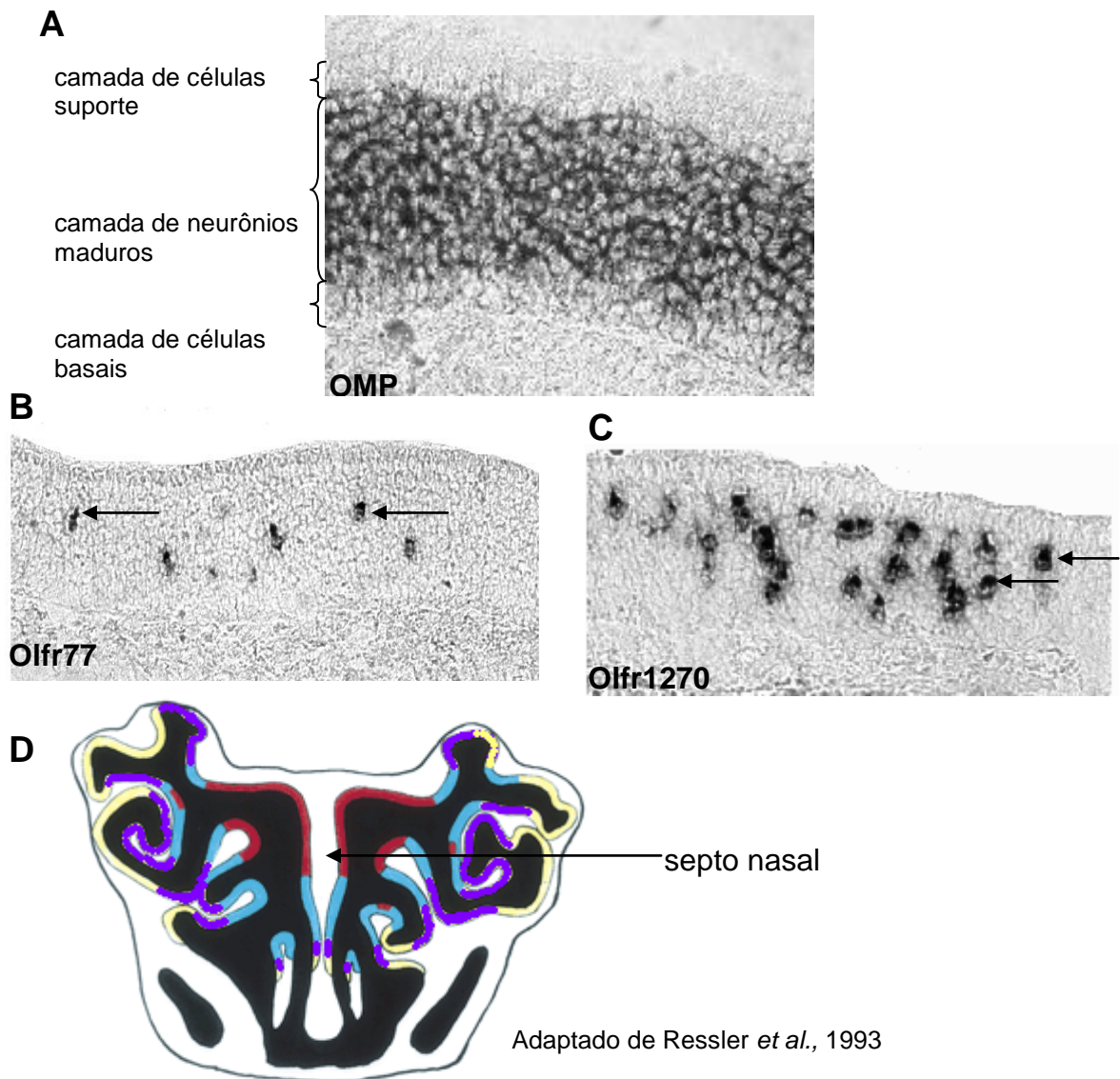


**Figura 2. Distribuição dos genes de receptores olfatórios (ORs) e receptores de feromônios do tipo 1 (V1Rs) nos cromossomos de camundongo.** Os *loci* contendo os genes de ORs (representados em vermelho) estão presentes em quase todos os cromossomos. A maioria dos genes de ORs encontra-se agrupada em *clusters* nos cromossomos 2, 7, 9 e 11 enquanto que a maioria dos genes de V1Rs (representados em verde) agrupam-se, principalmente, nos cromossomos 6, 7, 13 e 17. Não há genes de ORs e V1Rs no cromossomo Y. Adaptado de Kambere & Lane, 2007.

### 1.2.1 Padrão de expressão dos receptores olfatórios

Estudos utilizando técnicas de hibridização *in situ* (Ressler *et al.*, 1993; Vassar *et al.*, 1993) no qual vários genes de ORs foram clonados e hibridizados em secções coronais da cavidade nasal de camundongos, mostraram que, aproximadamente, um em cada 1000 neurônios olfatórios expressa um determinado OR (figuras 3B e 3C) e que neurônios que expressam o mesmo tipo de OR ocupam uma mesma zona de expressão no epitélio olfatório (figura 3D). Como existem, aproximadamente, 1000 ORs diferentes, estas observações sugerem que cada neurônio expressa somente um único tipo de receptor olfatório. Análises utilizando RT-PCR de uma única célula confirmaram que um neurônio olfatório expressa somente um único tipo de receptor olfatório (Malnic *et al.*, 1999). Além de cada

neurônio olfatório selecionar apenas um dentre mil genes de ORs para expressar, somente um único alelo do gene selecionado é transcrito (Chess *et al.*, 1994). Os mecanismos que regulam a expressão dos genes ORs ainda não são conhecidos.



**Figura 3. Expressão de receptores olfatórios no epitélio olfatório.** A-C) hibridizações *in situ* com sondas de RNA antisense em epitélio olfatório de camundongo. O gene OMP marca positivamente todos os corpos celulares da camada de neurônios maduros (A) enquanto que as sondas para os ORs Olfr77 (B) e Olfr1260 (C) marcam somente os neurônios olfatórios maduros que expressam os genes de ORs indicados, por exemplo pelas ponta das setas. D) Esquema mostrando as quatro diferentes zonas de expressão de ORs num corte coronal do nariz de camundongo. Um determinado receptor olfatório é expresso em somente uma das zonas de expressão (indicadas pelas cores vermelha, azul, roxa e amarela) e dentro desta zona um receptor é expresso em neurônios espalhados de forma aleatória.

Vários modelos têm sido propostos para explicar o padrão de expressão gênica dos receptores olfatórios. Num dos modelos sugeridos levantou-se a possibilidade de que eventos de rearranjo de DNA poderiam estar envolvidos na escolha de um gene de OR (Serizawa *et al.*, 2004 e Shykind, 2005) de maneira semelhante ao que ocorre nos genes codificadores para imunoglobulinas em linfócitos (Hozumi & Tonegawa, 1976). No entanto, experimentos onde camundongos foram clonados a partir do núcleo de neurônios olfatórios maduros indicaram que este não é o caso (Eggan *et al.*, 2004; Li *et al.*, 2004). Apesar desses camundongos terem sido gerados a partir de núcleos de neurônios olfatórios que expressavam um único gene de OR, o padrão de expressão de diferentes genes de ORs analisados no epitélio olfatório dos clones era normal. Estes resultados mostram que eventos de rearranjo de DNA não ocorrem na seleção dos genes de ORs.

Em outro modelo sugerido, uma cópia de um gene OR seria transferida, por conversão gênica, para uma única região de expressão ativa. Ishii e colaboradores (Ishii *et al.*, 2001) analisaram o núcleo de neurônios olfatórios expressando o gene de receptor olfatório MOR28 por FISH (*fluorescence in situ hybridization*). Foram usadas sondas de DNA-FISH capazes de hibridizar com ambos os alelos do gene MOR28 e sondas de RNA-FISH capazes de hibridizar com a região de um íntron do mRNA primário. A sonda de DNA-FISH detectou duas regiões genômicas no núcleo, cada uma correspondendo a um dos alelos, materno ou paterno, enquanto que a sonda de RNA-FISH detectou os transcritos primários do gene MOR28 co-localizando com somente um dos alelos do gene em questão. Se ocorresse conversão gênica, um terceiro sítio seria detectado pela sonda de DNA e esse seria o sítio transcricionalmente ativo o que não foi observado. Esses resultados reforçam

a expressão monoalélica dos genes de ORs e descartam o modelo de conversão gênica.

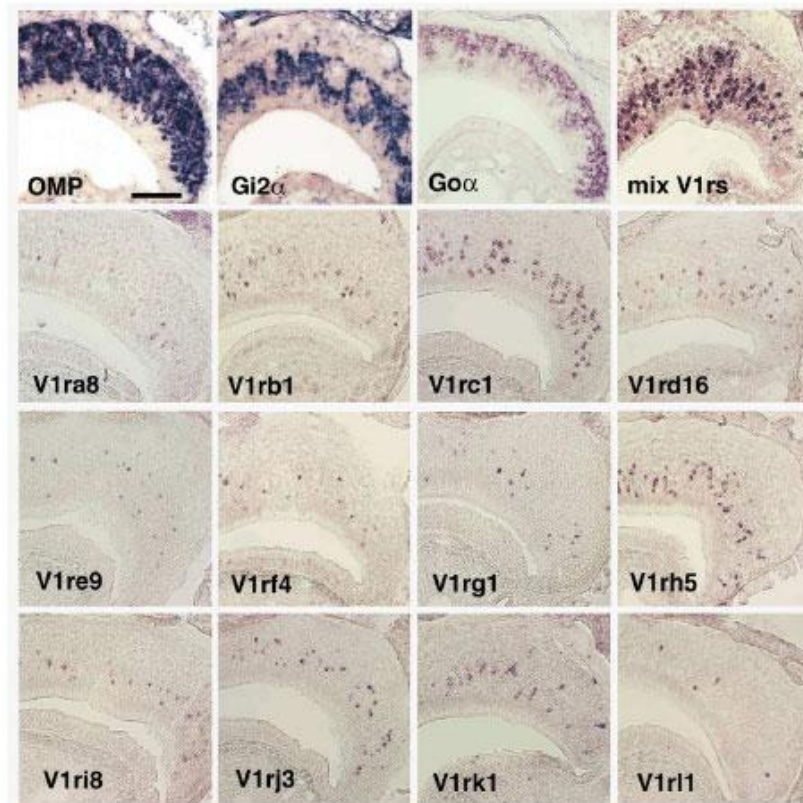
Estudos com camundongos transgênicos foram realizados com o objetivo de se determinar o menor fragmento de DNA genômico capaz de controlar a expressão de transgenes de maneira semelhante à expressão dos genes de ORs correspondentes endógenos. Foi demonstrado que regiões a 5' próximas aos sítios de início da transcrição (TSSs), com somente 395 pares de base a montante do gene de OR MOR23 e 161 pares de base a montante do gene de OR M71, e uma região genômica de 2700 pares de base a montante do gene de OR M4, possuem elementos capazes de comandar a expressão correta desses transgenes (Qasba & Reed, 1998; Vassalli *et al.* 2002; Rothmam *et al.* 2005). Isto indica que existem elementos *cis*-atuantes nos promotores proximais envolvidos na regulação da expressão dos genes ORs. No entanto, Serizawa e colaboradores (Serizawa *et al.*, 2000) mostraram que, para a expressão correta do transgene do gene de OR MOR28, é necessário um segmento de DNA genômico com mais de 70 mil pares de base a montante do sítio de início da transcrição. A comparação das seqüências genômicas a 5' da região codificadora do gene MOR28 de camundongo e do seu ortólogo humano revelou a existência de uma região homóloga com 2100 pares de base, chamada de elemento H (H referente a homólogo), localizada a 75 mil pares de base a montante do gene MOR28 (Serizawa *et al.*, 2003). Lomvardas e colaboradores (Lomvardas *et al.*, 2006) propuseram um modelo interessante mostrando que, num neurônio olfatório, o elemento reforçador H co-localiza com o gene de OR transcricionalmente ativo, mesmo quando o gene de OR ativo reside num cromossomo diferente do cromossomo onde está localizado o elemento H. Presente como cópia única no genoma e tendo um dos alelos metilado e, portanto,

inativo, foi sugerido que o elemento H age como um elemento funcionalmente monoalélico, *trans*-atuante, capaz de interagir com somente um promotor de um único gene OR por neurônio. Essa interação, mediada por proteínas, asseguraria a expressão de um único alelo de gene de OR por neurônio olfatório. Entretanto, Fuss e colaboradores (Fuss *et al.*, 2007) e Nishizumi e colaboradores (Nishizumi *et al.*, 2007) deletaram a região H em camundongos e observaram que, nos neurônios olfatórios, somente a expressão de genes de ORs do *cluster* do gene MOR28 é afetada. Esses resultados mostram que a região H atua no controle da expressão dos genes de ORs localizados em *cis* mas descarta a hipótese da região H atuar como regulador global na expressão de genes de ORs.

Em conjunto estes dados reforçam a presença de elementos *cis*-atuantes nas regiões promotoras dos genes de ORs que podem estar envolvidos nos mecanismos de controle da expressão desses genes.

### **1.3 Os Receptores Vomeronasais**

No VNO, os receptores de feromônios ou vomeronasais (VR: *vomeronasal receptors*) são codificados por duas superfamílias de genes não relacionadas, expressas em regiões distintas do neuroepitélio. Os neurônios da camada apical do neuroepitélio, como mostrado na figura 4, expressam genes de V1R e a subunidade G $\alpha$ 2 da proteína G (Dulac e Axel, 1995) enquanto que os neurônios da camada basal expressam genes de V2R e a subunidade G $\alpha$ o da proteína G (Herrada *et al.*, 1997; Matsunami *et al.*, 1997).

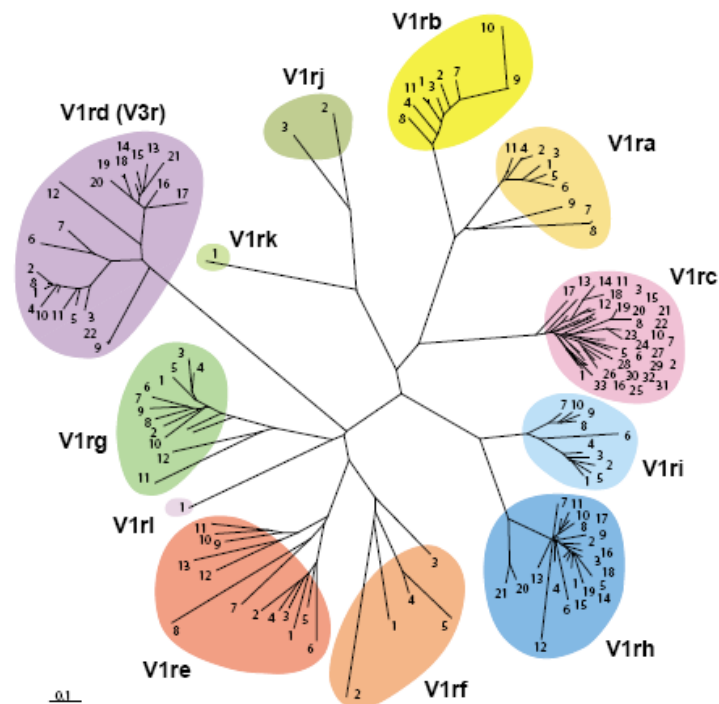


**Figura 4. Expressão de genes V1R no epitélio vomeronasal do VNO.** Hibridização *in situ* em secções coronais de VNO de camundongo utilizando sondas de RNA antisense para diferentes genes expressos. A marcação é visualizada em púrpura. A sonda de OMP marca todos os corpos celulares dos neurônios vomeronasais maduros presentes no epitélio. A sonda *Gi2* marca todos os corpos celulares dos neurônios que expressam genes de receptores V1Rs e estão na camada apical enquanto que a sonda *Go* marca todos os corpos celulares dos neurônios que expressam receptores V2Rs e estão na camada basal do epitélio. A mistura de sondas V1Rs confirma que somente os neurônios positivos para *Gi2* expressam V1Rs. Sondas para diferentes genes de V1Rs, de diferentes famílias, mostram a marcação de neurônios expressando diferentes receptores V1Rs na camada apical do VNO. Extraído de Rodriguez *et al.*, 2002.

Assim como os genes que codificam para receptores olfatórios, os genes de V1Rs são codificados por um único éxon. Ao contrário, os genes de V2R são codificados por múltiplos éxons. A presença de um único éxon codificador facilitou a pesquisa *in silico* de genes de V1Rs no genoma. O repertório de genes de V1Rs em camundongo foi estimado em aproximadamente 150 genes funcionais (Rodriguez *et al.*, 2002; Zhang *et al.*, 2004; Young *et al.*, 2005). Assim como os genes de ORs, os

genes de V1Rs estão distribuídos em *clusters* em vários cromossomos, como também está indicado na figura 2.

Os genes de V1R são classificados, conforme identidade de sequência de aminoácidos, em 12 famílias que apresentam grande diversidade entre si (Rodriguez *et al.*, 2002; Yang H. *et al.*, 2005) como mostrado na figura 5.



**Figura 5. Diversidade das famílias de genes V1Rs.** Um total de 137 genes de V1Rs foram agrupados conforme identidade de sequências de aminoácidos em 12 famílias nomeadas de a-l. Os membros anteriormente classificados como V3r foram alinhados como membros da família V1rd. Adaptado de Rodriguez *et al.*, 2002.

Existem poucos trabalhos publicados a respeito de sequências regulatórias nos genes de V1Rs. Análises comparativas de regiões genômicas de um único *cluster* de genes V1Rs sugeriram a presença de regiões a 5' bem conservadas (Lane *et al.*, 2002). O alinhamento de quatro promotores de genes V1Rs do mesmo *cluster*, obtidos experimentalmente por RACE-PCR, contra os genomas de camundongo e rato revelou que todos os genes de V1Rs que formam o *cluster* estudado possuem

duas regiões conservadas: um bloco conservado na região 5'UTR e outro bloco com 1000 pares de base conservado a uma distância de 5000 pares de base a montante das regiões codificadoras (Lane *et al.*, 2004). No entanto, a conservação de uma região tão grande (de 1000 pares de base) e contínua provavelmente não indica somente a conservação de sítios de ligação para fatores de transcrição individuais. A ampla identidade entre regiões promotoras de genes de um mesmo *cluster* pode ser, parcialmente, devido a eventos de conversão gênica responsáveis pela expansão do repertório de genes de V1Rs. A busca de seqüências conservadas em regiões promotoras putativas através do alinhamento de algumas poucas seqüências promotoras de genes de V1Rs do mesmo *cluster* contra o genoma pode não revelar a presença de motivos de ligação para fatores de transcrição comuns a regiões promotoras de diferentes genes de V1Rs envolvidos no controle geral da expressão dos genes de V1Rs.

Diferentemente, a busca de motivos conservados em regiões promotoras de diferentes genes de V1Rs pertencentes a diferentes famílias, é uma estratégia que pode revelar a presença de sítios de ligação de fatores de transcrição candidatos a atuarem na regulação geral da expressão dos genes de V1Rs.

#### **1.4 Semelhanças no padrão de expressão de genes de ORs e de V1Rs**

Apesar da falta de identidade em termos de seqüência, existem muitas semelhanças no padrão de expressão dos genes de ORs e dos genes de V1Rs:

- Tanto os ORs como os V1Rs, formam famílias multigênicas de GPCRs expressas exclusivamente nos neurônios sensoriais especializados do epitélio olfatório principal e do órgão vomeronasal, respectivamente.

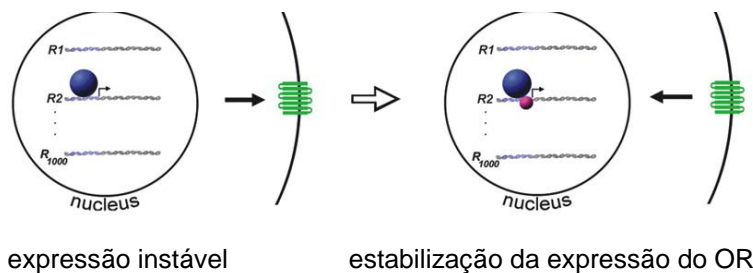


- Os genes de ORs e de V1Rs não apresentam íntrons nas regiões codificadoras e estão distribuídos, no genoma, agrupados em *clusters* localizados em diferentes cromossomos (Godfrey *et al.*, 2004; Young *et al.*, 2005).
- Neurônios olfatórios e vomeronasais selecionam um único alelo de um único gene de receptor OR ou de um único gene de V1R, respectivamente, para expressar (Malnic *et al.*, 1999; Chess *et al.*, 1994; Rodriguez *et al.*, 1999; Roppolo *et al.*, 2007).
- Foi demonstrado que a expressão monoalélica de um gene de OR é regulada por um mecanismo de retroalimentação negativo que requer um OR funcional expresso na membrana (Lewcock & Reed 2004; Serizawa *et al.*, 2003) e que a expressão de um gene de OR não funcional (pseudogene) permite que o neurônio olfatório escolha outro gene de OR funcional para expressar (Shykind *et al.*, 2004). O mesmo mecanismo foi demonstrado nos neurônios vomeronasais que expressam genes de V1Rs (Roppolo *et al.*, 2007). A figura 6 mostra um esquema do mecanismo de *feedback* que ocorre na expressão dos genes de ORs.

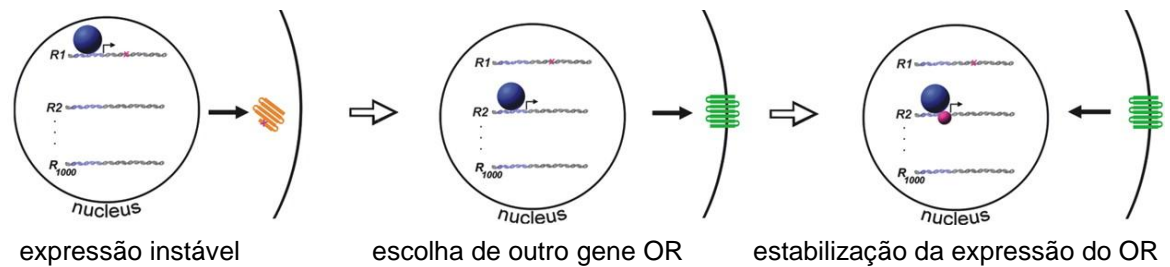
As semelhanças de expressão de genes de ORs e de V1Rs levantam a possibilidade de que o mecanismo de regulação da expressão desses genes possa ser comum. Foi proposto um modelo para a regulação da expressão dos genes de ORs no qual, após um alelo de um gene OR ser estocasticamente selecionado para expressão por um fator limitante presente no núcleo dos neurônios olfatórios, o receptor olfatório funcional na membrana é responsável pelo *feedback* que resulta na manutenção da expressão do receptor escolhido (Serizawa, 2004; Shykind, 2005) como mostrado no esquema da figura 6. O modelo proposto sugere que esse fator limitante no núcleo dos neurônios olfatórios reconhece elementos comuns presentes

na região promotora dos genes de ORs. Devido as semelhanças no padrão de expressão, esses elementos reguladores podem ser comuns aos genes de ORs e de V1Rs.

**A** O receptor olfatório funcional, expresso na membrana, sinaliza, via um mecanismo de retroalimentação negativo, estabilizando a escolha do gene OR expresso:



**B** Escolha de um receptor olfatório não funcional leva à seleção de outro gene OR:



**Figura 6. Modelo de seleção e manutenção de expressão de um gene OR em neurônios olfatórios.** A) Escolha de um receptor funcional. Um gene OR é selecionado por um mecanismo estocástico representado pela esfera azul. A proteína funcional de receptor olfatório (em verde) é responsável pelo mecanismo de retroalimentação (simbolizado pela esfera vermelha) que estabiliza a escolha do receptor expresso na membrana do neurônio olfatório. B) Expressão de um receptor não funcional. Um receptor olfatório não funcional é selecionado pelo mecanismo estocástico de escolha (esfera azul). A proteína de receptor de OR não funcional, representada em vermelho, falha na sinalização que estabiliza a escolha do OR e o mecanismo de seleção continua ativo. Outro gene de OR é escolhido até que um receptor funcional seja expresso na superfície da membrana e sinalize para inibir a escolha de outro gene OR. Adaptado de Shykind, 2005.

Uma forma de investigar a presença de elementos comuns envolvidos na expressão de genes co-regulados, como os genes de ORs e de V1Rs, é comparar diversas regiões promotoras de diferentes genes em busca de motivos conservados.

Como será apresentado em nossos resultados, determinamos, experimentalmente, a localização dos sítios de início da transcrição (TSSs) de vários genes de ORs e de V1Rs e pesquisamos as regiões a 5' dos TSSs em busca de motivos conservados potencialmente envolvidos na regulação da expressão dos genes de ORs e de V1Rs. Identificamos elementos conservados nos promotores analisados. Ensaios de EMSA mostram que estes motivos de DNA interagem com proteínas nucleares de OE e VNO indicando um potencial papel biológico.

## 2) OBJETIVOS

Conforme já apresentado, cada neurônio sensorial que expressa receptores olfatórios e receptores de feromônios do tipo 1, presente no epitélio olfatório principal (MOE) e no epitélio apical do órgão vomeronasal (VNO), seleciona um único alelo de um único gene de OR ou de V1R, respectivamente, para expressar. Os mecanismos que regulam a expressão dos genes de ORs e de V1Rs não estão esclarecidos.

O objetivo deste trabalho foi investigar a presença de potenciais elementos regulatórios em regiões promotoras de genes de ORs e genes de V1Rs. Para isso, realizamos uma análise na qual regiões a 5' de um grande número de genes de ORs e de genes de V1Rs foram obtidas experimentalmente e pesquisadas utilizando ferramentas de bioinformática com o objetivo de identificar elementos comuns potencialmente envolvidos na regulação da expressão desses genes. Alguns dos motivos identificados foram analisados quanto à sua capacidade de interagir com proteínas presentes no núcleo de neurônios olfatórios ou vomeronasais.

### **3) MATERIAL & MÉTODOS**

#### **3.1 Método de RLM-RACE para gerar cDNAs contendo a região 5' não traduzida**

Este trabalho foi realizado utilizando-se o sistema “GeneRacer™ RLM-RACE” (*RNA ligase-mediated and oligo-capping rapid amplification of cDNA ends*). Um esquema explicativo da técnica está apresentado na figura 7 do item 4-Resultados. Todo o procedimento foi realizado em condições livres de RNase.

##### **3.1.1 Dissecção de epitélio olfatório e órgão vomeronasal de camundongos**

Epitélio olfatório (OE) e órgão vomeronasal (VNO) de 10 camundongos da linhagem C57BL/6J, idade entre 5 e 7 semanas e ambos os sexos, foram dissecados e imediatamente submetidos a extração de RNA total.

##### **3.1.2 Extração de RNA total**

Os tecidos foram homogeneizados em 1,0 mL de reagente TRIzol (Invitrogen) /50-100 mg de tecido utilizando homogeneizador (Polytron). Após incubação por 5 minutos à temperatura ambiente para permitir a completa dissociação de complexos de nucleoproteína, foi adicionado 0,2 mL de clorofórmio para cada 1 mL de TRIzol às amostras homogeneizadas seguida de agitação vigorosa por 15 segundos e incubação por 2-3 minutos à temperatura ambiente. As amostras foram centrifugadas a 12.000Xg por 15 minutos a 4°C e o RNA presente na fase aquosa foi transferido para um novo tubo Falcon onde foi adicionado isopropanol (0,5 mL/ 1 mL de TRIzol usado na homogeneização inicial). Após incubação de 10 minutos e centrifugação a 10.000Xg por 20 minutos a 4°C, o sobrenadante foi eliminado restando o RNA precipitado. O RNA precipitado foi lavado com etanol a 75% (1ml de Etanol/ 1ml de TRIzol usado inicialmente) e centrifugado a 7.500Xg por 5 minutos a

4°C. O precipitado de RNA foi seco à temperatura ambiente, ressuspendido em 500 µL de água tratada com DEPC (Invitrogen).

A quantidade e qualidade das amostras de RNA total de OE e de VNO foram analisadas espectrofotometricamente e por eletroforese em gel de agarose-formaldeído. As amostras de RNA total foram estocadas a -70°C.

### **3.1.3 Produção de cDNAs de OE e de VNO contendo a extremidade 5' intacta**

#### **Desfosforilação do RNA**

A reação de desfosforilação da amostra de RNA total se deu nas seguintes condições:

1,5µl RNA total (2,7µg/µL)

1µL Tampão10X CIP (0,5M Tris-HCl, pH8,5; 1mM EDTA)

1µL RNaseOut (40 U/µL)

1µL CIP (10 U/µL)

5,5 µL água tratada com DEPC

A mistura foi incubada a 50°C em termociclador Eppendorf por 1 hora e, a seguir, mantida no gelo.

#### **Precipitação do RNA**

Adicionou-se 90 µL de água tratada com DEPC (Invitrogen) à mistura da reação anterior de defosforilação do RNA e,então, 100 µL de fenol:clorofórmio seguido de agitação vigorosa por 30 segundos. A amostra foi centrifugada a 16000Xg por 5 minutos à temperatura ambiente. A fase superior (fase aquosa) foi transferida para um novo tubo do tipo eppendorf e adicionou-se 2 µL de glicogênio, 10 µL de acetato de sódio 3M no pH 5,2 e, após misturar bem, adicionou-se 220 µL

de etanol 95%. A amostra foi vortexada brevemente e mantida em gelo por 10 minutos. A seguir, a amostra foi centrifugada a 20000Xg por 20 minutos a 4 °C. O sobrenadante foi cuidadosamente removido com auxílio de uma micropipeta. O precipitado foi lavado com 500 µL de etanol 70% e centrifugado a 20000Xg por 2 minutos a 4°C. Novamente, o sobrenadante foi removido. Nova centrifugação para coletar etanol remanescente. O precipitado foi deixado por cerca de dois minutos secando à temperatura ambiente e foi ressuscitado em 7 µL de água tratada com DEPC sendo usado no próximo passo.

### **Remoção do 5' *cap* dos RNAs mensageiros**

Aos 7 µL de RNA desfosforilado e precipitado anteriormente foram adicionados 1 µL de RNaseOut (40U/µL), 1µl de tampão TAP 10X (0,5M acetato de sódio, pH6,0; 10mM EDTA; 1% β-mercaptoetanol; 0,1% Triton X-100) e 1 µL de TAP (0,5U/µL). A reação foi misturada gentilmente e incubada a 37°C no termociclador Eppendorf por 1 hora. A seguir, a reação foi colocada no gelo e precipitada exatamente como descrito acima. O precipitado de RNA foi ressuscitado em 7 µL de água tratada com DEPC e usado no próximo passo.

### **Ligação do oligonucleotídeo de RNA aos RNAs mensageiros sem cap**

Os 7 µL de RNA sem *cap* anteriores foram colocados em um tubo que continha o oligonucleotídeo GeneRacer RNA liofilizado (0,25 µg). O oligonucleotídeo foi ressuscitado com o RNA pipetando a mistura algumas vezes. A mistura foi incubada a 65°C por 5 minutos para permitir o relaxamento das estruturas secundárias do RNA. A mistura foi colocada no gelo por 2 minutos e brevemente centrifugada. Adicionou-se ao tubo de reação: 1µL de tampão de ligase 10X (330mM

Tris-acetato, pH 7,8; 660mM acetato de potássio; 100mM acetato de magnésio; 5mM DTT); 1µL de 10 mM ATP; 1µL de RNaseOut (40U/µL) e 1µL de T4 RNA ligase (5U/µL). A reação foi incubada a 37°C por 1 hora. Colocada no gelo e novamente precipitada (como já descrito). O precipitado de RNA foi ressuspenso em 10 µL de água tratada com DEPC e usado no próximo passo.

### **Reação de transcrição reversa do mRNA (síntese de cDNA)**

Adicionou-se 1µL de oligo dT e 1µL de mistura de dNTP (25 mM cada) à amostra preparada no item anterior. A reação foi incubada a 65°C por 5 minutos, resfriada no gelo por 2 minutos e centrifugada brevemente. A seguir, adicionou-se:

4µL de 5X tampão RT (250 mM Tris-HCl (pH 8.3), 375 mM KCl, 15 mM MgCl<sub>2</sub>)

2µL 0,1M DTT

1µL Superscript II RT (200 U/µL)

1µL RnaseOut

A reação foi misturada e incubada a 42°C por 50 minutos. Para inativar a transcriptase reversa, a reação foi incubada a 70°C por 15 minutos; a seguir, a reação foi resfriada no gelo por 2 minutos. Adicionou-se 1µL de RNase H (2U) incubando a 37 °C por 20 minutos.

## **3.2 Obtenção de cDNAs correspondentes a genes de ORs e V1Rs**

### **3.2.1 Amplificação de cDNAs com oligonucleotídeos específicos para genes de ORs e de V1Rs**

As amostras de cDNAs contendo as regiões 5' dos mRNAs correspondentes intactas foram então usadas em reações de PCR com oligonucleotídeos degenerados e específicos que reconhecem seqüências conservadas entre



diferentes membros das famílias gênicas de receptores olfatórios (ORs) e de receptores de feromônios do tipo 1 (V1Rs).

Na amplificação de cDNAs 5' completos para genes de ORs, foi feita, primeiramente, uma reação de PCR utilizando-se o oligonucleotídeo degenerado P8 (oligonucleotídeo reverso para a região transmembrânica VII de ORs) ou o oligonucleotídeo degenerado P27 (*primer* reverso para a região transmembrânica VI de ORs) juntamente com o oligonucleotídeo forward GeneRacer 5'. A seguir, foi realizada uma PCR-*Nested* utilizando como alvo a primeira PCR e os oligonucleotídeos degenerados P27 (quando a primeira PCR foi com P8) ou P26R (quando a primeira PCR foi com P27).

Os oligonucleotídeos degenerados utilizados para seleção de genes de ORs foram:

**P8 (TM-VII): (GA)TTIC(TG)IA(AG)I(GC)(TA)(GA)TAIAT(AG)AAIGG(GA)TT**

**P27 (TM-VI): ACIACIGAIAG(GA)TGIGAI(GC)C(GA)CAIGT**

**P26R (TM-III): CAIATIGCIAC(AG)TAICG(GA)TCAIGTAIGC**

Na amplificação de cDNAs 5' completos para genes de V1Rs foi feita, primeiramente, uma reação de PCR com oligonucleotídeos degenerados para receptores de feromônios (V1Rs). Os oligonucleotídeos foram obtidos através do alinhamento de 130 sequências de genes de V1Rs (Rodriguez et al.,2002) disponíveis no banco de dados do NCBI (<http://www.ncbi.nlm.nih.gov/>) e utilizando a ferramenta ClustalW ([www.ebi.ac.uk/clustalw](http://www.ebi.ac.uk/clustalw)). A região mais conservada entre os genes de V1Rs analisados foi escolhida como alvo para os seguintes oligonucleotídeos degenerados:

**VNR1: GGI(C/G)(A/T)IA(G/T)IGTGA(C/T)IGC(C/T)TGIA(A/C)IAC(A/G)CTIAG**

**VNR2: A(C/G)(A/G)CAIGTIG(A/T)(A/G)CAIA(G/T)IGAI(A/G)(G/T)IC(C/T)IC(G/T)**

Para ampliar um grande número de diferentes cDNAs de genes de V1Rs também foram utilizados oligonucleotídeos reversos família-específicos que reconhecem regiões conservadas dentro de algumas das 12 famílias em que se agrupam, por identidade de seqüências, os genes de V1Rs. As condições de reação foram semelhantes às usadas para os oligonucleotídeos degenerados.

As seqüências dos oligonucleotídeos reversos família-específicos são:

V1RA-1 5' GAGACCAATGGGCAGGTC 3'  
 V1RA-2 5' GCCTGAGATCCCAATGCC 3'  
 V1RB-1 5' CCAAAGGACATTCAGCAG 3'  
 VR1B-2 5' CAGCTATGAGTCCCATAG 3'  
 V1RD-1 5' GTTATTTAAGACTGAAG 3'  
 V1RD-2 5' GTATGCAAGTTTACATTTG 3'  
 V1RE-1 5' GGTGCATGAGAATCAAATC 3'  
 V1RE-2 5' GAAATTTCCAGAATTCC 3'  
 V1RF-1 5' CAATTTGCATCCAATATC 3'  
 V1RF-2 5' GAATCCAGTCTATAGAC 3'  
 V1RG-1 5' GTCAAGTGCTCTATAATC 3'  
 V1RG-2 5' GAATTGCCAAGCATTCCC 3'  
 V1RJ-1 5' CTGGGAGTGATGGTGATG 3'  
 V1RJ-2 5' GATTGTCAACATATTGAC 3'  
 V1RL-1 5' GACTTGGAATTGTTGAGG 3'  
 V1RL-2 5' CCAGTGAAGTCAGCGATG 3'

Condições da reação para os PCRs:

reagente	concentração inicial	concentração final	volume
água estéril			17,8 $\mu$ l
tampão de PCR	10X	1X	2,5 $\mu$ l
MgCl <sub>2</sub>	50 mM	1,5 mM	0,75 $\mu$ l
dNTP's	10 mM	0,2 mM	0,5 $\mu$ l
oligonucleotídeo foward	10 $\mu$ M	0,5 $\mu$ M	1,25 $\mu$ l
oligonucleotídeo reverse	50 $\mu$ M	2,0 $\mu$ M	1,0 $\mu$ l
TaqDNA polimerase Platinum (Invitrogen)	5U/ $\mu$ l	~1U	0,25 $\mu$ l
cDNA obtido por RLM-RACE			1,0 $\mu$ l
volume total			25 $\mu$ l

Termociclador da marca Eppendorf foi programado do seguinte modo:

- 1) 95°C por 2 minutos
  - 2) 95°C por 1 minuto
  - 3) 50°C por 3 minutos
  - 4) 72°C por 2 minutos
- 40 ciclos de repetição dos passos 2 a 4
- 5) 72°C por 10 minutos

A reação de PCR foi analisada por eletroforese em gel de agarose 1,5% e a amostra armazenada a -20°C.

*Nested-PCR:*

Uma segunda reação de PCR, utilizando como molde 1 $\mu$ L da reação de PCR anterior, diluída 200X, foi feita com oligonucleotídeos específicos para genes de ORs e para os genes de V1R que anelam em regiões mais internas que os oligonucleotídeos utilizados no primeiro PCR juntamente com o oligonucleotídeo *GeneRacer 5' Nested*. Apenas o número de ciclos das condições de PCR anteriores foi alterado para 25.

A reação de PCR foi analisada por eletroforese em gel de agarose 1,5% e a amostra armazenada a  $-20^{\circ}\text{C}$ .

### **3.2.2 Purificação dos produtos de PCR**

Os produtos do *Nested*-PCR foram resolvidos em gel de agarose 1,5%. Para os produtos de genes de ORs, a distância calculada entre o ATG predito e a região reconhecida pelos os oligos P8 e P27, é 780 pares de base, então, somente os produtos >800pb foram gel purificados. Para o oligo P26R, a distância entre o ATG predito e a região reconhecida pelo oligo P26R, é 390 pares de base, então, somente os produtos >400pb foram gel purificados.

Como a distância entre o ATG inicial dos genes de V1Rs e a região reconhecida pelos oligonucleotídeos degenerados VNR1 e VNR2 é aproximadamente 400pb somente produtos de PCR >400pb foram gel purificados.

### **3.2.3 Clonagem e determinação da orientação dos insertos**

Os fragmentos de PCR gel purificados foram subclonados no plasmídeo pCR II<sup>®</sup> (Invitrogen).

Para a seleção de cDNAs provenientes de genes de ORs:

As colônias de bactéria *E coli* DH5 $\alpha$ , contendo seqüências de cDNA de tamanho >800pb ou >400pb foram selecionadas por PCR de colônia usando as combinações de oligos SP6 – P27 e T7 – P27 ou SP6 – P26R e T7 – P26R, respectivamente.

Para a seleção de cDNAs provenientes de genes de V1Rs:

As colônias de bactéria *E coli DH5 $\alpha$* , contendo seqüências de cDNA de tamanho >400pb foram selecionadas por PCR de colônia usando as combinações de oligos SP6 – VNR1 e T7 – VNR1.

As condições de ciclagem para a reação de PCR foram:

- 1) 95°C por 1 minuto
  - 2) 50°C por 1 minuto
  - 3) 72°C por 1 minuto
- 25 ciclos de repetição dos passos 1 a 3
- 4) 72°C por 10 minutos

### **3.2.4 Purificação dos plasmídeos e seqüenciamento de DNA**

O DNA plasmidial foi purificado das colônias de bactéria positivas por lise alcalina das bactérias seguida por filtração em placas *Multiscreen (Millipore MAGV N22)* e precipitação com isopropanol. A seguir descrevemos o método utilizado:

Colônias isoladas de bactérias foram incubadas em 1 ml de meio LB com ampicilina (100 $\mu$ g/ml) em placas de crescimento de 96 poços e mantidas a 37°C, sob agitação constante (200 rotações por minuto), por 20 horas. A seguir, as bactérias foram centrifugadas a 14000Xg por 8 minutos à temperatura ambiente e o sobrenadante foi desprezado. As bactérias foram, então, ressuspensas, utilizando vórtex, em 240 $\mu$ l de tampão GET (0,01% glicose; 0,01M EDTA; 0,03M Tris HCl pH 7,4) e novamente centrifugadas nas mesmas condições anteriores. O sobrenadante foi desprezado e as bactérias foram ressuspensas em 85 $\mu$ l de solução de RNase (0,5mg/ml RNase; 10mM Tris pH 7,5; 15mM NaCl). Após, adicionou-se 60 $\mu$ l de uma mistura de NaOH (0,2M) e SDS (0,01%). A solução foi gentilmente misturada. Em seguida, foi adicionado 60 $\mu$ l de solução gelada 3M de acetato de potássio e a

amostra foi incubada por 10 minutos à temperatura ambiente. A amostra foi centrifugada por 1 minuto a 200Xg e incubada por 30 minutos a 90°C em estufa. Em seguida, a amostra foi deixada por 15 minutos no gelo e, então centrifugada a 3000Xg por 15 minutos à temperatura de 4°C. Um volume de 110µl de isopropanol na temperatura ambiente foi adicionado em cada poço de uma placa de 96 poços com fundo em V. Sobre essa placa de fundo em V (placa coletora) foi colocado o filtro adaptado para placas de 96 poços Millipore MAGV N22 (placa filtrante). Após centrifugação da amostra um volume de 100µl foi transferido para a placa filtrante e todo o sistema (placa filtrante+placa coletora com isopropanol) foi centrifugado a 3000XG, por 40 minutos, a 4°C. A placa coletora foi mantida na posição invertida sobre papel absorvente por 1 minuto. Adicionou-se 200µl de etanol 70% gelado na placa coletora e esta foi centrifugada a 3000Xg por 10 minutos a 4°C. O sobrenadante foi descartado e a placa coletora foi centrifuga na posição invertida por 1 minuto na velocidade de 200Xg. A placa coletora foi mantida por 30 minutos em fluxo laminar à temperatura ambiente. Os plasmídeos foram ressuspensos em 80µl de água deionizada livre de DNase ou RNase (*Gibco*) com agitação suave por 10 minutos a 37°C.

Os plasmídeos foram seqüenciados utilizando-se os oligos SP6 ou T7 e reagentes Big Dye (Applied Biosystems) em máquina de sequenciamento automático ABI PRISM 3100 Genetic Analyzer (Hitachi).

### 3.3 Análise das seqüências de cDNA obtidas

Utilizamos a ferramenta BLAST disponível na página do NCBI (<http://www.ncbi.nlm.nih.gov/>) para definir quais membros da família de genes ORs e de genes V1Rs obtivemos.

O alinhamento das seqüências com o BLAST e sim4, assim como o agrupamento das seqüências em *clusters* foi realizado pelo aluno de doutorado Pedro Galante, que também participou deste projeto.

#### 3.3.1 Alinhamento das seqüências de cDNA obtidas

Utilizando a ferramenta BLAST (disponível na página do NCBI <http://www.ncbi.nlm.nih.gov>) as seqüências de cDNA foram alinhadas com a seqüência genômica de camundongo. Somente os alinhamentos com identidade maior que 93% foram considerados. A posição em que cada cDNA alinhou no genoma foi determinada e um intervalo de 50 kb de seqüência genômica flanqueantes à cada seqüência de cDNA alinhada foram extraídos dos contigs genômicos correspondentes. Cada seqüência foi novamente alinhada com sua correspondente seqüência genômica extraída usando o programa sim4 (Florea *et al.*, 1998). Os alinhamentos do sim4 com identidade maior que 93%, aqueles em que mais de 50% da seqüência estava totalmente alinhada, foram selecionados.

#### 3.3.2 Agrupamento das seqüências de cDNA em *clusters*

As seqüências de cDNA foram agrupadas em 198 *clusters* baseados nas coordenadas genômicas de cada seqüência. Seqüências que apresentavam ao menos um mesmo éxon ou íntron foram incluídas no mesmo *cluster*. Quando

nenhum íntron ou éxon foi definido, seqüências com ao menos 30 pb de sobreposição na mesma região genômica foram incluídas no mesmo *cluster*.

### 3.4 Análise das seqüências promotoras

Determinou-se a localização exata dos TSSs para os 198 genes de ORs e extraiu-se seqüências de 600 pares de base a 5' dos TSSs. As seqüências promotoras dos 198 genes de ORs foram analisadas usando três diferentes programas que utilizam diferentes algoritmos de procura de motivos conservados em seqüências não alinhadas. A tabela a seguir resume as informações sobre cada programa:

Programa	Parâmetro modificado/ comprimento dos motivos pesquisados	Página disponível na Web
Gibbs recursive sampler	6,8,10,12,14	<a href="http://bayesweb.wadsworth.org/gibbs/gibbs.html">http://bayesweb.wadsworth.org/gibbs/gibbs.html</a>
Consensus	6,8,10,12	<a href="http://bifrost.wustl.edu/consensus">http://bifrost.wustl.edu/consensus</a>
Weeder	6,8,10,12	<a href="http://159.149.109.16/Tool/ind.php">http://159.149.109.16/Tool/ind.php</a>

Os demais parâmetros foram mantidos nos valores padrão para eucariotos.

Para os genes de V1Rs, foram obtidas 39 seqüências de cDNA com a região 5'UTR intacta. Extraiu-se seqüências de 1000 pares de base a 5' dos TSSs determinados para as 39 seqüências e estas regiões promotoras foram analisadas utilizando-se os mesmos programas e parâmetros usados na análise dos promotores de genes de ORs, conforme descrito acima.



### 3.5 Preparação de extratos protéicos nucleares

Preparamos extratos nucleares conforme método descrito por Kudrycki *et al.*, 1993, com algumas modificações.

Dissecamos, separadamente, epitélio olfatório e órgão vomeronasal de 30 camundongos C57BL/6J com idade entre 4 a 7 semanas e de ambos os sexos. Após lavagens com 1X PBS para retirar o muco, homogeneizamos os tecidos em 4 volumes de tampão de homogeneização (0,25M sacarose; 15mM Tris-HCl pH7,9; 60mM KCl; 15mM NaCl; 5mM EDTA; 1mM EGTA; 0,5 mM espermidina; e os inibidores de proteases adicionados antes do uso: 1mM DTT; 0,1mM PMSF; 2 $\mu$ g/ml leupeptina; 0,015U/ml aprotinina) com homogeneizador Polytron. Deixamos decantando por minutos para que a parte óssea se depositasse no fundo do tubo Falcon. Coletamos o sobrenadante e centrifugamos 1000Xg por 10 minutos, a 4°C. As células foram ressuspensas em 4 volumes de tampão hipotônico (10mM HEPES pH7,9; 1,5mM MgCl<sub>2</sub>; 10mM KCL; 0,5mM DTT; 0,05mM PMSF; 1,5 $\mu$ g/ml leupeptina; 0,006U/ml aprotinina) e homogeneizadas em homogeneizador de vidro (Dounce). As células foram incubadas no gelo por 10 minutos para permitir a lise celular. A seguir, centrifugamos a amostra a 25000Xg por 10 minutos a 4°C. O precipitado de núcleos foi ressuspensado em 1 volume do tampão hipertônico (20mM HEPES pH7,9; 25% glicerol; 1,5mM MgCl<sub>2</sub>; 1mM EDTA; 1mM DTT; 0,1mM PMSF; 2 $\mu$ g/ml leupeptina; 0,015U/ml aprotinina) e, a seguir, foi adicionado lentamente 1 volume de 1M KCl. Incubamos a amostra sob agitação lenta no gelo por 30 minutos para extração das proteínas nucleares para o sobrenadante. Centrifugamos a 25000Xg por 30 minutos a 4°C e coletamos o sobrenadante que é o extrato nuclear. Concentramos o extrato usando a coluna *Microcon YM-3 centrifugal filter device*

(Millipore). Trocamos o tampão hipertônico pelo tampão de ligação (10mM Tris-HCl, pH7,9; 10% glicerol, 3mM DTT, 0,3mM PMSF) utilizando a coluna *Micro Bio-Spin P-6 chromatography column* (Bio-Rad). Determinamos a concentração de proteínas no extrato nuclear pelo método de Bradford (Bio-Rad).

### **3.6 Ensaios de alteração de mobilidade eletroforética (EMSA : *Electrophoretic Mobility Shift Assay*)**

Utilizamos o sistema (*DIG*) *gel shift* (*Roche Applied Science*) para marcação dos oligonucleotídeos dupla-fita com digoxigenina conforme as instruções do fabricante.

As reações foram feitas no tampão contendo 20 mM HEPES, pH 7.6, 1 mM EDTA, 10 mM  $(\text{NH}_4)_2\text{SO}_4$ , 1mM DTT, Tween 20 0.2%, 30 mM KCl, 2  $\mu\text{g}$  poly[d(I-C)], 0.1  $\mu\text{g}$  poly L-lysine, 1.2 ng oligonucleotídeo marcado com digoxigenina e 15  $\mu\text{g}$  de extrato nuclear de proteínas. Após 10 minutos de incubação no gelo, e 15 minutos de incubação a temperatura ambiente, adicionamos 5  $\mu\text{l}$  de *loading buffer* (0.25 x TBE buffer 60 %, glicerol 40 %, azul de bromofenol 0.2 %) e submetemos as amostras a eletroforese em tampão 0.5 x TBE num gel não denaturante de 4% poli(acrilamida) . O gel foi pré-corrído por 1 hora a 80 V, antes da aplicação das amostras. Os experimentos de competição foram feitos incubando-se as reações de ligação com excesso de 100 x dos oligonucleotídeos correspondentes não marcados por 5 minutos antes da adição do oligonucleotídeo marcado. A transferência foi feita para uma membrana de Nylon usando o sistema electro-blotting (Bio-Rad). A detecção imunológica foi feita utilizando anticorpo anti-digoxigenina conjugado a fosfatase alcalina e o substrato quimiluminescente CSPD (*Roche Applied Science*).

Os oligonucleotídeos usados nos ensaios de EMSA estão listados abaixo. Estão grifadas as seqüências dos motivos e, em negrito, os nucleotídeos mutados, nos casos necessários. Em parêntesis são os genes de ORs e de V1Rs de cujos promotores as seqüências foram extraídas.

### Oligonucleotídeos extraídos de promotores de genes de ORs:

M1foward (olfr720) 5' TCTCAGACTTTTCTCCTGGGAGACATCTCAG 3'  
M1reverse (olfr720) 5' CCTGAGATGTCTCCAGGAGAAAAGTCTGAG 3'

M2foward (olfr165) 5' TAAGATGCTAAATTCCCTGGAGAAAATTGTAA 3'  
M2reverse (olfr165) 5' TTTACAATTTCTCCAGGGAATTTAGCATCTT 3'

M3foward (olfr211) 5' CCTGGCATCTCCCACTGGGGCTTATATTCTG 3'  
M3reverse (olfr211) 5' ACAGAATATAAGCCCCAGTGGGAGATGCCAG 3'

M4foward (olfr1339) 5' CTTCAGCTTCATCCTCCCTGAGGACAGGGAG 3'  
M4reverse (olfr1339) 5' GCTCCCTGTCTCAGGGAGGATGAAGCTGAA 3'

Olf1-sitefoward 5' AAGGGAGAGTCCCTAGGGAGCTTGGGAGGGG 3'  
Olf1-sitereverse 5' CCCCTCCCAAGCTCCCTAGGGACTCTCCCTT 3'

Olf1-site\*fow (mutado) 5' AAGGGAGAGT**GC**CTAGGGAGCTTGGGAGGGG 3'  
Olf1-site\*rev (mutado) 5' CCCCTCCCAAGCTCCCTAGG**CA**CTCTCCCTT 3'

M1\* (MUTADO) foward 5' TCTCAGACTTTT**AG**CCTGGGAGACATCTCAG 3'  
M1\* (MUTADO) reverse 5' CTGAGATGTCTCCAGG**CT**AAAAGTCTGAGA 3'

M2\* (MUTADO) foward 5' TTTACAATTT**ATCCATTT**CATTTAGCATCTT 3'  
M2\* (MUTADO) reverse 5' AAGGATGCTAAAT**GAAATGGAT**TAAATTGTAAA 3'

M3\* (MUTADO) foward 5' CCTGGCATCTC**ACAAGTGGT**CTTATATTCTG 3'  
M3\* (MUTADO) reverse 5' CAGAATATAAG**ACCACTTGT**GAGATGCCAGG 3'

M4\* (MUTADO) foward 5' CTTCAGCTTCATCC**GACCTGAT**TACAGGGAG 3'  
M4\* (MUTADO) reverse 5' CTCCCTGT**AA**TCAGG**TC**GGATGAAGCTGAAG 3'

### Oligonucleotídeos extraídos de promotores de genes de V1Rs:

MV3foward (vlrc5) 5' AGAATTTCTTCTCCAGATGACATGCTAATTA 3'  
MV3reverse (vlrc5) 5' AATTAGCATGTCATCTGGAGAAGGAAATTCTG 3'

MV7foward (vlrb1) 5' TGCCCAGTCCTCCCCAGATGGTTGCTATCCA 3'  
MV7reverse (vlrb1) 5' GGATAGCAACCATCTGGGGAGGACTGGGCAG 3'

MV12foward (vlrc3) 5' CACACGGAAACCCAGAGCTCACCTGTCACTA 3'  
MV12reverse (vlrc3) 5' AGTGACAGGTGAGCTCTGGGTTTCCGTGTGT 3'

MV21foward (vlri4) 5' GAGAGCAGAGAGCTGACATCTAGAATGGAG 3'  
MV21reverse (vlri4) 5' CTCCATTCTAGATGTCAGCTCTCTGCTCTC 3'

MV28forward(v1rc26) 5' GTTGTGTTCAACAAGGACTGGGATTAGTTAATT 3'  
 MV28reverse(v1rc26) 5' ATTAATAATCCCAGTCCTTGTGAACACAACA 3'

### **3.7 Hibridização *in situ* em secções coronais de epitélio olfatório de camundongo**

#### **3.7.1 Preparo das sondas de RNA usadas na hibridização *in situ*:**

Os plasmídeos com os clones foram digeridos com as enzimas adequadas para sua linearização e, a seguir, purificados com o sistema “*Rapid PCR purification System*”. Sondas de RNA marcadas com digoxigenina foram construídas utilizando reagentes da Roche e purificadas usando colunas “*Quick spin G-50*” (Roche). As sondas foram construídas nas seguintes condições livres de RNase :

- Até 13 µl de DNA template ou volume suficiente para 1µg de plasmídeo linearizado;
- 2 µl de 10X tampão de transcrição (400mM Tris-HCl, pH 8,0 (20°C); 60mM MgCl<sub>2</sub>; 100mM DTT; 20mM espermidina);
- 2 µl de mistura 10X concentrada de DIG RNA marcado (10mM ATP,10mM CTP,10mM GTP, 6,5 mM UTP, 3,5 mM DIG-11-UTP);
- 2 µl de SP6/T7 RNA polimerase (20 U/µl);
- 1 µl de inibidor de RNase.

Incubação a 37°C por 2 horas.

#### **3.7.2 Preparação das lâminas contendo cortes da cavidade nasal:**

Narizes de camundongos da linhagem C57BL/6J com 2 - 3 semanas de idade foram dissecados e mantidos congelados em “OCT compound Tissue-Tek®” (Sakura®) a – 80°C. Cortes coronais de 16 µm de espessura foram feitos em criostato (MICRON) a –20°C e coletados em lâminas de vidro “SectionLock™ Slides” (Polyscience,Inc).

### 3.7.3 Hibridização *in situ*

Em condições livres de RNase, as lâminas foram secas e fixadas por 10 minutos em solução recém-preparada de paraformaldeído 4% em 1XPBS. A seguir, as lâminas foram lavadas em 1XPBS por 3 vezes de 5 minutos cada. As lâminas foram imersas em solução contendo 7 ml de trietanolamina em 600 ml de água estéril mantendo agitação leve. Em seguida gotejou-se 1,5 ml de anidrido acético na solução onde as lâminas estavam imersas e incubou-se por 10 minutos para acetilar os tecidos. Novamente as lâminas foram lavadas com 1XPBS por 3 vezes de 5 minutos cada. Foi adicionado em cada lâmina 500µl de solução de hibridização previamente preparada com 50% de formamida; 5XSSC; 5X Denhardt's e 250 µg/ml de "bakers yeast RNA" e as lâminas foram incubadas em placas de Petri umidificadas com 5XSSC por 2 horas a 72°C. As sondas (200 ng/ml) foram misturadas com 500 µl de solução de hibridização e aquecidas por 5 minutos a 85°C para linearização das mesmas e adicionou-se a mistura nas lâminas incubando em placas de Petri a 72°C "overnight".

Após a hibridização das sondas com o RNA do tecido, os próximos passos não precisaram ser em condições livres de RNase uma vez que o RNA dupla-fita não é degradado pela RNase. Lavaram-se as lâminas em 5XSSC a 72°C por 10 minutos. A seguir, lavaram-se as lâminas em 0,2XSSC a 72°C por 30 minutos, 2 vezes, e uma vez em 0,2XSSC a temperatura ambiente por 30 minutos. Incubaram-se as lâminas a 37°C em solução de 20 µg/ml de RNase +10mM de Tris + 400mM de NaCl por 30 minutos para eliminação de sondas não hibridizadas.

A detecção imunológica foi feita com o kit "DIG Wash and Block Buffer Set" (Roche Diagnostics). As lâminas foram lavadas em 1X "washing buffer" (10X ácido

maléico com 3% Tween) por 5 minutos; em 1X “blocking reagent” (10% Blocking reagent em tampão de ácido maléico) em tampão 1X ácido maléico por 1 hora à temperatura ambiente; em 1X “blocking reagent” em tampão 1X ácido maléico contendo anticorpo anti-digoxigenina (Roche Diagnostics) na diluição de 1:5000 por 1 hora. As lâminas foram lavadas 3 vezes de 20 minutos cada com 1X “washing buffer”. As lâminas foram incubadas por 5 minutos em tampão de detecção (100mM Tris pH 9,5 + 100mM NaCl) e, a seguir, em solução de detecção preparada com 10 ml de 1X “detection buffer” + 2,4 mg de levamisole + 50 mM MgCl<sub>2</sub>+ 200 µl de NBT/BCIP (Roche Diagnostics).

A reação foi interrompida em solução 10 mM Tris pH 8,0 + 1 mM EDTA duas vezes por 5 minutos. Após lavagem com água destilada por 5 minutos as lâminas foram contra-coradas com Hoescht (10 µg/ml em PBS) por 5 minutos. Finalmente as lâminas foram lavadas com água destilada cobertas com gelvatol e cada lâmina foi coberta com uma lamínula.

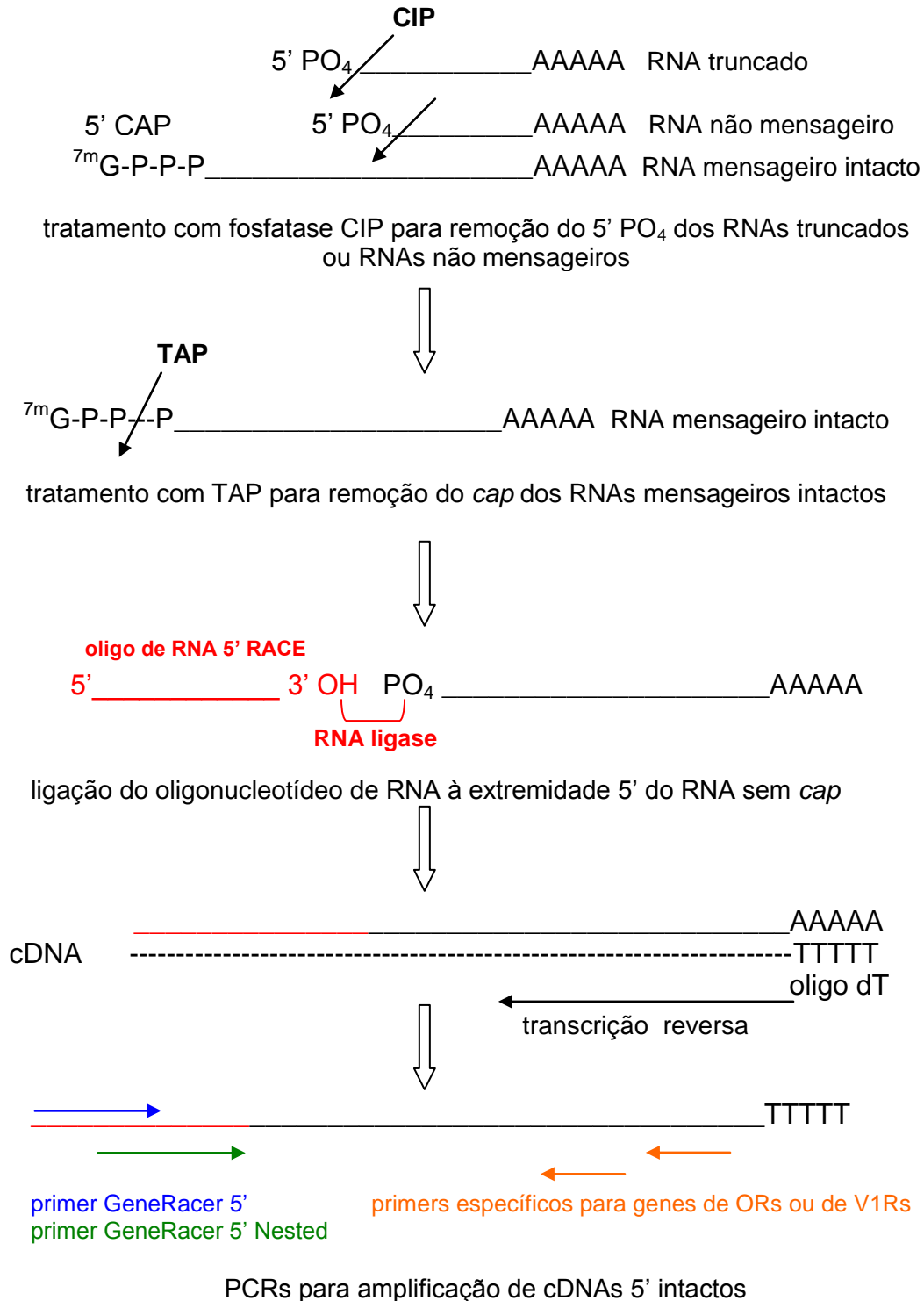
## 4) RESULTADOS

### 4.1 Estratégia utilizada para determinar regiões promotoras de genes de receptores olfatórios (ORs) e de receptores de feromônio do tipo 1 (V1Rs) de camundongo.

Com o objetivo de investigar se genes de ORs e V1Rs compartilham elementos *cis*-reguladores em suas regiões promotoras, utilizamos a estratégia de comparar regiões genômicas a 5' dos sítios de início da transcrição (TSSs) de um grande número de diferentes genes de ORs e de genes de V1Rs. Para isso, determinamos experimentalmente a exata localização dos TSSs utilizando o sistema “GeneRacer™ RLM-RACE” (*RNA ligase-mediated and oligo-capping rapid amplification of cDNA ends*). Este método permite gerar cDNAs contendo a região 5' não traduzida (5'UTR) completa dos correspondentes mRNAs. RNA total extraído de epitélio olfatório (OE) e também de órgão vomeronasal (VNO) de camundongos foram utilizados para a síntese de cDNA através do método RLM-RACE.

A figura 7 apresenta um esquema explicativo da estratégia adotada. Os RNAs mensageiros truncados (sem *cap*) e outros tipos de RNA são tratados com fosfatase de intestino de bezerro (CIP) para eliminação do grupo fosfato 5'. A seguir, os RNAs intactos, que contêm o *cap*, são tratados com a enzima TAP, uma pirofosfatase que retira o *cap* e expõem o grupo fosfato da extremidade 5', desse modo, somente os RNAs que perderam o *cap* terão grupos fosfatos livres para se ligarem à extremidade 3' OH do oligonucleotídeo de RNA fornecido no sistema. Assim, os cDNAs sintetizados com oligo dT contêm a região 5' não traduzida dos transcritos ligada à seqüência conhecida do oligonucleotídeo de RNA. Reações de PCR utilizando o oligonucleotídeo *GeneRacer* 5', alvo para o oligonucleotídeo de RNA que foi inserido na extremidade 5' dos transcritos, e oligonucleotídeos específicos que reconhecem seqüências conservadas entre diversos genes de ORs

ou de genes de V1Rs, permitem o enriquecimento da amostra com cDNAs contendo a região 5'UTR completa para genes de ORs e de V1Rs.



**Figura 7. Esquema explicativo do método de RLM-RACE.**



Um segundo PCR (*nested-PCR*), utilizando, como alvo, o produto obtido na primeira reação de amplificação juntamente com os oligonucleotídeos internos *GeneRacer 5' nested* e oligonucleotídeos específicos para genes de ORs ou para genes de V1Rs foi realizado com o objetivo de aumentar a especificidade de seleção de cDNAs provenientes de genes de ORs e genes de V1Rs. Os produtos de PCR foram, então, purificados, subclonados, inseridos em bactéria para amplificação e isolados para sequenciamento. As seqüências de cDNAs de genes de ORs e de V1Rs, contendo as extremidades 5'UTRs completas dos mRNAs, foram alinhadas contra o genoma de camundongo permitindo a exata localização dos TSSs. Regiões a 5' dos TSSs foram, então, extraídas do genoma e pesquisadas para a presença de motivos conservados.

A seguir, descreveremos os resultados obtidos nas análises das regiões promotoras de genes de ORs e de genes de V1Rs, separadamente.

## **4.2 Análise de regiões promotoras de genes de receptores olfatórios (ORs)**

### **4.2.1 Amplificação de cDNAs correspondentes a transcritos de genes de ORs contendo a região 5'UTR intacta.**

Como já descrito anteriormente, utilizamos o método de RLM-RACE para gerar seqüências de cDNA de genes de ORs que contêm a região 5'UTR completa a partir de RNA total de epitélio olfatório de camundongos. Este método permite que transcritos com a extremidade 5' completa sejam preferencialmente amplificados.

Para amplificar cDNAs de diferentes genes de ORs, utilizamos os oligonucleotídeos degenerados P8, P27 e P26R que reconhecem regiões conservadas entre diferentes membros da superfamília de genes de ORs (Malnic et al., 1999). Como indicado na figura 8A, a região transmembrânica III é alvo do

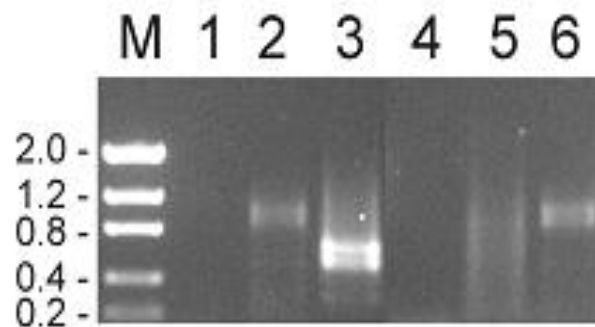
oligonucleotídeo P26R, a região transmembrânica VI é alvo do oligonucleotídeo P27 e a região transmembrânica VII é alvo do oligonucleotídeo P8.

**A**

REGIÃO CODIFICADORA DE RECEPTORES OLFATÓRIOS



**B**



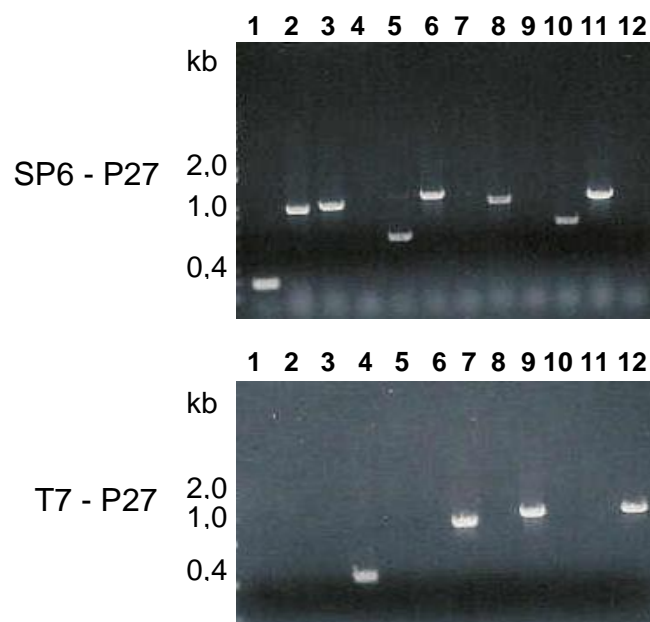
**Figura 8. Amplificação de cDNAs 5'UTR completos para genes de ORs.** A) Esquema representativo das sete regiões transmembrânicas (TM I-VII) dos receptores olfatórios mostrando os alvos dos oligonucleotídeos degenerados P26R, P27 e P8. B) Eletroforese em gel de agarose 1,5% marcado com brometo de etídeo mostrando os resultados de PCRs obtidos com diferentes combinações de oligonucleotídeos degenerados para genes de ORs. *lane 1*: controle negativo da reação do *lane 2*; *lane 2*: oligonucleotídeos GeneRacer 5' e P27 utilizando cDNA de epitélio olfatório de camundongo como alvo; *lane 3*: GeneRacer 5' Nested e P26R utilizando o produto de PCR mostrado no *lane 2* como alvo; *lane 4*: controle negativo da reação do *lane 5*; *lane 5*: oligonucleotídeos P8 e GeneRacer 5' utilizando cDNA de epitélio olfatório de camundongo como alvo; *lane 6*: oligonucleotídeos GeneRacer 5' Nested e P27 utilizando o produto de PCR mostrado no *lane 5* como alvo; M: peso molecular ( de 200 a 2000 pares de base)

A figura 8B, *lane 5*, apresenta o resultado da reação de PCR na qual utilizamos o oligonucleotídeo 5' *GeneRacer* e o oligonucleotídeo degenerado P8. Observamos que esta reação gerou uma mistura de tamanhos variados de produtos

de PCR o que era esperado já que a região 5'UTR dos transcritos varia de tamanho. Para aumentar a especificidade da seleção de transcritos de genes de ORs, realizamos um segundo PCR (*Nested-PCR*) com o oligonucleotídeo 5' *GeneRacer Nested* e o oligonucleotídeo degenerado P27, usando como alvo a primeira reação de PCR, resultado mostrado na figura 8B, *lane* 6. O mesmo procedimento foi realizado utilizando os oligonucleotídeos 5' *GeneRacer* e P27 na primeira reação de PCR (figura 8B, *lane* 2) e os oligonucleotídeos 5' *GeneRacer Nested* e P26R na segunda reação de PCR (figura 8B, *lane* 3).

O alinhamento de várias seqüências de ORs mostra que a distância entre o ATG inicial predito e a região em que os oligonucleotídeos degenerados P8 e P27 reconhecem nos genes de ORs compreende por volta de 780 pares de base (pb) e a distância entre o ATG predito e a região reconhecida pelo oligo degenerado P26R é de 390 pb. Deste modo, somente os produtos de PCR maiores que ~800 pb (do *lane* 6, figura 8B) e maiores que ~400 pb (do *lane* 3, figura 8B) foram gel purificados. Os fragmentos de cDNA purificados do gel de agarose foram clonados no plasmídeo pCRII. Colônias isoladas de bactérias contendo as seqüências de cDNAs obtidas por RLM-RACE inseridas no plasmídeo foram selecionadas por PCR de colônia usando os oligonucleotídeos T7 e SP6, que reconhecem regiões que flanqueiam o sítio múltiplo de clonagem do plasmídeo, juntamente com os primers degenerados P27 e P26R. A combinação de primers utilizada foi: T7– P27 e SP6 – P27 ou T7 – P26R e SP6 – P26R. Cada colônia foi usada como alvo para duas reações de PCR. A figura 9 mostra um exemplo de resultados obtido para 12 colônias. Neste exemplo, cada uma das 12 colônias foi usada em duas reações de PCR: uma contendo os oligonucleotídeos SP6 e P27 e outra reação contendo os oligonucleotídeos T7 e P27. Como podemos observar na figura 9, somente uma das reações amplifica o

fragmento de PCR flanqueado pelo oligonucleotídeo *forward* T7 ou SP6 e o oligonucleotídeo reverso P27. Por exemplo, no *lane* 3 dos dois géis de agarose mostrados na figura 9, vemos um produto de PCR com aproximadamente 1000 pb na reação com SP6 e P27 e nenhum produto de PCR na reação com T7 e P27 o que indica que o inserto foi clonado, no plasmídeo, na orientação “SP6 – 5’ inserto 3’ – T7”.



**Figura 9. Determinação da orientação dos insertos clonados no vetor pCR II.** Eletroforese em gel de agarose 1% corado com brometo de etídeo. As amostras 1 a 12 são produtos de PCR obtidos de colônias de bactéria transformadas com o vetor pCR II contendo como insertos seqüências de cDNA de epitélio olfatório com extremidade 5’ completa amplificadas pelo método RLM-RACE. Cada colônia isolada foi usada como alvo de duas reações de PCR, cada reação com uma combinação de oligonucleotídeos diferente: SP6-P27 e T7-P27. Os oligonucleotídeos SP6 e T7 anelam com regiões do vetor que flanqueiam o inserto e os insertos, neste exemplo, anelam com o oligonucleotídeo degenerado para receptores olfatórios P27. Como esperado, para cada colônia, somente em uma das combinações de oligonucleotídeos ocorre amplificação.

Este PCR permitiu a determinação da orientação dos insertos no plasmídeo, deste modo, pudemos selecionar o oligonucleotídeo SP6 ou T7 para sequenciarmos completamente a região 5'UTR dos cDNAs obtidos. Além disso, os resultados do PCR de colônia nos assegurou que insertos improváveis de serem cDNAs de genes de ORs fossem eliminados pois, pelo alinhamento de vários ORs, sabe-se que a distância entre o ATG e a região reconhecida pelo oligonucleotídeo P27 tem por volta de 780 pb, logo, produtos de PCR menores que ~800 pb foram descartados. Os plasmídeos das colônias selecionadas foram purificados e seqüenciados.

#### **4.2.2 Análise das seqüências de cDNAs de genes de ORs obtidas por RLM-RACE**

Sequenciamos um total de 1012 clones contendo a região 5'UTR dos cDNAs selecionados e 80% destes correspondiam a genes de ORs, o que indica que nossa estratégia, preferencialmente, amplificou seqüências de genes de ORs. Além disso, 96% dos cDNAs de genes de ORs que obtivemos contêm seqüências 5'UTR a montante do ATG inicial predito indicando que realmente transcritos com a região 5' íntegra foram amplificados. Análises utilizando a ferramenta BLAST (<http://www.ncbi.nlm.nih.gov>) mostrou que obtivemos seqüências de cDNA com 5' UTR completo para 198 diferentes genes de OR, o que corresponde a 17% do repertório total de genes ORs preditos em camundongo.

Na tabela 1 comparamos, nos cromossomos de camundongo, a distribuição dos genes de OR para os quais obtivemos seqüências de cDNA com a distribuição de genes de ORs previamente anotados (Godfrey *et al.*, 2004). Exceto para o cromossomo 3, obtivemos seqüências de genes de ORs que estão distribuídas em todos os cromossomos onde se encontram genes de ORs sendo que a maioria das

nossas seqüências correspondem a genes de ORs nos cromossomos 2 e 7, onde estão localizados o maior número de genes de ORs (Godfrey *et al.*, 2004; Young *et al.*, 2002; Zhang & Firestein, 2002). Consideramos que este resultado mostra que nossa amostra de cDNAs de genes de ORs representa satisfatoriamente o repertório de genes de ORs em camundongo.

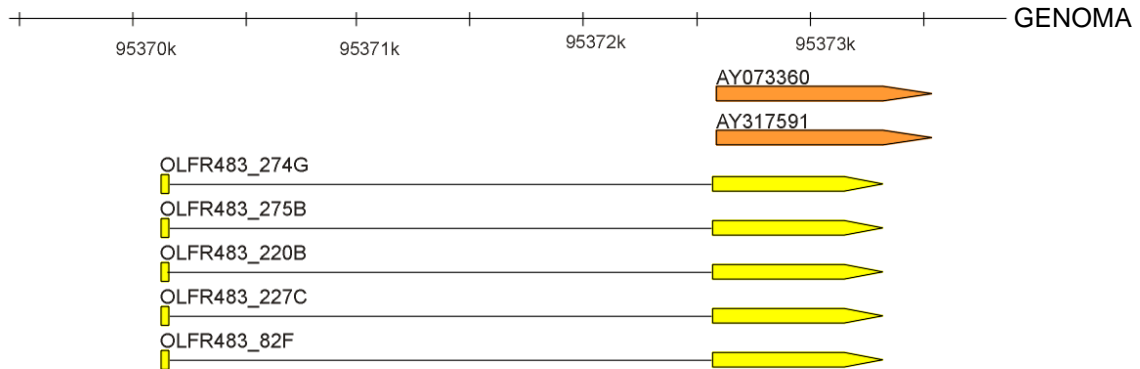
**Tabela1:** Distribuição nos cromossomos de camundongo dos genes de ORs previamente anotados e dos genes de ORs para qual obtivemos seqüências de cDNA. O número de pseudogenes estão indicados em parênteses.

<b>cromossomo</b>	<b>Número de genes anotados (Godfrey <i>et al.</i>, 2004)</b>	<b>Número de genes representados por cDNAs</b>
1	23(7)	6
2	275(81)	44(5)
3	2	-
4	19(7)	4
5	-	-
6	22(9)	5
7	198(51)	32
8	4	2
9	118(37)	20(1)
10	47(13)	9
11	35(20)	21(2)
12	-	-
13	12(2)	4
14	32(4)	6
15	6(3)	3(1)
16	29(8)	18
17	36(17)	5
18	-	-
19	52(19)	9
X	2	1
Y	-	-
<b>número total</b>	<b>1190</b>	<b>198</b>

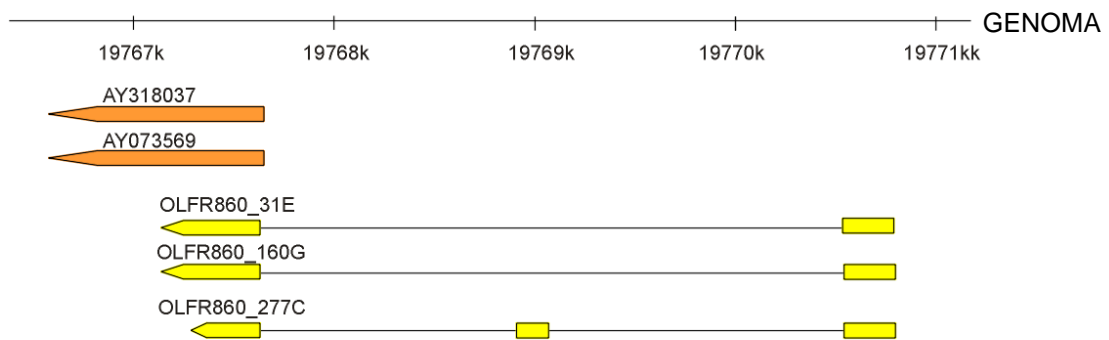
O aluno de doutorado Pedro A. F. Galante participou também deste projeto e realizou todo o trabalho de alinhamento e agrupamento das seqüências de cDNAs obtidas, além da análise da estrutura gênica 5' (caracterização de éxons e íntrons) e extração das seqüências promotoras.

As seqüências de cDNA de genes de ORs obtidas foram alinhadas contra a seqüência genômica de camundongo. Regiões codificadoras de genes de ORs, previamente anotadas (Young *et al.*, 2002; Zhang & Firestein, 2002), foram incluídas no alinhamento para auxiliar na definição das regiões 5' e codificadoras das nossas seqüências de cDNA de genes de ORs. Esse alinhamento permitiu que seqüências de cDNA correspondentes ao mesmo gene de OR fossem agrupadas em um mesmo *cluster*. Obtivemos um total de 198 *clusters* sendo que cada *cluster* contém as seqüências de cDNA de um gene de OR diferente. Cada cluster contém, ao menos, uma seqüência de cDNA e 54% dos clusters contém mais de 2 seqüências. A figura 10 representa dois exemplos de *clusters*. Na figura 10A vemos o *cluster* correspondente ao gene de OR olfr483 que contém 5 seqüências de cDNA. Em laranja vemos o alinhamento das seqüências codificadoras preditas dos ORs, que não possuem íntrons (Young *et al.*, 2002; Zhang & Firestein, 2002) e em amarelo o alinhamento das seqüências de cDNA, que obtivemos experimentalmente. O resultado mostra que este gene apresenta um único íntron e um éxon não traduzido como OR a 5' da região codificadora do OR em questão. Este tipo de estrutura gênica é encontrada na maioria dos genes de ORs analisados. A figura 10B mostra o *cluster* para o gene de OR olfr860 que é composto de 3 seqüências de cDNA e podemos observar que este gene apresenta dois éxons a 5' da região codificadora do OR, dois íntrons e sofre *splicing* alternativo.

A



B



**Figura 10. Exemplo da organização de 2 *clusters* de genes de receptores olfatórios.** A) A região do cromossomo 7 que contém o gene *olfr483* está representada pela linha contínua superior. As anotações das regiões codificadoras (que não possuem íntrons) preditas AY073360 e AY317591 (Young *et al.*, 2002; Zhang & Firestein, 2002) estão representadas em laranja. O cluster *olfr483* é formado por 5 cDNAs que alinham na mesma região e sobrepõem-se com as regiões codificadoras preditas. Os éxons estão representados em amarelo e os íntrons como linhas. A extremidade 5' do gene localiza-se à esquerda. B) O cluster do gene *olfr860* está localizado no cromossomo 9 e é formado por 3 seqüências de cDNA. As anotações das regiões codificadoras preditas AY318037 e AY035369 (Young *et al.*, 2002; Zhang & Firestein, 2002) estão representadas em laranja. O gene *olfr860* contém 2 íntrons e 2 éxons a 5' da região codificadora que podem sofrer *splicing* alternativo gerando dois tipos de cDNAs que diferem na extremidade 5'. Neste caso, a extremidade 5' do gene está localizada à direita.



A seguir, analisamos a estrutura da região 5'UTR dos 198 genes de ORs. Em média, o tamanho da região 5'UTR é de 139 pb, variando de 32 a 659 pb. Os resultados apresentados na tabela 2A mostram que dos 198 genes de ORs somente dois não apresentam íntrons na região 5'UTR. Do total de 198 genes de ORs analisados 72% possuem um éxon na região 5'UTR, 23% possuem dois éxons e 3% possuem 3 éxons. A tabela 2B mostra que os éxons presentes na região 5'UTR variam de 9 a 403 pb em tamanho. O tamanho dos íntrons varia bastante, de 91 pb até 22,5 kb, mas a grande maioria (46%) possui íntrons com 2 a 4 kb de tamanho (tabela 2C). Dos clusters com mais de duas seqüências de cDNA, 39 apresentam *splicing* alternativo.

Em nossas análises encontramos 108 genes que possuem uORFs (*upstream Open Reading Frames*) nos éxons a 5' da região codificadora dos ORs. Acredita-se que, em geral, a tradução de um mRNA maduro inicia-se no primeiro códon de iniciação AUG (Kozak, 1999). Estas uORFs possuem ao menos 10 códons sendo capazes de codificar pequenos peptídeos. Sugeriu-se que uORFs na região 5'UTR de transcritos possam estar envolvidas num mecanismo de regulação negativa da eficiência de tradução das proteínas (Pesole *et al.*, 2001). O papel destas uORFs localizadas na região 5'UTR dos genes de ORs precisa ainda ser analisado.

Também constatamos que um total de 35 genes de OR analisados tem um AUG em fase localizado à esquerda do AUG predito indicando que estes genes podem codificar receptores olfatórios com uma região amino-terminal maior do que a predita somente pela análise das seqüências genômicas.

**Tabela 2. Estrutura da região 5' UTR dos genes de ORs.** São mostrados o número de genes de ORs contendo zero a quatro exons na região 5' UTR e o número de éxons e íntrons com diferentes tamanhos na região 5' UTR.

**A**

número de genes de ORs	número de éxons
2	0
143	1
46	2
6	3
1	4

**B**

número de exons	tamanho dos éxons (bp)
58	< 50
67	50 – 100
54	100 – 150
35	150 – 200
31	> 200

**C**

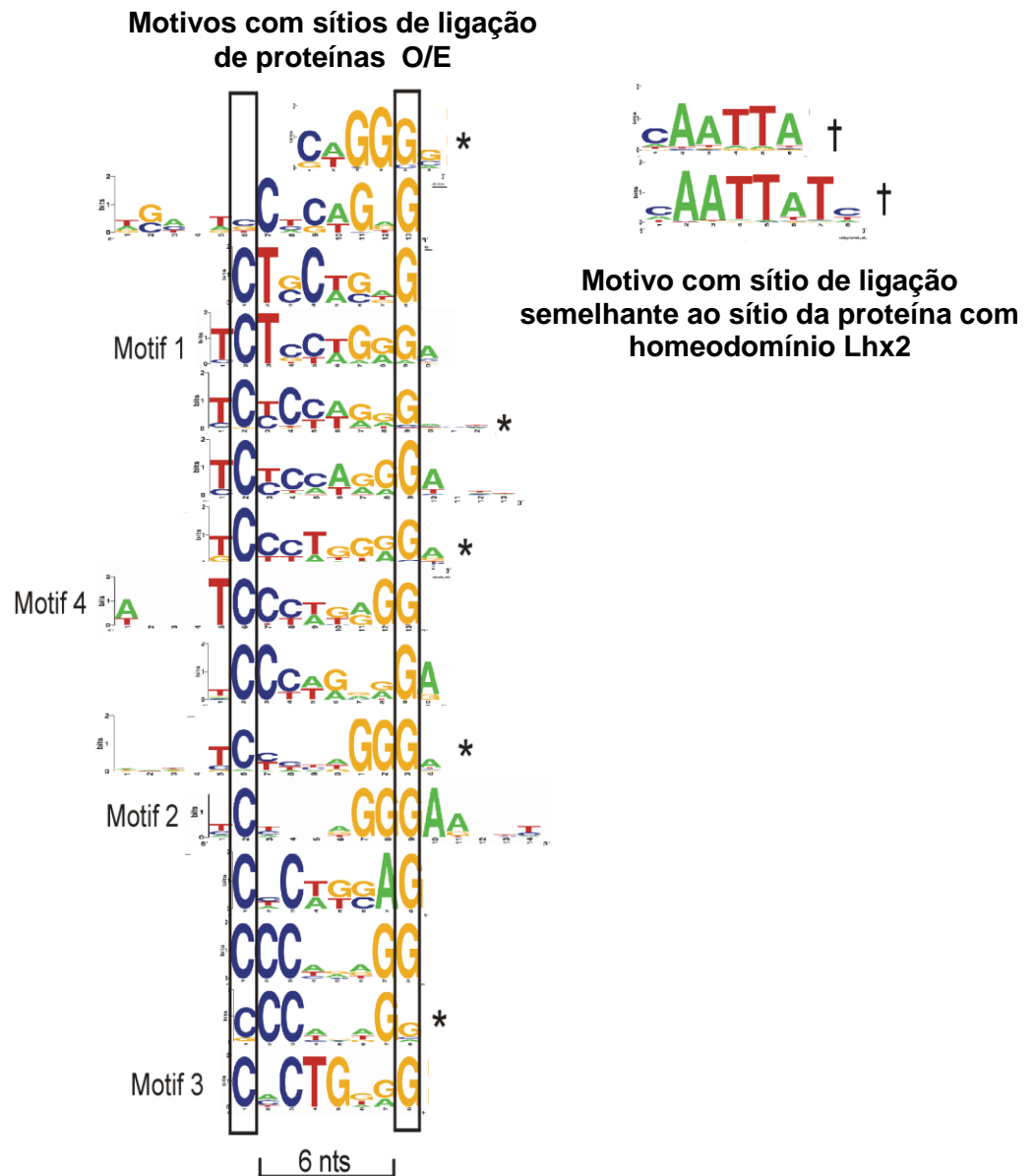
número de íntrons	tamanho dos íntrons (kb)
3	< 0,2
13	0,2 – 0,5
16	0,5 – 1
53	1 – 2
91	2 – 4
50	4 – 8
17	> 8

#### 4.2.3. Análise das regiões promotoras dos genes de ORs.

O alinhamento das seqüências de cDNA contendo a região 5'UTR completa contra o genoma de camundongo permitiu que determinássemos a localização exata dos TSSs para 198 genes de ORs. Extraímos seqüências genômicas de 600 pb a montante dos TSSs. Primeiramente, regiões repetitivas ou de baixa complexidade, correspondentes a 7% do total de 198 seqüências promotoras analisadas, foram mascaradas para a próxima análise.

Buscamos motivos conservados nas regiões promotoras utilizando três programas: Gibbs recursive sampler (Thompson *et al.*, 2003); Consensus (Hertz & Stormo, 1999) e Weeder (Pavesi *et al.*, 2004). Cada programa usa um método diferente para buscar motivos conservados em seqüências não alinhadas. A figura 11 mostra os resultados obtidos de várias análises com os três programas de bioinformática Gibbs, Consensus e Weeder. Os motivos conservados em várias das 198 regiões promotoras estudadas estão representados por Weblogs: representações gráficas que mostram o grau de conservação de nucleotídeos em cada posição na seqüência.

A figura 11 mostra que obtivemos, utilizando os programas Gibbs e Consensus, motivos semelhantes ao sítio de ligação para os fatores de transcrição da família O/E expressos no sistema olfatório (Wang & Reed, 1993). As proteínas O/E formam uma família com quatro membros chamados O/E1 (também chamada Olf1), O/E2, O/E3 e O/E4. Sítios de ligação para proteínas O/E foram identificados na região promotora de genes expressos especificamente em neurônios olfatórios (Kudrycki *et al.*, 1993; Wang *et al.*, 1993).



**Figura 11. Motivos encontrados nas regiões promotoras dos genes ORs.** Cada representação gráfica, gerada com o programa Weblogo, corresponde aos resultados obtidos com os programas Gibbs recursive sampler, Consensus(\*) e Weeder(†). Os motivos que lembram sítios de ligação de O/E foram alinhados e mostram o consenso CN<sub>6</sub>G. Os motivos 1 a 4, nomeados como motif1, motif2, motif3 e motif4, foram escolhidos como representantes de diferentes grupos de sítios O/E-like e foram usados em outras análises descritas nas figuras 13,14 e 15.

Todos os membros da família O/E ligam-se a sítios-Olf1 (seqüências mostradas na figura 12 para os genes *OMP*, *Golf*, *AcIII* e *OcNC*) formando homodímeros ou heterodímeros e são expressos em neurônios olfatórios maduros e nas células progenitoras desses neurônios (Wang *et al.*,1997; Wang *et al.*, 2002). Pesquisamos por motivos entre 6 e 14 nucleotídeos de comprimento e o alinhamento mostrado na figura 11 evidencia que mesmo diminuindo ou aumentando o comprimento do motivo a ser pesquisado pelos *softwares* Gibbs e Consensus, sempre encontramos o mesmo *core* de sítio de ligação semelhante ao sítio de ligação de O/E.

Analisando atentamente estes resultados observamos que, embora estes motivos lembrem sítios de ligação O/E e apresentem o consenso geral CN<sub>6</sub>G, eles também apresentam outros nucleotídeos conservados que permitem dividi-los em quatro grupos, dos quais selecionamos 4 representantes chamados de motivo 1 (M1), motivo 2 (M2), motivo 3 (M3) e motivo 4 (M4) que foram utilizados em outras análises, como descreveremos.

A figura 12 mostra o alinhamento de 10 seqüências escolhidas ao acaso para cada grupo de motivos semelhantes a sítios de ligação de O/E (M1-M4) e os sítios O/E já conhecidos para os genes de *OMP* (*Olfactory Marker Protein*), *Golf* (proteína G olfatória), adenilil ciclase III (*AcIII*) e do canal dependente de nucleotídeos cíclicos (*OcNC*). Cada grupo M1 a M4 apresenta diferentes nucleotídeos conservados nas posições marcadas com asteriscos (figura 12).

```

OPM  GTCCTCAGGGAG
OPM  CTTCCAGGGGAG
ACIII TTTCCAGGGGAG
OcNC  GTCCTCAGGGAG
GOLF  TTTCCCTCAGGAG
      ***** **

M1_720  TTCTCCTGGGAG  M2_165  TTCTCCAGGGAA
M1_1000 TTCTCCTAGGAG  M2_457  TCCCACAGGGAA
M1_56   CTCTCTTGGGAG  M2_1501 TCCTAGAGGGAA
M1_1273 ACCTCCTGGGAG  M2_1270 TACTAGAGGGAA
M1_339  ATCTCCTGGGAT  M2_181  TTCTGCAGGGAA
M1_1377 ATCTCCAGGGAC  M2_1352 TTCCATGGGGAA
M1_1045 ACCTCCAGGAAT  M2_27   ATCCATAGGGAA
M1_1356 CTCTTCAGAGAA  M2_741  TTTACTGGGGAA
M1_123  GTCTTCAGAGAA  M2_181  AACACTGGGGAA
M1_191  CCCTCCAGGGAA  M2_56   CTCTCTTGGGAG
      **          *
M3_211  CCCACTGGGGCT  M4_1339 ATCCTCCCTGAG
M3_1377 TCCACTGTGGTT  M4_389  ATCCCTGAGGAA
M3_1352 TTCCCTGTGGGA  M4_392  ATCCCTGAGGAG
M3_1308 ACCACTGAAGCA  M4_214  TTCCAGAGGAT
M3_171  TGCCTGTAGCA  M4_1    ATCCCAGAGGAG
M3_32   CACACTGTAGAC  M4_1000 TTCTTGGGGAC
M3_1434 AACACTGGGGAA  M4_24   TTCTATGGGGT
M3_90   CCCACTGGAGTA  M4_1301 GTCCATGGGGT
M3_1415 CTCCTGGAGGA  M4_1261 GTCCTTGGGTA
M3_413  ATCCCTGGAGTC  M4_702  ATCCTTTAGGTA
      *  ***  *          ***      *

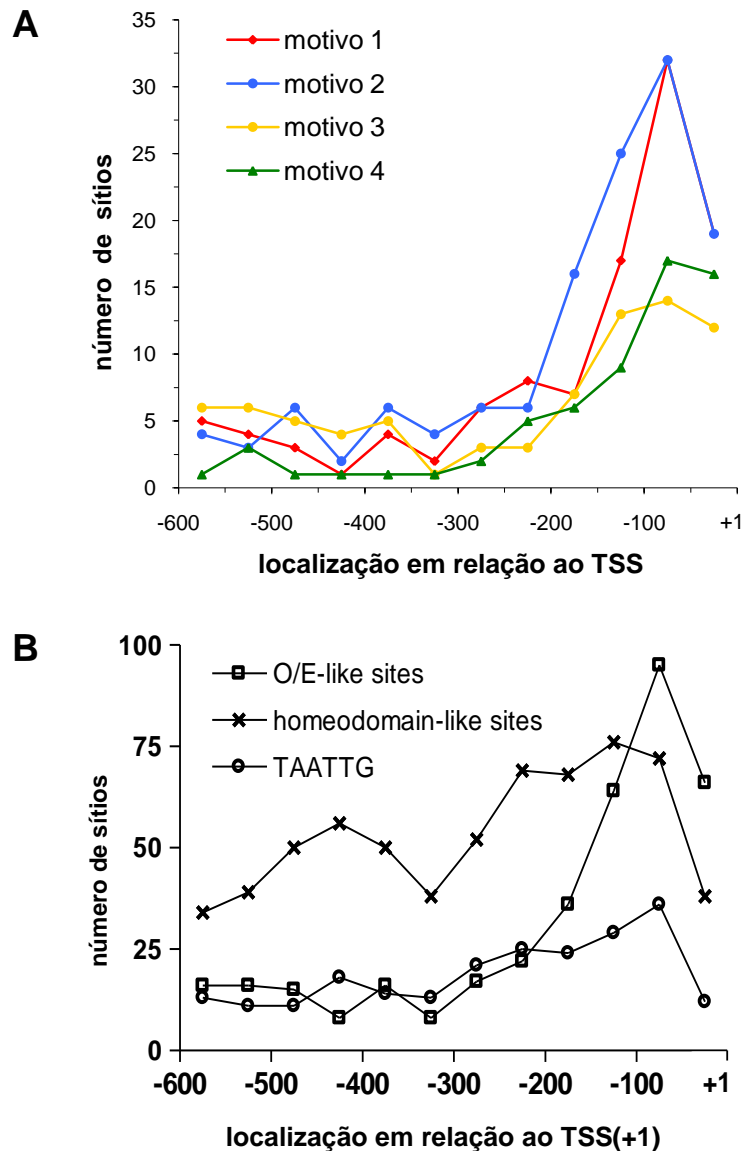
```

**Figura 12. Alinhamento de seqüências de nucleotídeos dos motivos 1 a 4 comparadas com os sítios de ligação de O/E conhecidos em outros genes.** Dez seqüências foram selecionadas de forma aleatória para cada um dos motivos (M1-M4) e alinhadas. Cinco seqüências, anteriormente identificadas como sítios de ligação de O/E nos promotores dos genes OPM, ACIII, OcNC e Golf (Wang et al. 1993), também foram alinhadas. Os nucleotídeos conservados em cada alinhamento estão indicados pelos asteriscos. Os nomes dos genes de ORs de onde vieram estes motivos estão indicados (por exemplo: M1\_720 corresponde ao motivo 1 encontrado no promotor do gene olfr720).

Identificamos, com uso do programa Weeder, sítios de ligação para proteínas com homeodomínios, mostrados na figura 11. Estes sítios foram anteriormente identificados na região promotora de alguns genes de ORs em proximidade com sítios de ligação de O/E (Rothman *et al.*, 2005; Vassalli *et al.*, 2002) e foi demonstrado que a proteína com homeodomínio Lhx2 interage com este sítio em ensaios de monohíbrido e ensaios de mobilidade eletroforética (Hirota & Mombaerts, 2004).

O gráfico da figura 13A mostra a distribuição dos 4 motivos semelhantes a sítios de ligação de O/E que identificamos dentro do intervalo de -600 a +1 das regiões promotoras de ORs obtidas. Esses motivos estão concentrados em regiões promotoras (-200 até +1) próximas aos TSSs.

O gráfico da figura 13B compara a distribuição de todos os motivos semelhantes a sítios de O/E, chamados de *O/E-like sites*, com a distribuição dos sítios de ligação de homeodomínios, chamados de *homeodomain-like sites*. Os sítios de ligação de homeodomínios apresentam uma distribuição mais espalhada nas regiões de 600 pb dos 198 promotores analisados. A seqüência 5' TAATTG 3', representada no gráfico da figura 8B, foi encontrada próxima a um sítio de O/E no promotor do gene de OR M71 e identificada como alvo da proteína Lhx2 (Rothman *et al.*, 2005; Hirota & Mombaerts, 2004).



**Figura 13. Gráfico da distribuição dos motivos na região de -600 até o TSS (+1) dos genes de receptores olfatórios analisados.** Cada ponto no gráfico representa o número de motivos presentes, naquele intervalo, para o total de 198 promotores analisados. A) Distribuição dos motivos de 1 a 4 semelhantes a sítios de ligação de O/E. Estes motivos aparecem mais representados na região entre -200 e +1. B) Localização dos motivos no total de 198 regiões promotoras analisadas. O número total de sítios de ligação de O/E (*O/E-like sites*) corresponde à soma dos sítios dos motivos M1-M4 e os sítios de ligação de homeodomínios (*homeodomain-like sites*) são baseados no consenso (C/A)AATTA. A distribuição do sítio de homeodomínio TAATTG, anteriormente identificado na região promotora de alguns genes de ORs (Vassalli *et al.* 20002) e envolvido na regulação da expressão do OR M71 (Rothman *et al.* 2005) é mostrada.



Nossos resultados mostram que a grande maioria dos genes de OR analisados compartilham elementos comuns nas suas regiões promotoras. Utilizando três programas diferentes identificamos somente dois sítios de ligação, sítios de ligação semelhantes aos sítios de proteínas O/E e sítios de ligação de proteína com homeodomínio Lhx2, presentes em 87% e 95% dos 198 promotores analisados, respectivamente, como apresentado na tabela 3.

**Tabela 3. Distribuição dos motivos nos promotores de genes de ORs.** \*Número de genes ORs do total de 198 genes analisados que contêm os motivos *O/E-like* e *homeodomain-like* nas regiões promotoras. †Número total de sítios encontrados no total de 198 promotores analisados.

<b>Motivo</b>	<b>Número de genes ORs* (% do total de 198 promotores)</b>	<b>Número total de sítios no total de 198 promotores†</b>
O/E-like	173 (87%)	382
Homeodomain-like	188 (95%)	1029

Importante lembrar que não podemos excluir a possibilidade de que existam outros motivos conservados, menos freqüentes do que os motivos que encontramos em nossa análise, e que não foram identificados pelos programas utilizados.

#### **4.2.4. Ensaio de EMSA para avaliar o significado biológico dos motivos encontrados nos promotores de genes de ORs.**

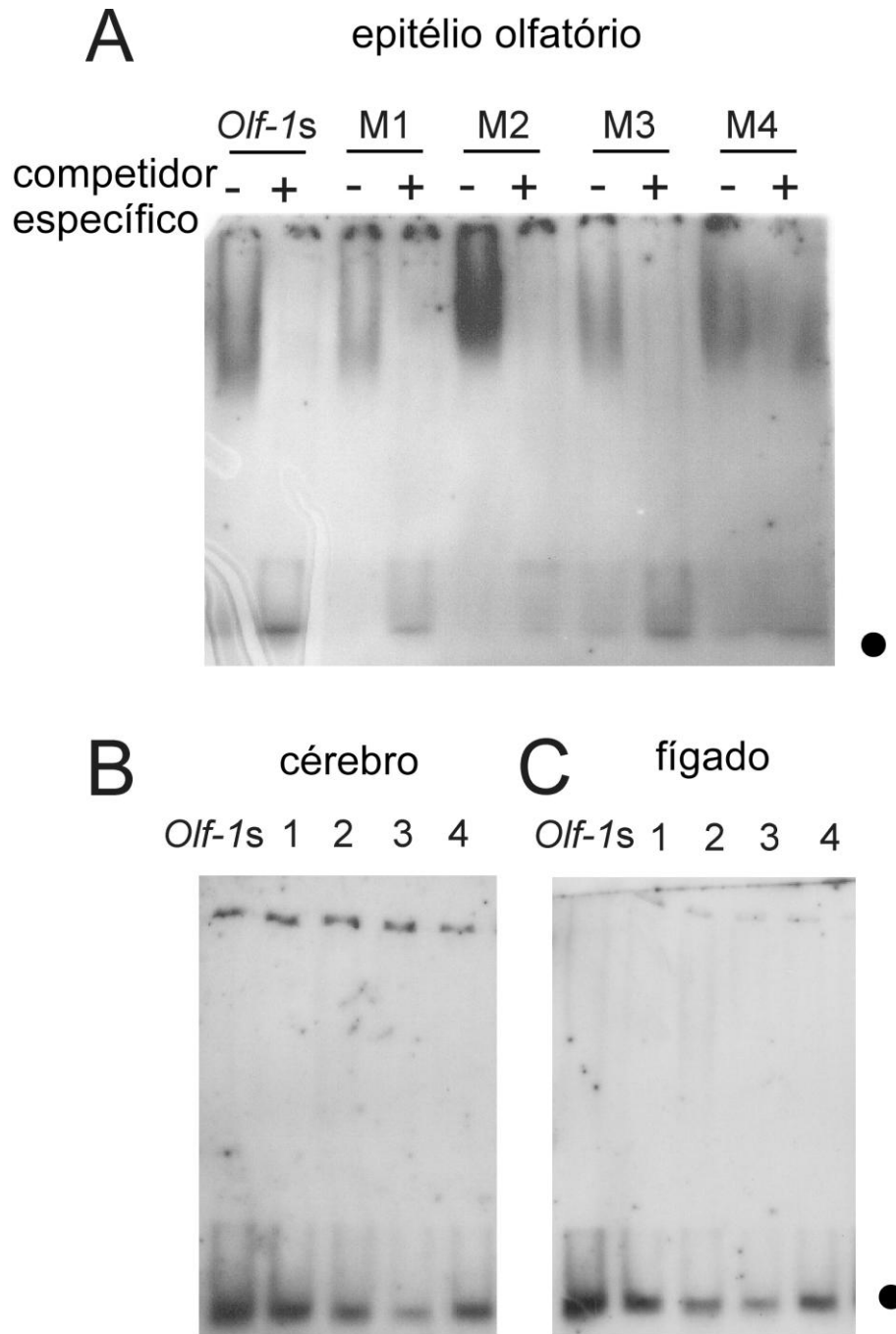
Para avaliar o significado biológico dos diferentes motivos semelhantes a sítios de ligação de O/E, realizamos ensaios de mobilidade eletroforética em gel de poliacrilamida (EMSA).

Desenhamos oligonucleotídeos contendo seqüências representativas para cada um dos motivos M1-M4. As extremidades 3' dos oligonucleotídeos dupla-fita foram marcadas com ddUTP ligado à molécula de digoxigenina. Esses oligonucleotídeos marcados foram incubados com extratos protéicos nucleares de epitélio olfatório, cérebro e fígado e, a seguir, realizamos eletroforese em gel não denaturante de poliacrilamida 4%. Os complexos formados entre DNA (oligo) e proteínas migram mais lentamente que os oligonucleotídeos livres e, dessa forma, podemos detectar se proteínas do extrato nuclear interagiram com determinada seqüência de DNA. Após imobilizarmos as proteínas e o DNA presentes no gel em membranas *Hybond-N+*, realizamos a detecção imunológica utilizando anticorpo anti-digoxigenina conjugado à fosfatase alcalina e o substrato quimiluminescente CSPD.

Como mostrado na figura 14A, os motivos de DNA M1-M4 formam complexos com proteínas nucleares de epitélio olfatório. Como controle do extrato nuclear de epitélio olfatório, usamos o sítio de ligação para Olf-1 (O/E1), que já havia sido caracterizado pelo mesmo ensaio (Kudrycki *et al.*, 1993; Wang *et al.*, 1993). A formação dos complexos entre proteínas nucleares de epitélio olfatório e os oligos marcados não é mais visualizada quando um excesso de oligos não marcados (competidores específicos) é adicionado à reação. Nesses casos, observamos que o oligonucleotídeo marcado presente no complexo é deslocado pelo excesso do mesmo oligonucleotídeo só que não marcado e vemos o aumento da intensidade da banda correspondendo ao oligonucleotídeo marcado na extremidade inferior do gel (figura 14A) indicando a especificidade de formação destes complexos. Não observamos a formação de complexos entre DNA e proteínas quando os motivos são incubados com extratos nucleares de fígado e cérebro, como mostrado nas

figuras 14B e 14C, indicando que os motivos analisados são alvos de interações com proteínas presentes no extrato nuclear de epitélio olfatório.

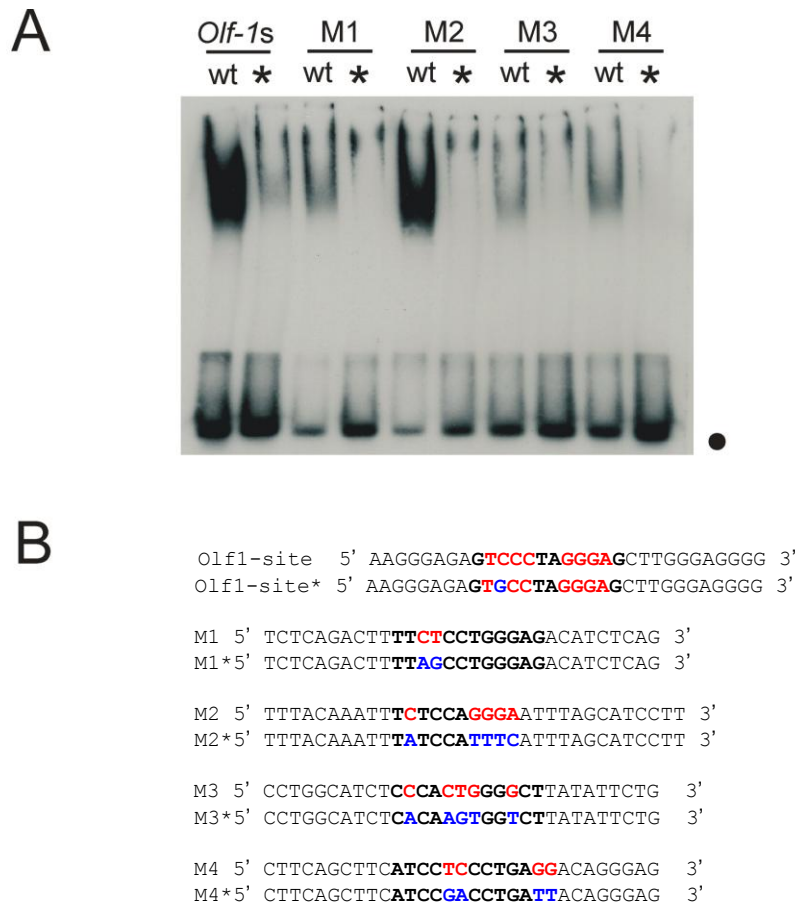
Como estes motivos são sequências de DNA altamente relacionadas aos sítios de ligação de O/E, nos perguntamos se os complexos formados na figura 14A seriam devido a interações das proteínas O/E com os motivos M1-M4. Observamos (dados não apresentados) que o oligonucleotídeo não marcado para Olf1, quando usado como competidor em reações de ligação com os motivos M1-M4, é capaz de deslocar os complexos formados entre os motivos marcados e as proteínas do extrato nuclear de epitélio olfatório da mesma forma que os competidores específicos mostrados na figura 14A. Além disso, o oligonucleotídeo marcado do motivo M2 também é capaz de deslocar o complexo formado entre o oligonucleotídeo marcado contendo o sítio de ligação de Olf1 e as proteínas do extrato nuclear (dado não mostrado). Esses dados sugerem que as mesmas proteínas, provavelmente as O/E, devem estar interagindo com a sequência de DNA olf1 e os motivos 1-4.



**Figura 14. Ensaio de mobilidade eletroforética em gel não denaturante de poliacrilamida 4%.** Oligonucleotídeos dupla-fita marcados com digoxigenina correspondendo aos motivos M1-M4 foram incubados com extratos nucleares de epitélio olfatório (A), cérebro (B) e fígado (C), como indicado. Os oligonucleotídeos marcados livres estão indicados pelo símbolo •.

#### 4.2.5 Especificidade das interações dos motivos semelhantes a sítios O/E

Para avaliar se as interações entre DNA e proteínas observadas na figura 14 eram realmente específicas e dependentes de seqüências conservadas nos motivos M1-M4, os nucleotídeos conservados em cada motivo foram mutados e usados em ensaios de ligação. Como mostrado na figura 15A, os oligonucleotídeos mutados não são capazes de formar complexos entre DNA e proteínas com extrato nuclear de epitélio olfatório indicando que as interações entre os oligonucleotídeos e as proteínas nucleares de epitélio olfatório dependem dos nucleotídeos conservados nos motivos M1-M4.



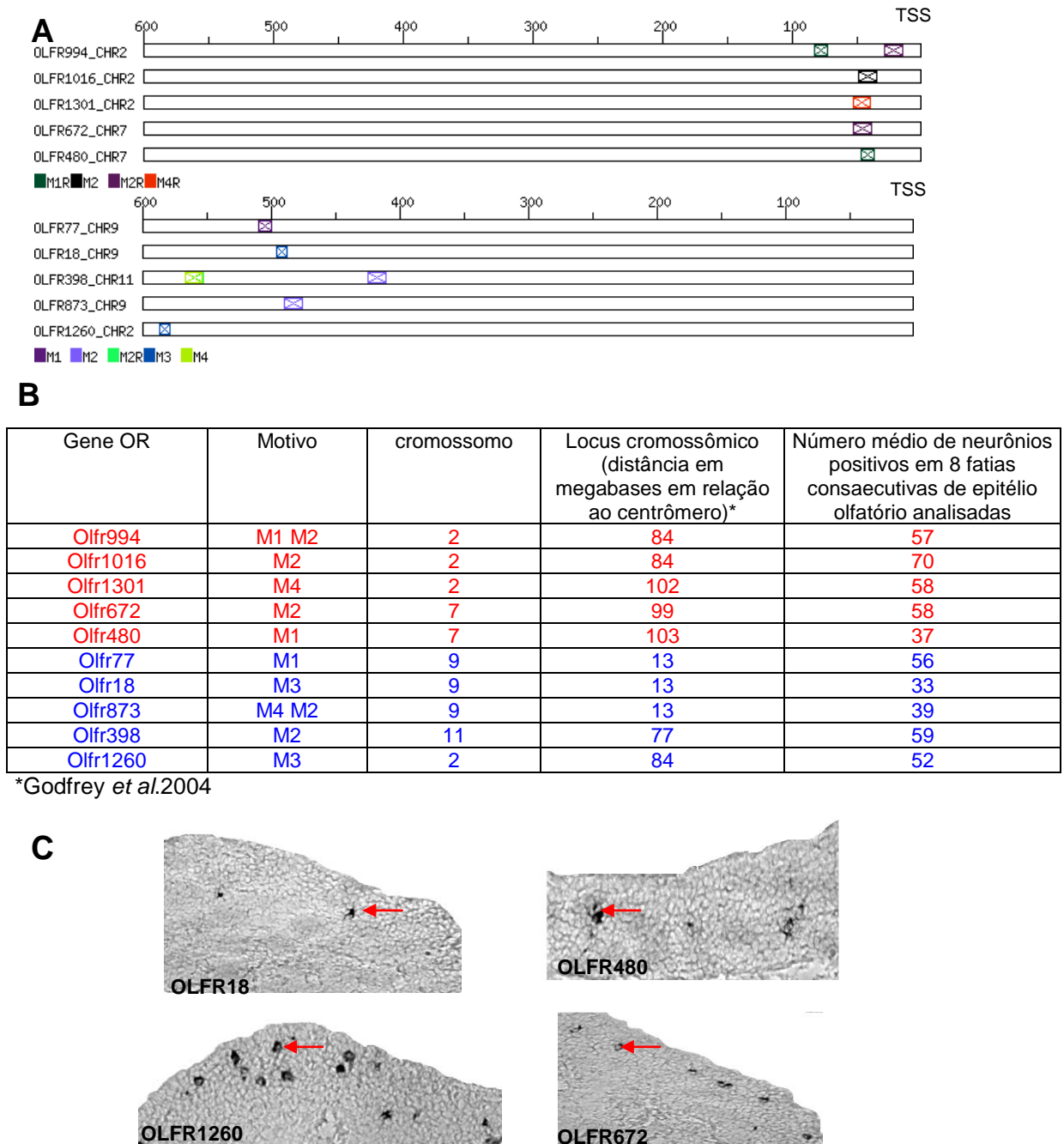
**Figura 15. Especificidade dos complexos formados entre DNA e proteínas nucleares. Ensaios de mobilidade eletroforética em gel de poliacrilamida 4%. A) Proteínas nucleares de epitélio olfatório formam complexos com os oligonucleotídeos contendo os motivos sem mutações (wt), mas não formam complexos com os oligonucleotídeos correspondentes mutados (\*). B) As seqüências dos oligonucleotídeos e seus correspondentes sítios mutados são mostradas. A região dos oligonucleotídeos em negrito (preto e vermelho) corresponde aos motivos. Os nucleotídeos em vermelho são os conservados em cada motivo, e em azul os nucleotídeos mutados.**

Nossos resultados mostram que diferentes promotores de genes de ORs compartilham sítios de ligação de proteínas O/E. Esses sítios possuem diferenças na seqüência que podem permitir a interação de diferentes combinações ou quantidades de proteínas O/E.

#### **4.2.6 Análise da escolha de expressão de um gene OR em relação aos motivos O/E presentes nos promotores**

Anteriormente foi demonstrado que diferentes proteínas O/E possuem propriedades de ligação ao DNA semelhantes (Wang *et al.*, 1997). Na figura 14A mostramos que os motivos M1-M4 semelhantes a sítios de ligação de proteínas O/E interagem diferentemente com proteínas presentes no extrato nuclear de epitélio olfatório. Esta diferença observada poderia ser explicada pela interação de cada motivo com uma proteína O/E diferente ou com a mesma proteína O/E mas com afinidades diferentes. Sabe-se que diferentes receptores olfatórios apresentam diferentes níveis de expressão (Young *et al.*, 2003), ou seja, o número total de neurônios expressando um determinado receptor olfatório, assim como o nível de mRNA produzido, é diferente. Com o objetivo de avaliar se o tipo de motivo semelhante a sítios O/E (M1, M2 M3 ou M4) presente no promotor de um gene de OR, assim como a sua proximidade em relação ao TSS, afeta a probabilidade de escolha de expressão deste gene pelo neurônio olfatório, analisamos a expressão de determinados ORs utilizando a técnica de hibridização *in situ* com sondas de RNA que hibridizam com transcritos de genes de OR. Com exceção do gene de OR Olfr994 que apresenta os motivos M1 e M2 no promotor e do gene de OR Olfr393 que apresenta os motivos M3 e M4 no seu promotor, os outros oito genes de ORs analisados possuem somente um tipo de motivo O/E no seu promotor. Além disso,

cinco genes de ORs escolhidos possuem sítios de ligação de O/E ( M1, M2 ou M4) próximos aos respectivos TSSs e cinco genes de ORs possuem sítios O/E (M1 ou M2 ou M3 ou M4) mais afastados aos TSSs nas regiões promotoras como representado na figura 16A. Fatias consecutivas de 16 $\mu$ m de espessura de narizes de camundongo com três semanas de idade foram depositadas em lâminas seqüenciais e cada lâmina foi incubada com uma sonda específica de RNA antisense contra transcritos de um dos genes de ORs selecionados. Realizamos contagens dos neurônios que hibridizaram com as sondas de genes ORs nas fatias de epitélio olfatório. Os resultados obtidos, apresentados na figura 16B, mostram que, em cada fatia de epitélio olfatório analisada, o número de neurônios positivamente marcados variou de 33 a 70. Observamos que a presença de um único tipo de motivo O/E (M1, M2, M3 ou M4) no promotor não parece influenciar, para mais ou para menos, a escolha de um gene OR ser transcrito. Na verdade, na maioria dos casos analisados (7 em 10), o número neurônios positivos para cada gene de OR mantém uma média de 59. Entre os cinco genes de ORs com sítios O/E próximos aos TSSs, somente o gene Olfr480 localizado no cromossomo 7 apresenta um número menor que quarenta (37) de neurônios positivamente marcados enquanto que os outros quatro genes de ORs analisados apresentam mais de 50 neurônios positivamente marcados (dados em vermelho na tabela da figura 16B). Entre os cinco genes de ORs com sítios O/E distantes dos TSSs (dados em azul na tabela da figura 16B) dois, olfr18 e olfr873, apresentam 33 e 39 neurônios marcados, respectivamente, por fatia de epitélio analisada e os outros três genes de OR apresentam mais de 50 neurônios positivos.



**Figura 16. Hibridizações *in situ* em epitélio olfatório de camundongo utilizando sondas de RNA antissense que hibridizam com RNAs mensageiros correspondentes a diferentes genes de receptores olfatórios (ORs).** A) esquema das regiões promotoras, de 600pb, a 5' dos 10 genes de ORs escolhidos. São mostrados o nome dos genes de ORs (nomenclatura *Olf*) e o cromossomo (CHR) em que residem. Os quadrados coloridos inseridos dentro dos retângulos mostram a localização dos putativos sítios de ligação para proteínas O/E (representados pelos motivos M1, M2, M3 e M4) dentro dessas regiões promotoras. B) tabela indicando o nome dos genes ORs analisados, os motivos que aparecem nos promotores, o cromossomo em que residem, a região genômica em que se encontram em relação à distância do centrômero e o número de neurônios positivos por fatia de epitélio olfatório analisado. Os dados em vermelho são referentes a genes de ORs que apresentam no promotor motivos O/E-like próximos aos TSSs e os dados em azul são referentes a motivos O/E-like distantes aos TSSs. C) Imagens de secções coronais de epitélio olfatório de camundongo obtidas em microscópio convencional com objetiva de 20X mostram marcação positiva (ponta da seta vermelha) em neurônios do epitélio olfatório indicando expressão de transcritos correspondentes aos diferentes genes de ORs analisados.



Resolvemos, então, verificar em que cromossomo e dentro de que locus (distância em megabases do gene de OR em relação ao centrômero) residem os genes de OR escolhidos (tabela da figura 16B). Do total de 10 genes de ORs analisados por hibridização *in situ*, quatro residem no cromossomo 2, três no cromossomo 9, dois no cromossomo 7 e um no cromossomo 11. Os genes de ORs residentes no cromossomo 2 apresentam mais de 50 (entre 52 e 70) neurônios positivamente marcados por fatia de epitélio olfatório analisada. Somente um dos genes de OR, olfr1301, está no locus 102 enquanto que os outros três encontram-se no locus 84 (significa que estão no mesmo cluster localizado 84 megabases do centrômero do cromossomo 2). Dos três genes de ORs localizados no locus 13 do cromossomo 9, dois apresentam menos de 40 neurônios marcados positivamente na hibridização *in situ* enquanto que o outro gene de OR, olfr77, possui 56 neurônios marcados. Em relação aos dois genes de ORs que residem no cromossomo 7 o que se localiza no locus 103 apresenta menos de 40 neurônios marcados e o que está no locus 99 apresenta 58 neurônios positivos. E finalmente o único gene de OR analisado do cromossomo 11 está no locus 77 e apresentou 59 neurônios positivamente marcados na hibridização *in situ*. Portanto, estes resultados indicam que os genes de ORs localizados em uma destas regiões do genoma não têm a probabilidade de expressão aumentada.

Em conclusão, estes resultados sugerem que nem o tipo nem a distância dos sítios O/E na região promotora de 600 pares de base a montante dos TSSs parece afetar para mais ou para menos a probabilidade de escolha de um gene de OR. No entanto, esta análise foi de uma pequena amostra (10 genes de ORs somente) e para confirmar a tendência observada novas análises precisariam ser feitas.

### **4.3 Análise de regiões promotoras de genes de receptores de feromônios do tipo 1 (V1Rs).**

Como já foi apresentado, a expressão dos genes que codificam para receptores V1Rs, semelhante à expressão dos genes de ORs, é monoalélica e mutuamente exclusiva (Rodriguez *et al.*, 1999; Ropollo *et al.*, 2007).

As semelhanças no padrão de expressão dos genes de V1Rs e dos genes de ORs levanta a possibilidade de existência de um mecanismo de regulação gênica comum aos dois neurônios sensoriais envolvendo elementos *cis*-atuantes conservados.

A estratégia que adotamos para determinar e analisar regiões promotoras de diferentes genes de ORs mostrou-se útil na identificação de elementos conservados. Empregamos a mesma abordagem para analisar também os promotores de genes de V1Rs.

#### **4.3.1 Desenho de primers degenerados e família-específicos para genes de V1Rs.**

Para obter regiões promotoras de diferentes genes de V1Rs, utilizamos o método de RLM-RACE, bem descrito no item 4.1. Como já mencionado, esse método permite gerar cDNAs correspondentes a mRNAs com a extremidade 5'UTR intacta.

Com o objetivo de amplificar produtos correspondentes a diversos genes de V1Rs, desenhamos oligonucleotídeos que reconhecem regiões conservadas entre vários genes de V1Rs. Identificamos, através do alinhamento de genes de V1Rs preditos (Rodriguez *et al.*, 2002), uma sequência conservada localizada na região transmembrânica 3 (TMIII) dos receptores V1Rs e escolhemos essa região para

desenhar oligonucleotídeos degenerados, nomeados VNR1 e VNR2 como mostrado na figura 17A. Estes oligonucleotídeos foram usados para amplificar diferentes membros da superfamília de genes V1Rs.

O repertório de genes de V1Rs divide-se em 12 famílias nomeadas de A-L (figura 5) e, ao contrário dos genes de ORs que possuem regiões altamente conservadas entre todos os membros, os genes de V1Rs são mais diversos (Rodriguez *et al.*, 2002) o que torna um desafio maior amplificar, a partir de oligonucleotídeos degenerados, um número representativo dos genes de V1Rs. Para a busca de elementos reguladores comuns em regiões promotoras de diferentes genes de V1Rs é importante pesquisar uma amostra representativa do repertório de genes expressos, no entanto, o uso dos oligonucleotídeos degenerados VNR1 e VNR2 não foi suficiente para enriquecer nossa amostra com cDNAs provenientes de diferentes famílias de genes de V1Rs. Portanto, decidimos adotar a estratégia de desenhar oligonucleotídeos específicos para as diferentes famílias de genes de V1Rs que não estavam representadas em nossa primeira amostra de cDNA. Seqüências correspondentes a membros que formam diferentes famílias de genes de V1Rs foram separadamente alinhadas. Regiões conservadas entre os membros da mesma família de genes de V1Rs foram selecionadas e usadas no desenho de oligonucleotídeos família-específicos. Estes foram usados na seleção de cDNAs 5' UTR intactos correspondentes a genes de diferentes famílias de V1Rs.

#### 4.3.2 Amplificação de cDNAs contendo a região 5'UTR completa para genes de V1Rs.

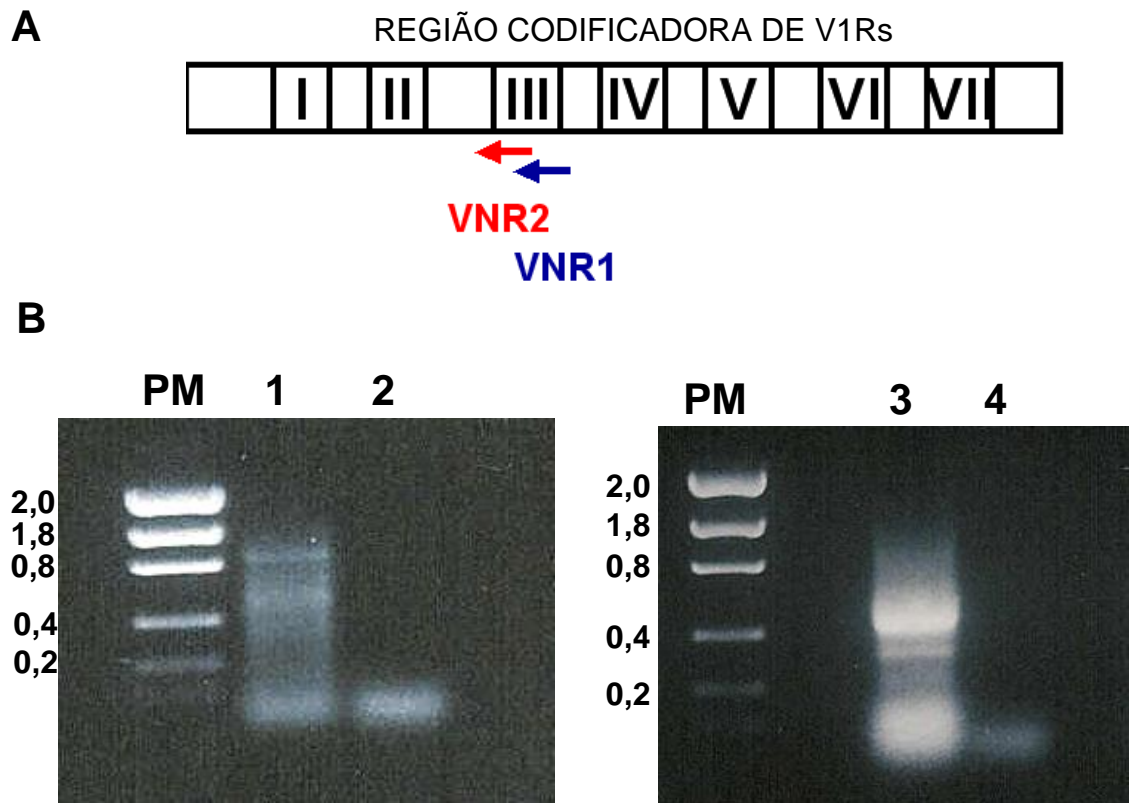
Utilizamos RNA total proveniente de órgão vomeronasal de camundongos para a síntese de cDNA pelo método RLM-RACE.

Na figura 17B (*lane 1*) vemos que os produtos de PCR obtidos com os oligonucleotídeos VNR1 e GeneRacer5' apresentam uma mistura de fragmentos de diferentes tamanhos. Isso mostra, como esperado, que a extremidade 5' de diferentes cDNAs codificantes para V1Rs varia em tamanho. Como a distância entre o oligonucleotídeo VNR1 e o ATG inicial predito para os genes de V1Rs é de aproximadamente 400 pares de base, o tamanho mínimo esperado dos produtos é de 400 pares de base.

Para diminuir a presença de produtos inespecíficos, realizamos um PCR *nested*, utilizando como alvo o produto obtido com os oligonucleotídeos GeneRacer 5' e VNR1 em uma reação de PCR com os oligonucleotídeos GeneRacer 5'*nested* e VNR2. A figura 17B (*lane 3*) mostra este resultado. Bandas com tamanho por volta de 500pb são preferencialmente amplificadas neste segundo PCR. Como o tamanho mínimo esperado para os produtos amplificados com GeneRacer 5' e VNR1 é ~400 pares de base, gel extraímos os produtos com tamanho acima de 400 pb e descartamos os fragmentos inespecíficos de tamanho inferior.

Os produtos de PCR purificados foram subclonados no plasmídeo pCRII e inseridos em bactéria para amplificação. Colônias isoladas de bactérias contendo as seqüências de cDNAs obtidas por RLM-RACE inseridas no plasmídeo, foram selecionadas por PCR de colônia usando os pares de oligonucleotídeos T7 – VNR2 e SP6 – VNR2. O procedimento adotado é semelhante ao PCR de colônia descrito

no item 4.2.1 para a seleção de clones contendo insertos de genes de ORs. Os plasmídeos das colônias selecionadas foram purificados e seqüenciados.



**Figura 17. Seleção de cDNAs 5'UTR completos correspondentes a genes de V1Rs.** A) Esquema representativo das sete regiões transmembrânicas (TM I ao TM VII) dos receptores de feromônios do tipo 1 mostrando os alvos dos oligonucleotídeos degenerados VNR1 e VNR2. B) Eletroforese em gel de agarose 1,5% marcado com brometo de etídeo mostrando os resultados de PCRs obtidos com os primers VNR1 e GeneRacer 5' (lane 1) utilizando cDNA de VNO de camundongo como alvo e com os primers VNR2 e GeneRacer 5' Nested (lane 3) utilizando o produto de PCR mostrado no lane 1 como alvo. Os lanes 2 e 4 são os respectivos controles negativos dos PCRs em 1 e 3. PM: peso molecular ( 2,0 a 0,2 quilo bases).

Como já descrito, para obter uma amostra de cDNAs 5' intactos representativa de todas as 12 famílias que compõem o repertório de genes de V1Rs no genoma de camundongo desenhamos oligonucleotídeos específicos que reconhecem regiões conservadas em diferentes membros de algumas das famílias de genes de V1Rs. De forma semelhante ao que descrevemos na seleção com os oligonucleotídeos degenerados VNR1 e VNR2, os produtos de PCR obtidos com os oligonucleotídeos família-específicos foram gel purificados, subclonados, inseridos em bactéria, selecionados e sequenciados.

#### **4.3.3 Análise das seqüências de cDNA 5'UTR completas de genes de V1Rs**

De um total de quase 650 seqüências analisadas, obtivemos 438 seqüências de cDNAs 5'UTR intactos (67%) correspondendo a 39 genes de V1Rs diferentes abrangendo 26% do repertório de 150 genes funcionais de V1Rs em camundongos.

Na tabela 4 comparamos a distribuição dos genes de V1Rs para os quais obtivemos seqüências de cDNA 5'UTR intactos, com a distribuição de 149 genes de V1Rs preditos nos cromossomos de camundongo retirados do banco de dados do NCBI. Exceto para o cromossomo X que possui somente dois genes de V1R, obtivemos seqüências de genes de V1Rs que estão distribuídas em todos os cromossomos onde se encontram genes de V1Rs, sendo que a maioria das nossas seqüências correspondem a genes de V1Rs no cromossomo 6 onde estão localizados o maior número de genes de V1Rs.

**Tabela 4. Número de genes por cromossomo.** O resultado está dividido em dois grupos: um contendo as regiões codificadoras previstas; outro contendo apenas a amostra de cDNAs obtida.

<b>regiões codificadoras previstas</b>		<b>amostra de cDNAs obtida</b>	
cromossomo	número de genes	cromossomo	número de genes
CHR6	52	CHR6	21
CHR7	47	CHR7	8
CHR13	32	CHR13	8
CHR17	16	CHR17	2
CHRX	2	CHRX	-
Total	149	Total	39

A distribuição do repertório de genes de V1R anotados nas 12 famílias de genes de V1R (Rodriguez *et al.*,2002) assim como a distribuição da nossa amostra de cDNAs é mostrada na tabela 5. Algumas famílias com muitos membros, como, por exemplo, as famílias v1rd e v1re, estão representadas em nossa amostra por somente uma seqüência de cDNA apesar dos oligos família-específicos que desenhamos, em teoria, serem capazes de reconhecer todos os membros dessas famílias. Uma provável explicação para este fato é que o método de RLM-RACE que usamos para gerar cDNAs favorece a amplificação de mensagens muito expressas e também com 3'UTR curto e, portanto, as mensagens menos expressas e mais longas não são amplificadas com freqüência, o que pode ter acontecido com os outros membros das famílias pouco representadas na amostra.

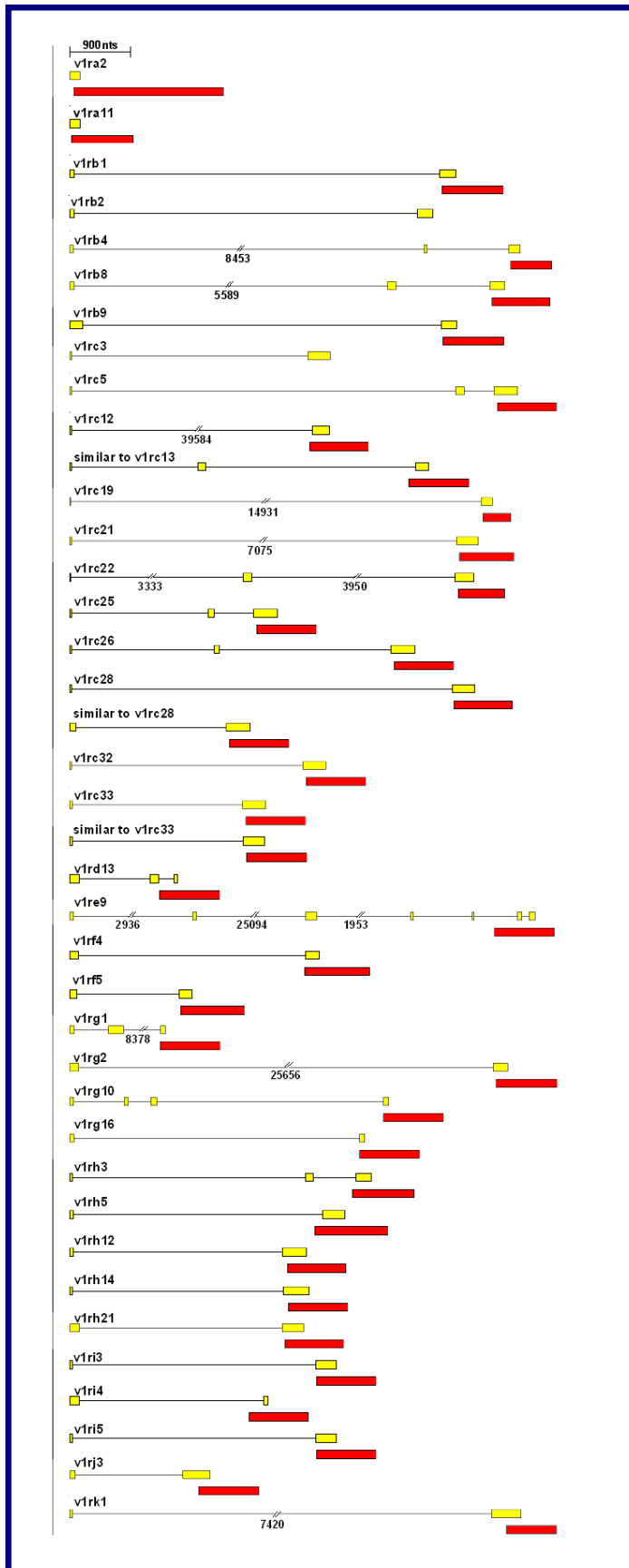
**Tabela 5. Distribuição dos genes de V1Rs obtidos em relação às famílias gênicas.** O repertório de 137 genes de V1Rs preditos foram agrupados em 12 famílias de A até L (\*Rodriguez *et al.*,2002). O número de genes de V1Rs que obtivemos para cada família está indicado na terceira coluna da tabela.

<b>família</b>	<b>repertório de genes V1R anotados*</b>	<b>cDNAs 5' completos obtidos experimentalmente</b>
<b>A</b>	10	1
<b>B</b>	9	5
<b>C</b>	31	14
<b>D</b>	22	1
<b>E</b>	13	1
<b>F</b>	5	2
<b>G</b>	12	4
<b>H</b>	21	5
<b>I</b>	10	3
<b>J</b>	2	1
<b>K</b>	1	1
<b>L</b>	1	0
<b>TOTAL</b>	137	39



As 438 seqüências de cDNA 5' intactas foram alinhadas, pelo aluno de doutorado Pedro Galante, contra o genoma de camundongo, permitindo a análise da estrutura 5' dos genes de V1Rs e a determinação dos TSSs.

A figura 18 mostra em esquema da estrutura 5' dos 39 genes de V1Rs analisados. Uma seqüência representativa de todos os cDNAs 5' intactos que obtivemos para cada um dos 39 diferentes genes de V1R é mostrada na figura 18. Os éxons estão representados pelos retângulos amarelos e os íntrons pelas linhas pretas. Para auxiliar o mapeamento dos genes de V1R, seqüências preditas depositadas no banco de dados do NCBI foram também utilizadas no alinhamento e estão representadas, na figura 18, como retângulos vermelhos. Em alguns casos, como claramente mostrado para os genes v1rd13 e v1re9, foi constatado que houve erro na predição das seqüências codificadoras dos genes de V1Rs. Em muitos casos o ATG predito não corresponde ao ATG experimental.



**Figura 18. Representação esquemática da estrutura 5' dos cDNAs obtidos pelo método RLM-RACE correspondentes a 39 diferentes genes de V1R.**

As regiões codificadoras preditas dos genes de V1R estão representadas pelos retângulos vermelhos. Os éxons das sequências de cDNAs obtidas estão representados pelos retângulos amarelos e os íntrons pelas linhas pretas. O tamanho dos íntrons fora da escala está indicado em pares de base. A localização exata dos sítios de início da transcrição (TSSs) foi determinada pelo alinhamento destas sequências contra o genoma de camundongo e corresponde à posição do primeiro nucleotídeo do éxon mais a 5', no esquema localizado à esquerda.

A tabela 6 apresenta os resultados obtidos em relação à região 5' não traduzida. A maioria dos genes de V1R apresenta um único éxon e íntron na região 5' não traduzida (5'-UTR), assim como os genes de ORs. Um total de 37 dos 39 genes de V1Rs apresentam ao menos 1 éxon e um íntron na região 5'UTR. O tamanho médio dos éxons é de 94 pares de base (pb) variando entre 22 a 518 pb. Os íntrons da região 5' apresentam uma grande variação em tamanho de 237 a 39876 pb. A grande maioria dos genes de V1R apresenta íntrons entre 2000 a 4000 pb.

Somente para três genes de V1Rs em nossa amostra analisada detectamos eventos de splicing alternativo sendo eles v1rc28, V1rc5 e V1re9.

**Tabela 6. Estrutura da região 5' dos genes de V1R analisados.** A maioria dos genes de V1R apresenta um único éxon e íntron na região 5' não traduzida (5'UTR). O tamanho médio dos éxons é de 94 pares de base (pb) variando entre 22 a 518 pb. Os íntrons da região 5' apresentam uma grande variação em tamanho de 237 a 39876 pares de base. A grande maioria dos genes de V1R apresenta íntrons entre 2000 a 4000 pares de base.

éxon 5'UTR		íntron	
tamanho (pb)	freqüência	tamanho (kb)	freqüência
< 50	8	< 0,5	5
50 –100	20	0,5 – 1,0	3
100 –150	8	1,0 – 2,0	6
150 –200	2	2,0 – 4,0	24
> 200	15	4,0 – 8,0	9
		> 0,8	6

#### 4.3.4 Busca de elementos conservados na região promotora dos genes de V1Rs.







Determinamos experimentalmente a posição exata dos TSSs para 39 genes de V1Rs diferentes. Seqüências de 1000 nucleotídeos a montante dos TSSs dos 39 genes de V1Rs obtidos foram extraídas e utilizadas na busca de elementos conservados.

Buscamos motivos conservados, nas 39 regiões promotoras, utilizando os programas *Gibbs recursive sampler* (Thompson *et al.*, 2003), Consensus (Hertz & Stormo, 1999) e Weeder (Pavesi *et al.*, 2004) que possuem algoritmos de busca diferentes.

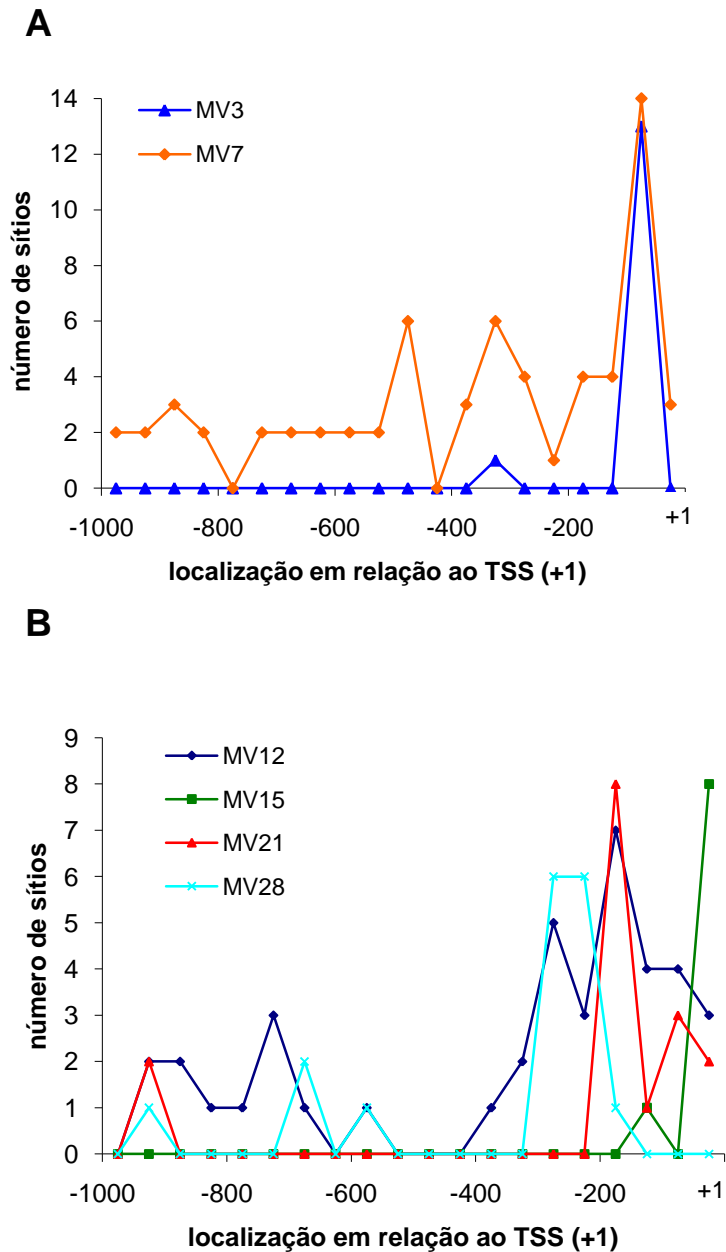
Selecionamos seis motivos. Os critérios para a seleção desses motivos foram: 1) aparecerem com alta freqüência em diferentes promotores analisados; 2) estarem concentrados próximos aos TSSs; 3) serem motivos diferentes, ou seja, apresentarem consensos diferentes. Os motivos selecionados estão representados pelos Weblogos na tabela 7.

Os motivos MV3 e MV7 assemelham-se aos sítios de ligação de proteínas O/E os quais também foram identificados na região promotora dos genes de ORs. Já os motivos MV12, MV15, MV21 e MV28 não se assemelham a sítios de ligação conhecidos presentes nos bancos de dados de fatores de transcrição Jaspar e Transfac que foram consultados utilizando o programa STAMP (Mahony & Benos, 2007) disponível no endereço eletrônico: <http://www.benoslab.pitt.edu/stamp/>.

**Tabela 7. Motivos selecionados na região promotora dos 39 genes de V1R analisados.** Os motivos selecionados estão indicados na tabela pela representação obtida com o programa Weblogo. A seleção baseou-se na escolha dos motivos mais frequentes em diferentes promotores e que estavam concentrados próximos aos TSSs.

motivo	logos	Número de genes V1Rs que contêm os motivos nas regiões promotoras (% do total de 39 promotores)
MV3		14 (36%)
MV7		35 (90%)
MV12		30 (77%)
MV15		9 (23%)
MV21		16 (41%)
MV28		17 (43%)

A seguir, nos gráficos da figura 19, mostramos a distribuição desses motivos nas regiões promotoras pesquisadas. Contamos o número de vezes que cada motivo aparece em intervalos de 50 pares de base, do TSS (+1) até -1000, na região promotora utilizando o programa *SiteSeer* (Boardman *et al.*, 2003) disponível no endereço eletrônico <http://rocky.bms.umist.ac.uk/SiteSeer/>. Este programa permite visualizar a posição ocupada pelo motivo em todas as seqüências promotoras. O gráfico da figura 19A mostra a distribuição dos motivos MV3 e MV7, que se assemelham a sítios de ligação de proteínas O/E. Estes motivos aparecem mais concentrados nas regiões entre -200 e +1 dos 39 promotores analisados assim como os motivos semelhantes a sítios O/E anteriormente identificados nos promotores de genes de ORs. Os motivos MV12, MV15, MV21 e MV28 mostrados na figura 19B, também estão concentrados próximos aos TSSs.



**Figura 19. Gráfico da distribuição dos motivos na região de  $-1000$  até o TSS ( $+1$ ) dos genes de V1R analisados.** Cada ponto no gráfico representa o número de motivos presentes naquele intervalo para o total de 39 promotores analisados. Os motivos aparecem mais concentrados na região entre  $-300$  e  $+1$  (TSS). A) Os motivos MV3 e MV7 assemelham-se aos sítios de ligação *O/E-like* que anteriormente identificamos nos promotores de genes de receptores olfatórios. B) Os motivos MV12, MV15, MV21 e MV28 não se assemelham a sítios de ligação de fatores de transcrição conhecidos e também estão concentrados próximos aos TSSs.

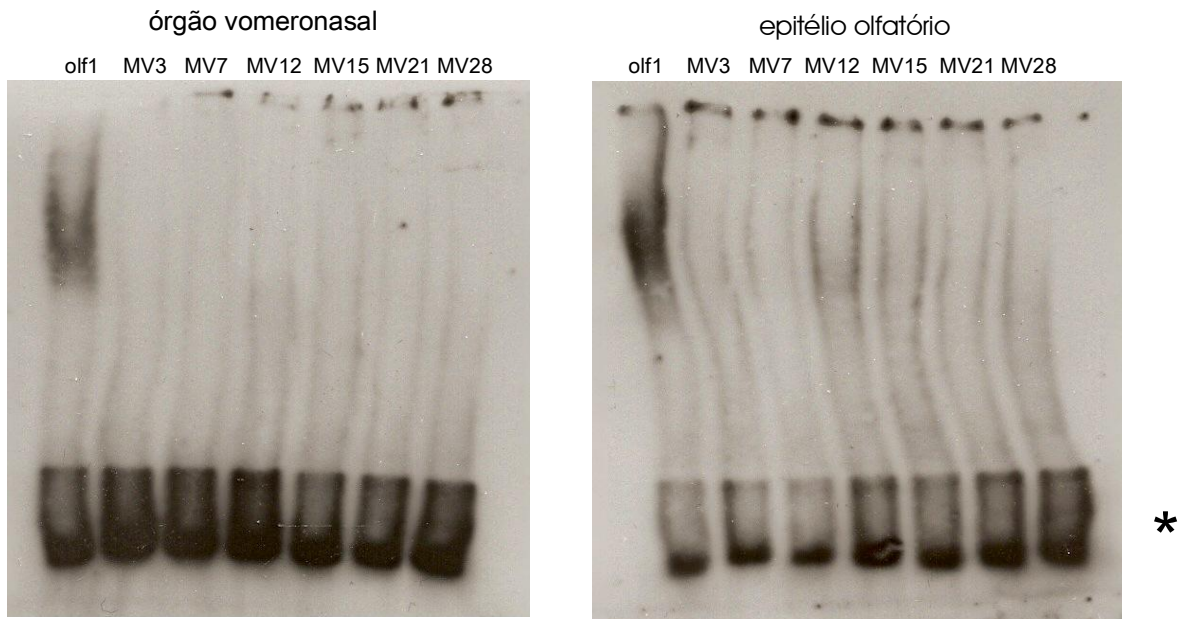
#### **4.3.5 Ensaios de EMSA com os motivos encontrados nos promotores de V1Rs.**

Realizamos ensaios de EMSA, semelhante ao apresentado no item 4.2.4 nos resultados da análise dos motivos encontrados em promotores de genes de ORs. Como já mencionado, esses ensaios de EMSA permitem observar interações entre DNA e proteínas de extratos proteicos nucleares de tecidos.

Oligonucleotídeos marcados contendo seqüências representativas para os seis motivos selecionados, sendo esses MV3 e MV7 (sítio de ligação de O/E) e MV12, MV15, MV21 e MV28, foram usados nos ensaios de ligação.

Os oligonucleotídeos foram incubados com extratos protéicos nucleares de órgão vomeronasal (VNO) ou de epitélio olfatório (OE) e, a seguir, a amostra foi resolvida em gel não denaturante de poliacrilamida 4%. Como controle dos extratos nucleares de VNO e OE, usamos o sítio de ligação para o fator de transcrição OLF-1 (olf1) que é expresso tanto em neurônios olfatórios quanto em neurônios vomeronasais.

Como mostrado na figura 19, não visualizamos interações entre os motivos MV3, MV7, MV15, MV21 e MV28 com proteínas nucleares de VNO e de OE. O motivo MV12 formou um complexo com proteínas nucleares de OE. A interação do motivo MV12 com proteínas nucleares de VNO, que só pode ser visualizada após longa exposição do filme fotográfico, é mostrada com mais clareza na figura 20.

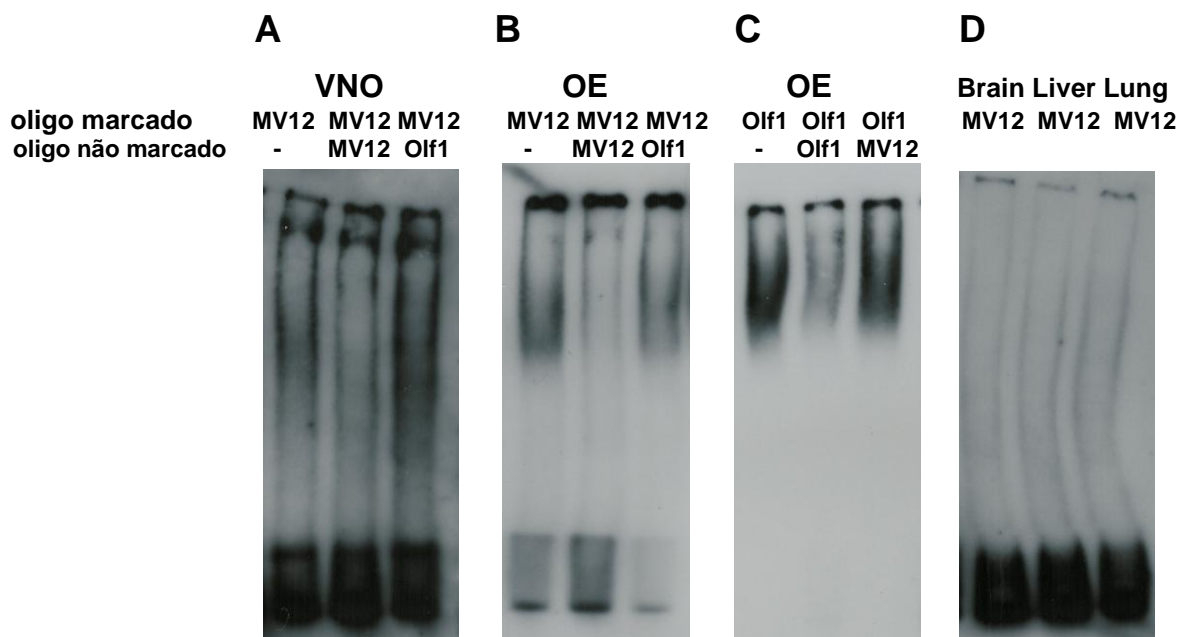


**Figura 19. Interação entre os motivos de DNA e proteínas nucleares de VNO e OE.** Ensaio de mobilidade eletroforética em gel não denaturante de poliacrilamida 4%. Oligonucleotídeos dupla-fita, marcados com digoxigenina, correspondendo aos motivos selecionados MV3, MV7, MV12, MV15, MV21 e MV28 foram incubados com 10  $\mu$ g de extratos nucleares de órgão vomeronasal (VNO) e de epitélio olfatório (OE) de camundongo. Os oligonucleotídeos marcados livres estão indicados pelo asterisco. Em extrato nuclear de VNO, a formação de complexo entre DNA e proteínas na presença do oligo MV12 é identificada fracamente após exposição de 30 minutos do filme fotográfico na etapa de detecção do ensaio. Ensaio posterior (figura 20), com quantidades de proteína nuclear aumentada mostrarão com mais clareza a interação observada entre MV12 e proteínas de extrato nuclear no VNO. O oligonucleotídeo contendo o sítio de ligação para o fator de transcrição OLF1, abundante no OE e VNO, foi usado como controle dos extratos nucleares. No ensaio realizado com extrato nuclear de OE é evidente a presença de complexos entre DNA e proteínas para os motivos olf1 e MV12.

A figura 20 mostra que o oligonucleotídeo MV12 forma complexos com proteínas nucleares de VNO e também com proteínas nucleares de OE. Os complexos formados por MV12 marcado são deslocados quando um excesso de oligonucleotídeo MV12 não marcado é adicionado às reações de ligação, indicando a especificidade da formação desses complexos (figuras 20A e B). Não observamos o deslocamento de MV12 marcado dos complexos entre DNA e proteínas quando



excesso de olf1 não marcado é adicionado (figuras 20A e B). Extratos nucleares de cérebro, fígado e pulmão não interagem com o motivo MV12 (figura 20D). Observamos, na figura 20C, que o complexo entre proteínas e o oligo olf1 não é deslocado por excesso de MV12 não marcado. Estes resultados indicam que MV12 é alvo de proteínas nucleares específicas do sistema olfatório (VNO e OE) e não relacionadas ao fator de transcrição OLF1.

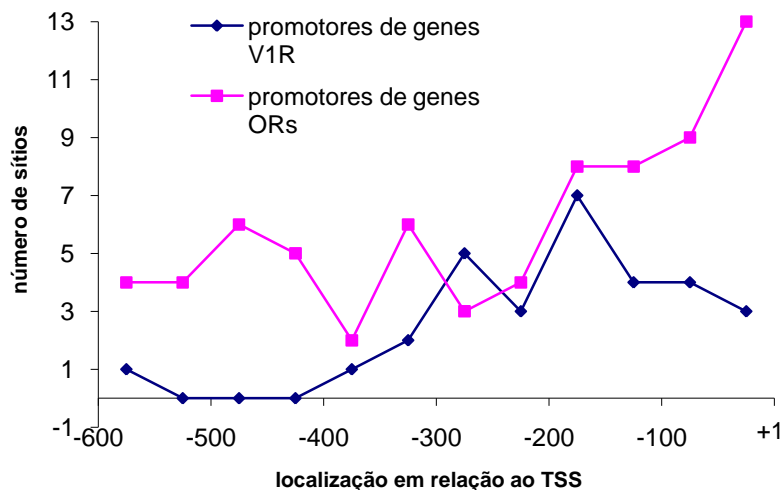


**Figura 20. Ensaio de EMSA com competidor específico e inespecífico.** Nestes ensaios foram utilizados 30µg de extratos de proteínas nucleares. A-B) Os complexos formados pelo motivo MV12 e proteínas nucleares de VNO (A) e de OE (B) de camundongo são especificamente deslocados pela adição de excesso de oligonucleotídeo MV12 não marcado, no entanto, não são deslocados pelo excesso de oligonucleotídeo inespecífico olf1 não marcado. C) Da mesma forma o complexo entre DNA e proteínas formado pelo motivo olf1 não é deslocado por excesso do oligonucleotídeo MV12 não marcado. D) Não se observa formação de complexos entre MV12 e proteínas nucleares de cérebro (Brain), fígado (Liver) e pulmão (Lung). O tempo de exposição do filme fotográfico na detecção foi de 40 minutos.

A descoberta de que o motivo MV12, presente na maioria dos promotores de genes de V1Rs, é capaz de interagir com proteínas nucleares tanto de neurônios vomeronasais quanto de neurônios olfatórios levou-nos a pesquisar a presença e distribuição do motivo MV12 nos promotores de genes de ORs.

Encontramos o motivo MV12 presente em 30 dos 39 (77%) promotores de genes de V1Rs, e em 104 dos 198 (52%) promotores de genes de ORs. O gráfico da figura 21 mostra que MV12 aparece com maior frequência nas regiões promotoras próximas aos TSSs dos promotores analisados, entre  $-200$  a  $+1$ .

O fato de não havermos identificado o motivo MV12 na análise anterior dos promotores de genes de ORs pode ser devido à alta frequência de aparecimento de sítios de ligação de proteínas O/E nesses promotores. Além de serem bastante abundantes, aparecendo em 87% do total de 198 promotores de genes de ORs analisados, esses sítios O/E são divididos em 4 grupos, conforme a conservação de diferentes nucleotídeos. Isso acaba favorecendo o aparecimento de muitos resultados (*outputs*) com alta frequência para os diferentes sítios O/E gerados pelos programas de busca *Gibbs Recursive Sampler* e *Consensus* enquanto que outros sítios, comparativamente menos frequentes, deixam de ser selecionados.



**Figura 21. Gráfico da distribuição do motivo MV12 na região de  $-600$  até o TSS ( $+1$ ) dos genes de ORs e de V1Rs.** Cada ponto no gráfico representa o número do motivo MV12 presente naquele intervalo para o total de 39 promotores de genes de V1Rs e 198 promotores de genes de ORs analisados. O motivo aparece mais concentrado na região entre  $-200$  e  $+1$  (TSS).

Em conclusão, os resultados aqui apresentados indicam que genes de V1Rs e genes de ORs compartilham elementos comuns em suas regiões promotoras e que esses elementos são sítios de ligação de proteínas nucleares presentes em ambos os tecidos do sistema olfatório.

## 5) DISCUSSÃO

### 5.1. Motivos presentes em promotores de genes de ORs e de V1Rs.

Em nosso trabalho, realizamos uma análise na qual regiões a 5' de genes de ORs e de genes de V1Rs foram comparadas em busca de elementos promotores conservados, potencialmente envolvidos no mecanismo de regulação de expressão desses genes.

Utilizamos a estratégia de RLM-RACE combinada com o uso de oligonucleotídeos degenerados que reconhecem seqüências conservadas entre diversos membros das famílias de genes de ORs e de V1Rs; utilizamos também oligonucleotídeos específicos para diferentes famílias de genes de V1Rs. Esta abordagem experimental permitiu-nos gerar cDNAs contendo a região 5'UTR intacta para 198 diferentes genes de ORs, que corresponde a 17% do repertório de genes ORs, e para 39 diferentes genes de V1Rs, abrangendo 26% do repertório de genes de V1R. O alinhamento dessas seqüências contra o genoma de camundongo forneceu a estrutura gênica 5' e a exata localização dos sítios de início da transcrição. Seqüências genômicas próximas, a montante dos TSSs, foram extraídas e pesquisadas em busca de motivos de DNA conservados.

Do total de 198 regiões promotoras de genes de ORs analisadas, 87% contêm sítios de ligação para fatores de transcrição O/E, e 95% contêm sítios de ligação de proteínas com homeodomínios. O papel dos fatores de transcrição que interagem com os sítios de ligação O/E e de homeodomínios na regulação de expressão dos genes de ORs não está esclarecido.

Sítios de ligação para proteínas O/E já haviam sido identificados em genes expressos especificamente em neurônios olfatórios como os genes *Omp*, *OcNC*,

*AcIII*, *Golf* (Kudrycki *et al.*, 1993; Wang *et al.*, 1993) e em alguns genes de ORs (Rothman *et al.*, 2005; Vassalli *et al.*, 2002). Os quatro membros da família de fatores de transcrição O/E são expressos em neurônios olfatórios maduros e nas células progenitoras desses neurônios (Wang *et al.*, 1997; Wang *et al.*, 2002). A função desses fatores de transcrição ainda não está clara. Ensaios de duplo-híbrido em levedura, mostraram que O/E1 interage com a proteína ROAZ (*Rat O/E1-associated zinc finger*). ROAZ é expressa somente na camada de células basais no epitélio olfatório e sua interação com O/E1 impede que O/E1 se ligue ao sítio de DNA alvo (Tsai *et al.* 1997). Foi proposto que ROAZ funcione como um regulador negativo da atividade transcricional de O/E1 nos neurônios imaturos impedindo que nesses neurônios sejam expressos genes de neurônios diferenciados.

O camundongo *knockout* para O/E1 (*Olf1*) não possui alteração na expressão de genes de ORs (Lin & Grosschedl, 1995). Provavelmente este fenótipo seja devido ao efeito redundante de outros membros da família de proteínas O/E. Já os camundongos duplo *knockout* para O/E2 e O/E3 apresentam defeitos na projeção dos axônios dos neurônios olfatórios da zona mais dorsal do epitélio para a região mais dorsal do bulbo olfatório. O bulbo olfatório é a região do cérebro que primeiro recebe a informação sensorial trazida pelos axônios dos neurônios olfatórios. No entanto, os neurônios olfatórios em *O/E2<sup>-/-</sup>* e *O/E3<sup>-/-</sup>* expressam *Omp*, *OcNC* e genes ORs da zona dorsal mostrando que a maquinaria de transdução de sinal está presente nos neurônios que não expressam O/E2 e O/E3 (Wang *et al.*, 2003). Uma explicação para a expressão desses genes continua sendo a ação redundante das outras O/E (O/E1 e O/E4) mas, neste caso, a falta de projeção de axônios de neurônios olfatórios para o bulbo pode indicar uma função não redundante para O/E2 e O/E3.

Sítios de ligação para proteínas com homeodomínio e proteínas O/E (Olf1) foram identificados na região promotora mínima do gene de OR M71 (Rothman *et al.*, 2005). A mutação desses dois sítios no locus endógeno do OR M71 reduz o número de neurônios que expressam o gene OR M71 e o padrão de expressão desse OR no epitélio olfatório é alterado indicando que esses sítios podem ter um papel na expressão desse receptor no epitélio olfatório.

Havia sido demonstrado que todas as proteínas O/E interagem igualmente com o sítio de ligação de Olf-1 (Wang *et al.*, 1997; Wang *et al.*, 2002). Identificamos sítios de ligação de O/E que possuem diferentes nucleotídeos conservados formando quatro grupos e observamos que esses quatro motivos O/E interagem diferentemente com proteínas de extrato nuclear do epitélio olfatório. Estas diferenças estruturais podem ser responsáveis por diferentes afinidades desses sítios pelas diferentes O/E. Sabe-se que diferentes receptores olfatórios apresentam diferentes níveis de expressão (Young *et al.*, 2003). Uma explicação para esta variação no nível de expressão de diferentes genes de ORs é que os tipos de sítios de ligação de O/E na região promotora de genes de OR determinam a probabilidade de escolha de um gene de OR ser transcrito.

Como já mencionado, sítios de ligação para fatores com homeodomínios foram identificados em 95% dos promotores de genes de ORs analisados. Ensaios *in vitro* mostraram que a proteína com homeodomínio Lhx2 interage com o sítio de ligação de proteínas com homeodomínio presente no promotor do gene OR M71 (Hirota & Mombaerts, 2004; Rothman *et al.*, 2005) e mutações no sítio de ligação de Lhx2 dentro do promotor do gene OR M71 comprometem a expressão deste gene (Rothman *et al.*, 2005).

A proteína com homeodomínio Lhx2 é expressa fortemente na camada de células basais do epitélio olfatório, onde se encontram as células precursoras dos neurônios olfatórios maduros, e fracamente na camada neuronal onde estão os neurônios olfatórios maduros. O camundongo *knockout* para Lhx2 morre alguns dias após o nascimento devido à anemia severa. Em embriões E16.5 os neurônios olfatórios já expressam genes de ORs. Nos embriões E16.5 Lhx2<sup>-/-</sup> não há expressão de genes ORs assim como não há expressão de OMP (marcador de neurônios maduros) o que sugere que o fator Lhx2 é necessário nos estágios iniciais de desenvolvimento dos neurônios. Estes dados não são fortes evidências de que Lhx2 é necessária para a escolha de genes de ORs nos neurônios olfatórios mas sugerem que Lhx2 pode estar envolvida com o desenvolvimento do neurônio olfatório antes dele atingir o estágio de diferenciação em que passa a expressar determinado gene de OR.

Também identificamos sítios de ligação de O/E em 90% dos 39 promotores de genes de V1Rs analisados. A amostra de 39 promotores de genes de V1Rs, obtida experimentalmente, é a maior até hoje analisada e, pela primeira vez, sítios de ligação de O/E são descritos em promotores de genes V1Rs. Apesar da expressão de O/E-1, O/E-2 e O/E-3 no VNO já ter sido relatada (Wang *et al.*, 1997 e 2002) e existirem, como citamos acima, trabalhos analisando a expressão de genes de ORs nos *knockouts* para esses fatores de transcrição, nenhuma análise quanto ao padrão de expressão dos genes de V1Rs foi feita. O fato de não observarmos a interação de proteínas nucleares de VNO com as seqüências de DNA escolhidas (MV3 e MV7) como representantes dos motivos semelhantes a sítios O/E identificados nos promotores de V1Rs não descarta os sítios O/E como alvos dos fatores de transcrição O/E em neurônios vomeronasais. Observamos que a

seqüência *olf1*, utilizada como controle do extrato nuclear de VNO, apesar de não ter sido extraída de um promotor de gene V1R e sim ter sido retirada do promotor do gene que codifica o canal ligado a nucleotídeos cíclicos (*OcNC*) expresso somente em neurônios olfatórios (Wang *et al.*, 1993), interage com proteínas do extrato de VNO. Isto indica que as proteínas O/E expressas no VNO reconhecem sítios O/E. Mostramos que diferentes sítios O/E extraídos de promotores de genes ORs interagem diferentemente com extrato nuclear de epitélio olfatório (figura 14A do item 4-resultados). Uma possibilidade é que os motivos MV3 e MV7 interajam fracamente com proteínas O/E do VNO e não tenhamos visualizado esta interação. Outra possibilidade é que as seqüências MV3 e MV7 escolhidas contenham mutações no sítio O/E que abolem a interação esperada. Nossos resultados mostram que os sítios O/E são abundantes e aparecem concentrados próximos aos TSSs dos promotores V1Rs pesquisados num padrão semelhante ao observado nos promotores de genes ORs.

Nossos resultados indicam que os fatores de transcrição O/E são encontrados com freqüência em regiões promotoras proximais de genes de ORs e V1Rs. Esses dados indicam que os fatores de transcrição O/E possivelmente desempenham um papel na expressão de genes ORs e V1Rs.

Identificamos também a presença de um motivo de DNA, chamado MV12, em 77% dos promotores de genes de V1Rs e em 52% dos promotores de genes de ORs. O motivo MV12 não se assemelha a nenhum sítio de ligação de fatores de transcrição conhecido. Ensaio de interação *in vitro* indicam que MV12 interage com proteínas nucleares extraídas de epitélio olfatório e órgão vomeronasal mas não interage com proteínas nucleares de outros tecidos como cérebro, pulmão e fígado.



Pela primeira vez, um potencial sítio de ligação para fatores de transcrição desconhecido e conservado em ambos promotores de genes de ORs e V1Rs é relatado. As proteínas que se ligam a este sítio podem ser fatores de transcrição ainda desconhecidos que desempenham papel comum no mecanismo de regulação da expressão dos genes de ORs e V1RS. Nossos resultados mostram que existem proteínas nos núcleos de neurônios olfatórios e de neurônios vomeronasais que são capazes de se ligar a este sítio. A caracterização destas proteínas deverá contribuir para a elucidação do papel desempenhado pelo motivo MV12.

## **5.2. Regulação da expressão dos genes de ORs e de V1Rs.**

O mecanismo que permite que somente um único alelo de um único gene de OR ou de V1R seja expresso em cada neurônio olfatório ou vomeronasal ainda não foi esclarecido.

Trabalhos anteriores sugeriram um modelo de regulação muito interessante no qual se propõe a existência de um sítio único de expressão para os genes de ORs no núcleo, o que ajuda a explicar a escolha de um único gene de OR por neurônio (Lomvardas *et al.*, 2006). Experimentos de 3C (*chromosome conformation capture*) e FISH indicam que uma região conservada de 2kb próxima a um *cluster* de genes ORs no cromossomo 14 de camundongo, chamada de região H (Lane *et al.*, 2002; Serizawa *et al.*, 2003) está fisicamente associada com o gene de OR escolhido para expressão, independente do cromossomo onde se localiza o gene de OR expresso (Lomvardas *et al.*, 2006). Estes dados sugeriam que a região H era o regulador chave de todos os genes de ORs. Entretanto, esse modelo que explicava de maneira elegante a expressão monoalélica e mutuamente exclusiva dos genes de ORs foi derrubado quando se observou que o camundongo no qual a

região H é eliminada continua a expressar genes de ORs tendo somente os genes próximos, localizados em *cis* à região H, comprometidos em sua expressão (Fuss *et al.*, 2007; Nishizumi *et al.*, 2007). Apesar da região H não atuar como regulador geral da expressão de genes ORs, os resultados apresentados por Lomvardas e colaboradores mostram que a região H encontra-se fisicamente próxima ao gene de OR transcricionalmente ativo no núcleo dos neurônios olfatórios. Isto levanta a hipótese de que no núcleo dos neurônios olfatórios possa existir um único sítio ativo de transcrição no qual todos os elementos necessários à transcrição dos genes de ORs estejam reunidos e elementos reguladores como *enhancers* e promotores tenham que competir por este sítio único.

Um exemplo de modelo de controle de regulação de expressão por um único sítio ativo de transcrição é sugerido para os genes *var* no *Plasmodium falciparum*. No plasmódio, a família de 60 genes *var* expressa proteínas na superfície de eritrócitos de forma semelhante aos genes de ORs nos neurônios olfatórios: um único gene é ativado para transcrição enquanto que o restante permanece silenciado. No plasmódio ainda ocorre variação antigênica, ou seja, um gene da família de genes *var* pode ser ligado e depois desligado quando outro gene *var* é selecionado para a transcrição. Resultados recentes mostraram que elementos localizados *upstream* dos genes *var* aparentemente contêm toda a informação necessária para ativar genes *var* de modo mutuamente exclusivo (Voss *et. al*, 2006). Além disso, evidências sugerem que exista um único sítio ativo de transcrição na periferia do núcleo. Deste modo, um único gene *var* pode ser ativado movendo-se para este sítio ativo. Esse sítio ativo só pode acomodar um único promotor de gene *var* explicando porque todos os outros genes *var* permanecem silenciados.

No modelo proposto para a regulação da expressão dos genes ORs, cada gene OR seria controlado por uma região regulatória, semelhante à região H. Estas regiões estariam localizadas num único sítio específico do núcleo onde uma única maquinaria transcricional ativaria a expressão de um único gene OR (Johnston *et al.*, 2008). Elementos *enhancers* como o região H não foram ainda identificados em outros *clusters* de genes de ORs mas é interessante que a presença de sítios de ligação O/E e sítios de ligação para proteínas com homeodomínios, além de terem sido identificados na região promotora de vários genes de ORs, aparecem também na região H (Hirota & Mombaerts, 2004) indicando que fatores de transcrição comuns podem ajudar na interação destes elementos.

Em nosso trabalho mostramos que genes de ORs e genes de V1Rs apresentam elementos conservados em suas regiões promotoras e que estes interagem com proteínas nucleares expressas no sistema olfatório. Como estes elementos, assim como as proteínas que se ligam a eles, contribuem para a regulação da expressão dos genes OR e V1R ainda não está claro. Uma possibilidade é que eles sejam necessários para a repressão destes genes. Outra possibilidade é que eles direcionem o gene para um sítio nuclear único onde a transcrição ocorrerá. Portanto, a identificação de elementos conservados nas regiões promotoras dos genes de OR e V1R, assim como a identificação das proteínas que se ligam a esses elementos deverá contribuir para a elucidação deste mecanismo peculiar de regulação gênica.

## 6) CONCLUSÃO

A partir dos resultados obtidos neste trabalho concluímos que:

- A estratégia de comparar regiões promotoras de genes co-regulados, como é o caso dos genes de ORs e dos genes de V1Rs, mostrou-se útil na identificação de elementos comuns potencialmente envolvidos na regulação da expressão desses genes;
- Regiões promotoras de genes de ORs e de genes de V1Rs compartilham sítios de ligação de fatores O/E e sítios de ligação de fatores ainda não identificados indicando que elementos reguladores comuns estão envolvidos no mecanismo de expressão dos genes de ORs e V1Rs;
- Proteínas nucleares extraídas de epitélio olfatório e órgão vomeronasal interagem com elementos comuns presentes nos promotores de genes ORs e V1Rs reforçando o papel biológico destes elementos *cis*-atuantes.

## 7) REFERÊNCIAS

- Buck, L. B. 2000. The Molecular Architecture of Odor and Pheromone Sensing in Mammals, *Cell* **100**: 611-618.
- Buck, L. 2004. Olfactory receptors and odor coding in mammals. *Nutrition Rev.* **62**: S184-S188.
- Buck, L. and R. Axel. 1991. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**: 175-187.
- Boardman P.E., Oliver S.G., Hubbard S.J. 2003. SiteSeer: visualisation and analysis of transcription factor binding sites in nucleotide sequences. *Nucleic Acids Res.* **31**:3572–3575.
- Chess, A., I. Simon, H. Cedar, and R. Axel. 1994. Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78**: 823-834.
- Dulac, C., Wagner S. 2006. Genetic analysis of brain circuits underlying pheromone signaling. *Annu. Rev. Genet.* **40**: 449-67.
- Dulac, C. and A.T. Torello. 2003. Molecular detection of pheromone signals in mammals: from genes to behaviour. *Nat. Reviews Neuroscience* **4**:551-562.
- Eggan, K., K. Baldwin, M. Tackett, J. Osborne, J. Gogos, A. Chess, R. Axel, and R. Jaenisch. 2004. Mice cloned from olfactory sensory neurons. *Nature* **428**:44-49.
- Fuss, S. H., Momura, M., Mombaerts, P. 2007. Local and cis Effects of the H Element on Expression of Odorant Receptor Genes in Mouse. *Cell.* **130(2)**:373-84.
- Godfrey, P.A., B. Malnic, and L.B. Buck. 2004. The mouse olfactory receptor gene family. *Proc Natl Acad Sci U S A* **101**: 2156-2161.
- Hertz, G. Z.& Stormo, G. D. 1999. Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* **15**: 563-577.
- Herrada G. and Dulac C. 1997. A novel family of putative pheromone receptors in mammals with a topographically organized and sexually dimorphic distribution. *Cell.* **90(4)**:763-73.
- Hirota, J., Mombaerts, P. 2004. The LIM-homeodomain protein Lhx2 is required for complete development of mouse olfactory sensory neurons. *Proc Natl Acad Sci U S A* **101**:8751- 8755.
- Ishii, T., S. Shou, A. Kohda, H. Nakatani, T. Shiroishi, K. Okumura, Y. Iwakura, F. Nagawa, A. Tsuboi, and H. Sakano. 2001. Monoallelic expression of the odorant receptor gene and axonal projection of olfactory sensory neurons. *Genes to Cells* **6**: 71-78.

- Johnston, R.J., Desplan, C. 2008. Stochastic neuronal cell fate choices. *Curr Opin Neurobiol* **18**:1-8.
- Kambere, M.J., and Lane, R. P. 2007. Co-regulation of large and rapidly evolving repertoires of odorant receptor genes. *BMC Neurosci.* 8 (Suppl 3): S2.
- Kozak, M. 1999. Initiation of translation in prokaryotes and eukaryotes. *Gene* **234**:187-208.
- Kudrycki, K., C. Stein-Izsak, C. Behn, M. Grillo, R. Akesson, and F. Margolis. 1993. Olf-1 binding site: characterization of an olfactory neuron-specific promoter motif. *Molecular and Cellular Biology* **13**: 3002-3014.
- Lane R.P., Cutforth T., Axel R., Hood L., Trask B.J. 2002. Sequence analysis of the mouse vomeronasal receptor gene clusters reveals common promoter motifs and a history of recent expansion. *PNAS* **99**(1): 291-6.
- Lane, R.P., Young J., Newman T., Trask B.J. 2004. Species specificity in rodent pheromone receptor repertoires. *Genome Res* **14**(4):603-608.
- Lewcock, J.W. and R.R. Reed. 2004. A feedback mechanism regulates monoallelic odorant receptor expression. *Proc Natl Acad Sci U S A* **101**: 1069-1074.
- Li, J., T. Ishii, P. Feinstein, and P. Mombaerts. 2004. Odorant receptor gene choice is reset by nuclear transfer from mouse olfactory sensory neurons. *Nature* **428**: 393-399.
- Lin, H., Grosschedl, R., 1995. Failure of B-cell differentiation in mice lacking the transcription factor EBF. *Nature* **376**:263-267.
- Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R. 2006. Interchromosomal interactions and olfactory receptor choice. *Cell* **126**(2):403-413.
- Mahony S, Benos PV. 2007. STAMP: a web tool for exploring DNA-binding motif similarities. *Nucleic Acids Res* **35**: W253-8.
- Malnic, B., J. Hirono, T. Sato, and L.B. Buck. 1999. Combinatorial receptor codes for odors. *Cell* **96**: 713-723.
- Matsunami H, Buck LB. 1997. A multigene family encoding a diverse array of putative pheromone receptors in mammals. *Cell*.**90**(4):775-84.
- Matsunami, H., Amrein, H. 2003. Taste and pheromone perception in mammals and flies. *Genome Biol.* **4**(7):220.
- Nishizumi, H., Kumasaka, K., Inoue, N., Nakashima, A., Sakano, H. 2007. Deletion of the core-H region in mice abolishes the expression of the three proximal odorant receptor genes in cis. *Proc Natl Acad Sci U S A* **104**:20067-20072.

- Hozumi, N.; Tonegawa, S. 1976. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc Natl Acad Sci U S A* **73**:3628-3632.
- Pavesi, G., Mereghetti, P., Mauri, G. & Pesole, G. 2004. Weeder Web: Discovery of transcription factor binding sites in a set of sequences from co-regulated genes. *Nucleic Acids Res.* **32**, W199-W203.
- Pesole G, Liuni S, Grillo G, Saccone C. (1997). Structural and compositional features of untranslated regions of eukaryotic mRNAs. *Gene*:**205**:95-102.
- Qasba, P. and R.R. Reed. 1998. Tissue and zonal-specific expression of an olfactory receptor transgene. *J. Neuroscience* **18**: 227-236.
- Ressler, K.J., S.L. Sullivan, and L.B. Buck. 1993. A zonal organization of odorant receptor gene expression in the olfactory epithelium. *Cell* **73**: 597-609.
- Rodriguez I, Feinstein P, Mombaerts P. 1999. Variable patterns of axonal projections of sensory neurons in the mouse vomeronasal system *Cell*. **97(2)**:199-208.
- Rodriguez I, Del Punta K., Rothman A., Ishii T., Mombaerts P. 2002. Multiple new and isolated families within the mouse superfamily of V1r vomeronasal receptors. *Nat Neurosci.***5(2)**:134-40.
- Roppolo, D., Vollery, S., Kan, C., Luscher, C., Broillet, M., Rodriguez, I. 2007. Gene cluster lock after pheromone receptor gene choice. *EMBO* **26(14)**:3423-30.
- Rothman A, Feinstein P, Hirota J, Mombaerts P. 2005. The promoter of the mouse odorant receptor gene M71. *Mol Cell Neurosci.* **28(3)**:535-46.
- Serizawa, S., T. Ishii, H. Nakatani, A. Tsuboi, F. Nagawa, M. Asano, K. Sudo, J. Sakagami, H. Sakano, T. Ijiri, Y. Matsuda, M. Suzuki, T. Yamamori, Y. Iwakura, and H. Sakano. 2000. Mutually exclusive expression of odorant receptor transgenes. *Nature Neurosc.* **3**: 687-693.
- Serizawa, S., K. Miyamichi, H. Nakatani, M. Suzuki, M. Saito, S. Yoshihara, and H. Sakano. 2003. Negative feedback regulation ensures the one receptor-one olfactory neuron rule in the mouse. *Science* **302**: 2088-2094.
- Serizawa, S., K. Miyamichi, and H. Sakano. 2004. One neuron-one receptor rule in the mouse olfactory system. *Trends in Genetics* **20**: 648-653.
- Shykind, B.M., S.C. Rohani, S. O'Donnell, A. Nemes, M. Mendelsohn, Y. Sun, R. Axel, and G. Barnea. 2004. Gene switching and the stability of odorant receptor gene choice. *Cell* **117**: 801-815.
- Shykind, B.M. 2005. Regulation of odorant receptors: one allele at a time. *Hum Mol Genet* **14**: R33-R39.

- Thompson, W., E. Rouchka, and C. Lawrence. 2003. Gibbs recursive sampler: finding transcription factor binding sites. *Nucl. Acids Res.* **31**: 3580-3585.
- Tsai, R., Y., L., Reed, R., R., 1997. Cloning and functional characterization of ROAZ, a zinc finger protein that interacts with O/E-1 to regulate gene expression: implications for olfactory neuronal development. *J Neurosci* **17**: 4159-4169.
- Vassalli, A., A. Rothman, P. Feinstein, M. Zapotocky, and P. Mombaerts. 2002. Minigenes impart odorant receptor-specific axon guidance in the olfactory bulb. *Neuron* **35**: 681-696.
- Vassar, R., J. Ngai, and R. Axel. 1993. Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. *Cell* **74**: 309-318.
- Voss T. S., Healer J., Marty A. J., Duffy M. F., Thompson J. K., Beeson J. G., Reeder J. C., Crabb B. S. and Cowman A. F. 2006. A *var* gene promoter controls allelic exclusion of virulence genes in *Plasmodium falciparum* malaria. *Nature* **439**: 1004-1008.
- Wang, M.M., R.Y.L. Tsai, K.A. Schrader, and R.R. Reed. 1993. Genes encoding components of the olfactory signal transduction cascade contain a DNA binding site that may direct neuronal expression. *Molecular and Cellular Biology* **13**: 5805-5813.
- Wang, S., R. Tsai, and R. Reed. 1997. The characterization of the Olf-1/EBF-like HLH transcription factor family: implications in olfactory gene regulation and neuronal development. *J. Neuroscience* **17**: 4149-4158.
- Wang, S., A. Betz, and R. Reed. 2002. Cloning of a novel Olf-1/EBF-like gene, O/E-4, by degenerate oligo-based direct selection. *Mol Cell Neurosci* **20**: 404-414.
- Wang, S.S., J.W. Lewcock, P. Feinstein, P. Mombaerts, and R.R. Reed. 2003. Genetic disruptions of O/E2 and O/E3 genes reveal involvement in olfactory receptor neuron projection. *Development* **131**: 1377-1388.
- Yang H, Shi P, Zhang YP, Zhang J. 2005. Composition and evolution of the V2r vomeronasal receptor gene repertoire in mice and rats. *Genomics*. **86(3)**:306-15.
- Young, J.M., B.M. Shykind, R.P. Lane, L. Tonnes-Priddy, E.M. Ross, M. Walker, E.M. Williams, and B.J. Trask. 2003. Odorant receptor expressed sequence tags demonstrate olfactory expression of over 400 genes, extensive alternate splicing and unequal expression levels. *Genome Biology* **4**: 4:R71.
- Young, J.M., Friedman, C., Williams, E.M., Ross, J.A., Tonnes-Priddy, L., and Trask, B.J. 2002. Different evolutionary processes shaped the mouse and human olfactory receptor gene families. *Hum. Mol. Genet.* **11**: 535-546.



- Young, J.M., Kambere M., Trask B.J., Lane R.P. 2005. Divergent V1R repertoires in five species: amplification in rodents, decimation in primates, and a surprisingly small repertoire in dogs. *Genome Res* **15(2)**:231-40.
- Zhang, X. and Firestein, S. 2002. The olfactory receptor gene superfamily of the mouse. *Nat. Neurosci.* **5**: 124–133.
- Zhang X, Rodriguez I., Mombaerts P., Firestein S. 2004. Odorant and vomeronasal receptor genes in two mouse genome assemblies. *Genomics* **83(5)**:802-11.

## 8) LISTA DE ANEXOS

### Anexo - I) Súmula Curricular

### Anexo - II) Publicação resultante deste trabalho

**Michaloski, J.S.;** Galante,A.F.; Malnic,B. Identification of potential regulatory motifs in odorant receptor genes by analysis of promoter sequences. *Genome Res.*, 2006. **16(9):** 1091-8.

**ANEXO- I**  
**Súmula Curricular**

## SÚMULA CURRICULAR

### DADOS PESSOAIS:

Nome: Jussara Michaloski Souza.

Local e data de nascimento: Xanxerê – SC, 20 de abril de 1976.

### EDUCAÇÃO:

Fevereiro/2003 – atual

Doutorado em Ciências Biológicas

Orientadora: Dra. Bettina Malnic.

Projeto de Pesquisa: Caracterização de Genes Expressos no Epitélio Olfatório.

Instituição: Instituto de Química, Universidade de São Paulo (USP)

Julho/2001 - Dezembro/2002.

Estágio de Iniciação Científica

Projeto de Pesquisa: Caracterização de Receptores Olfatórios Humanos.

Instituição: Instituto de Química, Universidade de São Paulo (USP)

Março/1998 - dezembro/2002.

Graduação em Farmácia-Bioquímica, especialidade Análises Clínicas e Toxicológicas.

Instituição: Faculdade de Ciências Farmacêuticas (FCF), Universidade de São Paulo (USP)

Julho/2002 - novembro/2002

Estágio no Laboratório Clínico do Instituto do Coração (InCor), Faculdade de Medicina, Universidade de São Paulo (USP)

### BOLSAS RECEBIDAS:

#### Doutorado Direto

Bolsa concedida pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) de novembro/2001 a dezembro/2002.

#### Iniciação Científica

Bolsa concedida pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) de novembro/2001 a dezembro/2002.

### PUBLICAÇÕES:

**Michaloski, J.S.;** Galante, A.F.; Malnic, B. Identification of potential regulatory motifs in odorant receptor genes by analysis of promoter sequences. *Genome Res.*, 2006. **16(9):** 1091-8.

Kerr, D.S.; Von Dannecker, L.E.; Davalos, M.; **Michaloski, J.S.;** Malnic, B. Ric-8B interacts with Gaolf and Gg13 and co-localizes with Gaolf, Gb1 and Gg13 in the cilia of sensory neurons. *Mol. Cell. Neurosc.*, 2008. **38:** 341-348.

**Anexo – II**  
**Publicação resultante deste trabalho**

**Identification of potential regulatory motifs in odorant receptor genes by analysis of promoter sequences.**

# GENOME RESEARCH

## Identification of potential regulatory motifs in odorant receptor genes by analysis of promoter sequences

Jussara S. Michaloski, Pedro A.F. Galante and Bettina Malnic

*Genome Res.* 2006 16: 1091-1098; originally published online Aug 10, 2006;  
Access the most recent version at doi:[10.1101/gr.5185406](https://doi.org/10.1101/gr.5185406)

---

**Supplementary data**

*"Supplemental Research Data"*

<http://genome.cshlp.org/cgi/content/full/gr.5185406/DC1>

**References**

This article cites 49 articles, 20 of which can be accessed free at:  
<http://genome.cshlp.org/cgi/content/full/16/9/1091#References>

Article cited in:

<http://genome.cshlp.org/cgi/content/full/16/9/1091#otherarticles>

**Email alerting service**

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

---

---

To subscribe to *Genome Research* go to:  
<http://genome.cshlp.org/subscriptions/>

---



# Identification of potential regulatory motifs in odorant receptor genes by analysis of promoter sequences

Jussara S. Michalowski,<sup>1</sup> Pedro A.F. Galante,<sup>1,2</sup> and Bettina Malnic<sup>1,3</sup>

<sup>1</sup>Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, C.P.26077 CEP 05513-970, São Paulo, Brazil;

<sup>2</sup>Ludwig Institute for Cancer Research, São Paulo, 01509-010, SP, Brazil

Mouse odorant receptors (ORs) are encoded by >1000 genes dispersed throughout the genome. Each olfactory neuron expresses one single OR gene, while the rest of the genes remain silent. The mechanisms underlying OR gene expression are poorly understood. Here, we investigated if OR genes share common *cis*-regulatory sequences in their promoter regions. We carried out a comprehensive analysis in which the upstream regions of a large number of OR genes were compared. First, using RLM-RACE, we generated cDNAs containing the complete 5'-untranslated regions (5'-UTRs) for a total number of 198 mouse OR genes. Then, we aligned these cDNA sequences to the mouse genome so that the 5' structure and transcription start sites (TSSs) of the OR genes could be precisely determined. Sequences upstream of the TSSs were retrieved and browsed for common elements. We found DNA sequence motifs that are overrepresented in the promoter regions of the OR genes. Most motifs resemble O/E-like sites and are preferentially localized within 200 bp upstream of the TSSs. Finally, we show that these motifs specifically interact with proteins extracted from nuclei prepared from the olfactory epithelium, but not from brain or liver. Our results show that the OR genes share common promoter elements. The present strategy should provide information on the role played by *cis*-regulatory sequences in OR gene regulation.

[Supplemental material is available online at [www.genome.org](http://www.genome.org). The sequence data from this study have been submitted to GenBank under accession nos. DR065530–DR065963.]

Mammalian olfactory sensory neurons select, from >1000 possible choices, one single olfactory receptor (OR) allele to express (Ressler et al. 1993; Vassar et al. 1993; Chess et al. 1994; Malnic et al. 1999; but see Mombaerts 2004). The receptor type that is chosen will not only determine the range of odorants to which this neuron will respond, but also axonal targeting to specific glomeruli in the olfactory bulb (Mombaerts et al. 1996; Wang et al. 1998). OR gene choice is therefore fundamental for the functional organization of the olfactory system. How this choice is accomplished is, however, still unclear.

Little is known about the role of *cis*-regulatory sequences in the regulation of OR gene expression. In studies using transgenic mice, different sizes of genomic DNA segments containing OR genes were tested for their ability to drive an OR expression similar to that of the endogenous gene. It was demonstrated that short pieces of DNA located upstream of the coding region, ranging from 460 bp to 6.7 kb, are sufficient for expression of the ORs M4, M71, and MOR23 (Qasba and Reed 1998; Vassalli et al. 2002). However, large segments of ~200 kb are required to obtain expression of MOR28 (Serizawa et al. 2000). Sequence comparison of the mouse and human genome revealed a 2-kb conserved sequence located ~75 kb upstream of the *MOR28* gene cluster. This region, denominated the H region, was proposed to work as a *cis*-acting locus control region (LCR) that would activate the expression of one single OR gene member from within the *MOR28* gene cluster (Serizawa et al. 2003). Altogether, these re-

sults indicate that *cis*-regulatory sequences may play important roles in OR gene choice.

To date, different combinations of transcription factor binding sites (TFBSs) have been identified in promoters of OR genes (Hoppe et al. 2000, 2003; Sosinsky et al. 2000; Lane et al. 2001; Vassalli et al. 2002); however, there is no evidence yet that these sites are directly involved in OR gene choice. A strong consensus sequence, the *Olf-1* site (O/E-like site), was identified in the promoters of several olfactory specific genes, such as *Gnal* (formerly known as *Golf*), adenylyl cyclase III (*AcIII*), olfactory cyclic nucleotide gated channel (*Cnga2*), and olfactory marker protein (*Omp*) (Kudrycki et al. 1993; Wang et al. 1993), and was also found in the promoter regions from some OR genes (Glusman et al. 2000b; Sosinsky et al. 2000; Vassalli et al. 2002; Hoppe et al. 2003) but not from other OR genes (Hoppe et al. 2000; Lane et al. 2001).

There are >1000 OR genes dispersed throughout the genome (Young et al. 2002; Zhang and Firestein 2002; Godfrey et al. 2004). A genomic approach to identify potential regulatory *cis*-acting sequences is to search for DNA sequence elements that are conserved in a large number of OR gene promoters. Promoter sequences, which are usually located proximal to and upstream of the transcription start site (TSS), can be retrieved from the available mouse genome sequence. This can be done by aligning full-length OR mRNA sequences with their counterpart genomic sequences. It is important that full-length mRNA sequences are used, because the transcriptional start sites can be located far away from the translational start sites. It has been demonstrated, for example, that ORs in the mouse P2 OR cluster have 5'-untranslated regions (5'-UTRs) that range from 1.7 to 9 kb (Lane et al. 2001). Besides, it is known that many OR genes have 5'

### <sup>3</sup>Corresponding author.

E-mail [bmalnic@iq.usp.br](mailto:bmalnic@iq.usp.br); fax 55-011-38155579.

Article is published online before print. Article and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.5185406>.

non-coding exons, and some undergo 5' alternative splicing (Sosinsky et al. 2000; Lane et al. 2001).

To date, full-length 5' cDNA sequences are available only for ~30 different mouse OR genes, distributed over five different chromosomes (Bulger et al. 2000; Hoppe et al. 2000, 2003; Sosinsky et al. 2000; Lane et al. 2001; Vassalli et al. 2002). Another study produced cDNA sequences representing >400 OR genes, many of them containing 5'-UTR sequences. However, they may not be full-length sequences, since they originated from a regular cDNA library, and not from a full-length cDNA library (Young et al. 2003).

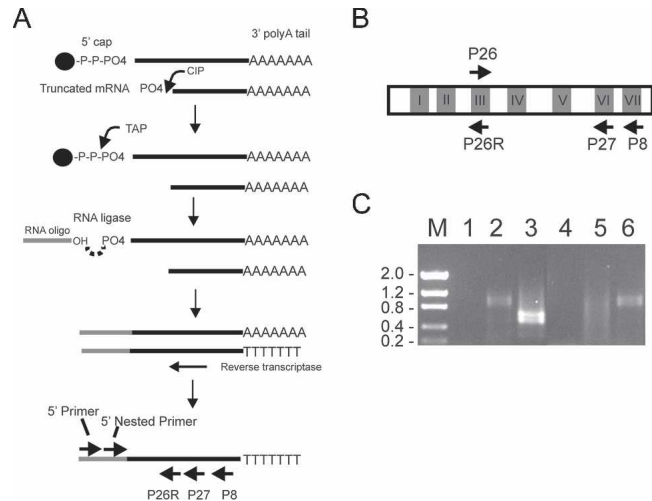
Here we have used RLM-RACE with degenerate primers to produce cDNAs containing the complete 5'-UTRs for a total number of 198 mouse OR genes. Then, we aligned these cDNA sequences to the mouse genome so that sequences corresponding to one same OR gene are organized in one same cluster, making a total of 198 clusters. For each OR gene cluster, we determined the 5'-gene structure (exon and intron distribution) and TSS. The 198 promoter sequences were retrieved and browsed for common elements. Motifs that are common to a large percentage of the OR genes were found. Most motifs resemble O/E- and homeodomain-like sites. The O/E-like sites are localized within 200 bp upstream of the TSS and specifically interact with proteins extracted from nuclei prepared from the olfactory epithelium, but not with proteins extracted from brain or liver.

## Results

### Amplification of OR cDNA 5'-ends

To obtain 5'-end sequences from OR cDNAs, we performed RNA ligase-mediated rapid amplification of 5' cDNA ends (RLM-RACE) using total RNA purified from mouse olfactory epithelium. This method has the advantage that only full-length transcripts (authentic 5'-end capped mRNAs) are amplified (Fig. 1A). In order to obtain 5' cDNA ends for a large number of OR genes, we used degenerate primers matching to conserved regions in ORs (Fig. 1B) because these primers can amplify the majority of the members of the OR family (Buck and Axel 1991; Malnic et al. 1999). The initial PCR reaction was performed using the 5' GeneRacer primer together with the reverse P27 degenerate primer. In order to eliminate PCR artifacts, a secondary (nested) PCR was performed using the first PCR product as a template and the 5' GeneRacer nested primer together with the P26R degenerate primer (Fig. 1A). The primary PCR product contained a heterogeneous mixture of cDNA fragments ranging from 0.8 to 2 kb in length (Fig. 1C, lane 2). This result was expected, since different OR cDNAs must have different 5'-UTR sizes. The secondary PCR reaction produced a similar range of cDNA fragments, except that their sizes were ~350 bp smaller (Fig. 1C, lane 3). This was also expected, since the P26R primer matches to a region in the OR coding sequence that is located ~350 bp upstream to the region matched by P27 (Fig. 1B). Since the region between the first AUG codon and TM-III in OR genes is ~380 bp long, only the nested PCR products  $\geq$ 380 bp long were gel-purified, cloned, and sequenced.

The same procedure was also performed using primers 5' GeneRacer and P8 degenerate primer for the primary PCR reaction (Fig. 1C, lane 5), and primers 5' GeneRacer nested and P27 for the secondary PCR reaction (Fig. 1C, lane 6). In this case, since the region between the first AUG codon and TM-VI in ORs



**Figure 1.** The strategy used for the generation of OR complete 5'-end cDNA sequences. (A) Truncated mRNAs were dephosphorylated using calf intestine phosphatase (CIP) so that they cannot participate in subsequent ligation reactions. The RNA was then treated with tobacco acid pyrophosphatase (TAP) to remove the 5'-cap structure (represented by black balls) from intact full-length mRNA, and the GeneRacer RNA oligonucleotide was ligated to the decapped mRNA. Reverse transcription was performed using oligo(dT) primers. To obtain 5'-ends, PCR was done using the GeneRacer 5'-primer and a degenerate primer directed toward conserved OR regions (P8 or P27). Nested PCR was then done using the GeneRacer 5' nested primer and another OR degenerate reverse primer (P27 or P26R). (B) Schematic representation of an OR coding region showing the seven transmembrane regions (I-VII) and the regions matched by the degenerate primers used in this study. (C) 1.5% agarose gel showing the PCR products obtained using different combinations of primers. (Lane 1) Negative control for reaction in lane 2 (no DNA added); (lane 2) GeneRacer 5'-primer and P27; (lane 3) GeneRacer 5' nested primer and P26R; (lane 4) negative control for reaction in lane 5 (no DNA added); (lane 5) GeneRacer 5'-primer and P8; (lane 6) GeneRacer 5' nested primer and P27. (M) Molecular weights are given in kilobases.

is ~750 bp long, only the nested PCR products  $\geq$ 750 bp were analyzed.

### OR gene clusters

We sequenced 1012 clones from their 5'-ends, and 80% of them correspond to ORs, indicating that our strategy preferentially amplifies OR sequences. In addition, 96% of the OR cDNAs contain 5' sequences upstream of the predicted initial AUG codon, indicating that full-length mRNAs were amplified. Sequence analysis showed that 5'-RACE products were obtained for a total number of 198 different OR genes, corresponding to ~17% of the complete mouse OR gene repertoire (Table 1). Only nine of the OR genes are pseudogenes.

OR sequences are classified into two phylogenetic classes, referred to as Class I (fish-like) and Class II (terrestrial-specific) ORs (Ngai et al. 1993; Freitag et al. 1995; Glusman et al. 2000a). The Class I ORs constitute ~12% of the mouse OR repertoire (Zhang and Firestein 2002). Nine of the 198 ORs (4.5%) are Class I ORs, indicating that the method we used amplifies members of the two OR Classes, although it may favor amplification of Class II ORs.

The OR genes for which cDNA sequences were obtained are distributed among all of the mouse chromosomes previously shown to contain OR genes (Godfrey et al. 2004), except for chromosome 3 (Table 1). The majority of the sequences corre-



**Table 1.** Chromosomal distribution of OR genes

Chromosome	No. of genes <sup>a</sup>	
	Annotated OR genes	OR genes in this study
1	23 (7)	6
2	275 (81)	44 (5)
3	2	—
4	19 (7)	4
5	—	—
6	22 (9)	5
7	198 (51)	32
8	4	2
9	118 (37)	20 (1)
10	47 (13)	9
11	35 (20)	21 (2)
12	—	—
13	12 (2)	4
14	32 (4)	6
15	6 (3)	3 (1)
16	29 (8)	18
17	36 (17)	5
18	—	—
19	52 (19)	9
X	2	1
Y	—	—
Total	1190	198

<sup>a</sup>The number of annotated mouse OR genes from Godfrey et al. (2004) and of mouse OR genes for which cDNA sequences were obtained in the present study. Numbers of pseudogenes are indicated in parentheses.

spond to genes in chromosomes 2 and 7, which had been previously shown to contain the higher numbers of OR genes (Young et al. 2002; Zhang and Firestein 2002; Godfrey et al. 2004). In addition, the 198 OR genes can be subdivided into 102 (41%) out of the total 248 mouse OR subfamilies (where all members of a subfamily are  $\geq 60\%$  identical to all other members in amino acid sequence, as described by Godfrey et al. 2004; see Supplemental Table 1). We therefore believe that our sequences are representative of a random sample of the mouse OR genes. However, it is important to note that the present method may favor the amplification of OR cDNAs that have short 3'-UTRs or that are highly expressed in the olfactory epithelium.

We next used BLAST (Altschul et al. 1990) and Sim4 (Florea et al. 1998) to align all of the cDNA sequences with the mouse genome sequence. The previously annotated mouse OR genes (Young et al. 2002; Zhang and Firestein 2002) were also included in the alignment, to help with the localization of the cDNAs 5'-UTR regions. Each cDNA sequence aligned to one single genomic region, and the cDNAs that aligned to the same genomic region as one of the annotated ORs were considered to correspond to that particular OR gene. In this way, we obtained a total of 198 OR clusters, where each cluster corresponds to one different OR gene. Each one of the clusters contains at least one cDNA sequence, the largest cluster contains 71 sequences, and 54% of the clusters contain more than two sequences (Supplemental Table 1).

The structural organization of each one of the 198 clusters can be visualized using the Olfactory Receptor cDNA Clusters Viewer (<http://gbrowser.compbio.ludwig.org.br/or/>) by entering the corresponding cluster numbers shown in Supplemental Table 1. In summary, of the 198 OR gene clusters, only two do not have introns in their 5'-ends (Supplemental Tables 1 and 2), and 39 showed alternative splicing (in this case, only clusters containing more than two cDNA sequences were analyzed).

### 5'-structure of the OR genes

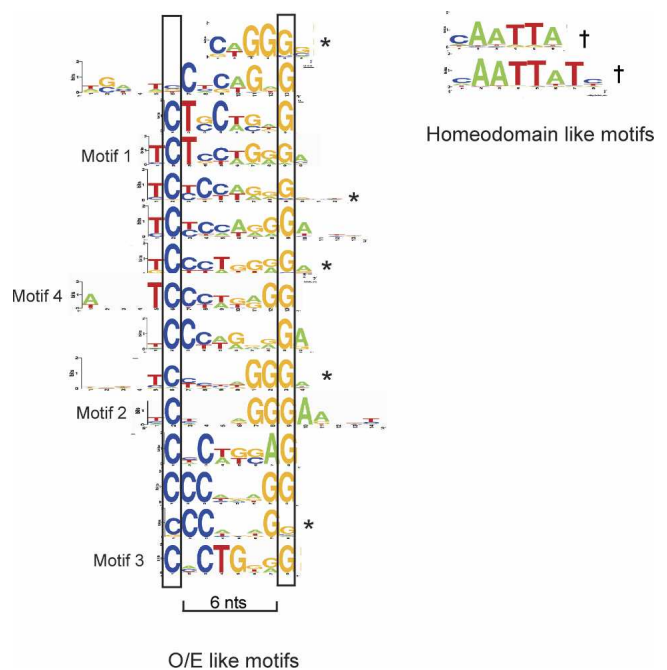
On average, the 5'-UTR is 189 bp long, ranging from 32 to 659 bp. The 5'-UTR exons range from 9 to 403 bp (Supplemental Table 2). Most of the OR genes (72%) have only one 5'-UTR exon, 23% have two, and 3% have three (Supplemental Tables 1 and 2). The 5'-UTR introns show a wide variation in size, ranging from 91 bp to 22.5 kb. A large fraction of the OR genes (46%) have introns with sizes between 2 kb and 4 kb, and 34% of the genes have introns  $>4$  kb.

It is predicted that in most eukaryotic mRNAs translation initiates at the first AUG starting from the 5'-cap (Kozak 1999). Therefore, one would expect not to find AUG codons or upstream ORFs (uORFs) in the 5'-UTRs. However, we found that 108 of the 198 OR genes have uORFs at least 10 codons long. It has been suggested that uORFs could be involved in down-regulation of protein translational efficiency (Pesole et al. 2001). Further analysis should clarify whether these uORFs play a role in OR gene regulation.

We also found that a significant percentage of the OR genes (18%) have an in-frame upstream AUG, indicating that these genes code for OR proteins with a longer N-terminal region than the one originally predicted from their genomic sequences.

### OR promoter regions

The generation of 5' full-length cDNAs allowed us to precisely determine the TSSs for the 198 OR genes. On average, the TSSs are located 4.3 kb upstream of the initial AUG codon, the furthest TSSs being located 22.5 kb away and the closest only 18 bp away. Sequences (600 bp) upstream of each TSS were excised from the mouse genomic sequence and analyzed. We first screened the



**Figure 2.** Motif patterns found in OR promoter regions. Logo representations were created using the software from <http://weblogo.berkeley.edu/logo.cgi>. Motifs were identified by using Gibbs Recursive Sampler, Consensus (\*) or Weeder (†). Motifs that resemble O/E-like sites show a conserved C<sub>n</sub>G sequence. Motifs 1–4 were further analyzed as described in Figures 3–5.

<i>Omp</i>	GTCCCAAGGAG		
<i>Omp</i>	CTCCCAGGGAG		
<i>AcIII</i>	TTCCCTTGAGGA		
<i>GnaI</i>	TTCCCCAAGAA		
	**** *		
M1_720	TTCTCTGGGAG	M3_211	CCCCTGGGGCT
M1_1000	TTCTCTAGGAG	M3_1377	TCCACTGTGGTT
M1_56	CTCTCTGGGAG	M3_1352	TTCCCTGTGGGA
M1_1273	ACCTCTGGGAG	M3_1308	ACCCTGAAGCA
M1_339	ATCTCTGGGAT	M3_171	TGCACTGTAGCA
M1_1377	ATCTCCAGGGAC	M3_32	CACACTGTAGAC
M1_1045	ACCTCCAGGAAT	M3_1434	AACACTGGGGAA
M1_1356	CTCTTCAGAGAA	M3_90	CCCCTGGAGTA
M1_123	GTCTTCAGAGAA	M3_1415	CTCACTGGAGGA
M1_191	CCCTCCAGGGAA	M3_413	ATCCCTGGAGTC
	** *		* ** *
M2_165	TTCTCCAGGGAA	M4_1339	ATCCTCCCTGAG
M2_457	TCCCACAGGGAA	M4_389	ATCCCTGAGGAA
M2_1501	TCCTAGAGGGAA	M4_392	ATCCCTGAGGAG
M2_1270	TACTAGAGGGAA	M4_214	TTCCAGAGGAT
M2_181	TTCTGCAGGGAA	M4_1	ATCCAGAGGAG
M2_1352	TTCCATGGGGAA	M4_1000	TTCTTGGGGAC
M2_27	ATCCATAGGGAA	M4_24	TTCTATGGGGT
M2_741	TTCACTGGGGAA	M4_1301	GTCCATGGGTT
M2_181	AACACTGGGGAA	M4_1261	GTCCTTGGGTA
M2_56	CTCTTGGGGAG	M4_702	ATCCTTAGGTA
	* ****		*** *

**Figure 3.** Nucleotide sequence alignment of the conserved sequences in O/E-like motifs M1–M4 and comparison with *Olf-1* binding sites. Ten motif sequences were randomly selected for each one of the motifs (M1–M4) and manually aligned to the *Olf-1* sites (from *Omp*, *AcIII*, *Cnga2*, and *GnaI*) (Wang et al. 1993). Identical nucleotide positions in each group of sequences are indicated by asterisks. The names of the OR genes from which the motifs were retrieved are indicated (e.g., M1\_720: motif 1 from olfr720).

sequences with RepeatMasker (<http://www.repeatmasker.org/>) and found that only 7% of the total sequences contain repeats or low-complexity regions. Typical TATA-boxes were found in only a small number of the OR gene promoters (35%), consistent with previous reports (Hoppe et al. 2000; Sosinsky et al. 2000; Lane et al. 2001).

Because most TFBS are usually short, they can occur very frequently in the sequences, making it difficult to identify significant sites. In order to reduce the false-positive predictions, we decided to search for motifs that are common to a large fraction of the promoter sequences and thereby identify elements that are more likely to be functionally important. To do this, we used the Gibbs recursive sampler (Thompson et al. 2003), Consensus (Hertz and Stormo 1999), and Weeder (Pavesi et al. 2004) tools, which were designed to locate common elements in collections of unaligned DNA sequences. We found several motifs that are shared by the OR gene promoter sequences (Fig. 2). A closer inspection of the motifs revealed that although they are diverse, the majority of them resemble *Olf-1* (O/E) like sites (Figs. 2 and 3; Wang et al. 1993). The O/E-like motifs can be divided into four groups, denominated M1–M4. Motifs in each one of these four groups show different conserved nucleotide sequences (Fig. 3). We also found motifs that resemble the homeodomain sites, previously shown to be located in proximity to O/E-like binding sites in OR promoter genes (Fig. 2; Vassalli et al. 2002; Rothman et al. 2005).

The spatial distribution of O/E-like motifs M1–M4 in the promoter sequences is shown in Figure 4A. All of the motifs are concentrated near the TSSs (between +1 and –200 bp). Differently, the homeodomain-like sites show a broader distribution over the entire extent of the 600-bp sequence (Fig. 4B).

The O/E-like motifs and the homeodomain-like motifs were found, respectively in 87% and 95% of the OR promoter sequences (Table 2).

## DNA motifs form complexes with olfactory nuclear proteins

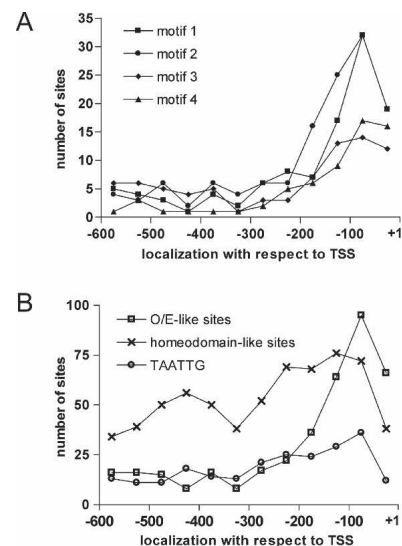
To evaluate the biological significance of the motifs, we performed gel shift assays. As shown in Figure 5, motifs M1–M4 formed DNA–protein complexes in the presence of the olfactory nuclear extracts, but not in the presence of brain and liver nuclear extracts. The different complexes show similar electrophoretic mobilities, but the intensities of the shifted bands vary. Formation of these complexes was inhibited by pre-incubation of the binding reaction with a 100-fold excess of the corresponding specific unlabeled oligonucleotides (Fig. 5A). We next evaluated the specificity of the DNA–protein complexes. Mutated motifs M1–M4, where the conserved nucleotides were changed into different ones, were unable to form stable DNA–protein complexes with olfactory nuclei proteins (Supplemental material 3). Altogether, these results show that proteins present in nuclei of olfactory epithelium cells, but not in liver and brain extracts, specifically bind to motifs M1–M4.

## Discussion

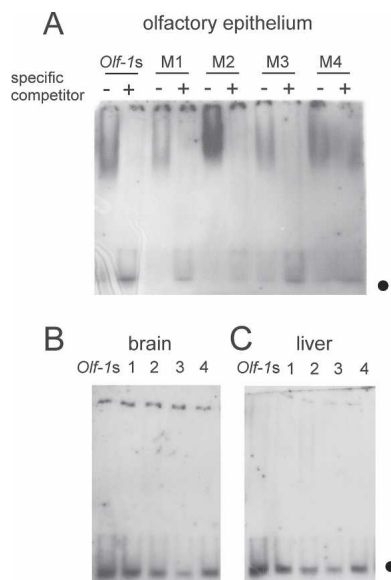
The role of *cis*-regulatory sequences in the regulation of OR gene expression is not understood. In the present study, we carried out a comprehensive analysis in which the promoter regions of 198 OR genes were compared.

### Common OR gene promoter motifs

By comparing a variable set of OR genes, we aimed to identify common promoter elements that may be involved in the general mechanisms of OR gene regulation. Strikingly, 87% of the OR gene promoter regions contain O/E-like sites, and 95% contain homeodomain-like sites. No other types of motifs were found to be overrepresented in these sequences, although we cannot ex-



**Figure 4.** Localization of the motifs with respect to the transcription start sites. (A) The number of sites for motifs M1–M4 found across the OR gene promoter regions. The positions of the TSS (+1) and 600 bp upstream of the TSS (–100 to –600) are indicated. (B) The number of total O/E-like sites (sum of M1–M4 motifs) and homeodomain-like sites (based on the HAATTA consensus) found across the OR promoter regions. The distribution of TAATTG homeodomain sites, previously shown to be present in many OR promoter regions (Vassalli et al. 2002) and to be involved in the regulation of M71 expression (Rothman et al. 2005), is also shown.



**Figure 5.** Binding of nuclear proteins to the DNA motifs. Labeled double-stranded oligonucleotides corresponding to motifs M1–M4 were incubated with nuclear extracts from (A) olfactory epithelium, (B) brain, or (C) liver, as indicated. The *Olf-1* binding site (*Olf-1s*) was used as a positive control. DNA–olfactory epithelium nuclear protein complexes are observed for motifs M1–M4. The DNA–protein interactions are completed by the addition of a 100-fold molar excess of the corresponding unlabeled specific oligonucleotide (A). The positions of the free probes are indicated by filled circles.

clude the possibility that the algorithms we used may have precluded us from finding additional sites.

The roles of the O/E-like and homeodomain-like proteins in OR gene expression are still unclear. It has been demonstrated that expression of the M71 OR gene can be driven by a short region upstream of the TSS (a minimal promoter region) containing an O/E-like site and a homeodomain-like site (Vassalli et al. 2002). Mutational studies using transgenic mice indicate that both sites are required for normal OR gene expression (Rothman et al. 2005). Yet, disruption of *olf-1*-like genes does not alter OR gene expression (Lin and Grosschedl 1995; Wang et al. 2003), possibly because of the functional redundancy of the multiple O/E family members expressed in the olfactory epithelium (O/E1, O/E2, O/E3, and O/E4) (Wang et al. 1997, 2002). Nevertheless, it was demonstrated that O/E2 and O/E3 mutant mice show defects in the projection of olfactory neurons to the olfactory bulb, indicating that the O/E genes' function may not be completely redundant (Wang et al. 2003).

The LIM-homeodomain protein *Lhx2* was shown to bind to the homeodomain site in the M71 OR promoter region (Hirota and Mombaerts 2004). *Lhx2* knockout mice do not express ORs, but since they also lack mature olfactory neurons, there is no evidence so far that this homeodomain protein is directly involved in OR gene expression (Hirota and Mombaerts 2004; Kolterud et al. 2004). It is also possible that other homeodomain proteins than *Lhx2* bind to the homeodomain sites in the OR gene promoters.

Although the vast majority of OR gene promoters have O/E-like sites, the structure of these sites is variable (see Figs. 2 and 3). It has been previously shown that the different O/E proteins possess similar DNA-binding properties (Wang et al. 1997). Our results indicate that the M1–M4 motifs interact differently with

proteins from olfactory epithelium nuclei (Fig. 5). Several possibilities could explain the different DNA–protein complex affinities. The motifs could bind to different O/E proteins, or to alternatively spliced versions of these proteins (Wang et al. 1997, 2002). It is also possible that the same O/E protein types could bind to the M1–M4 motifs, but with different affinities.

Our findings suggest that different OR gene promoters are bound by different combinations or amounts of O/E-like proteins. The consequences of these differential interactions for OR gene regulation are unknown. It is known that different ORs are expressed in different levels (Young et al. 2003). One interesting possibility is that the types of O/E-like sites in an OR gene promoter region may determine its probability of being transcribed. The identification of the proteins that interact with each one of the motifs and the analysis of the expression patterns of OR genes that have different motifs should clarify the role of these O/E-like sites in OR gene regulation.

### Promoter DNA elements and OR gene regulation

Different models for OR gene regulation have been considered to date (for review, see Sosinsky et al. 2000; Mombaerts 2004; Serizawa et al. 2005; Shykind 2005). It has been recently demonstrated that the monoallelic expression of an OR gene is regulated by a negative feedback mechanism that requires a functional OR protein (Serizawa et al. 2003; Lewcock and Reed 2004). In addition, it was shown that immature olfactory neurons that express a given odorant receptor can switch receptor expression at a low frequency, while neurons expressing a mutant (nonfunctional) OR can switch expression with a greater probability (Shykind et al. 2004). Based on these results, a new model has been proposed (Serizawa et al. 2004; Shykind 2005). In this model, after an OR gene is stochastically selected for expression by a limiting factor, its corresponding OR protein product mediates a feedback signal that results in the maintenance of the receptor choice.

Here we show that a collection of random OR genes will have the same types of *cis*-regulatory elements, suggesting that these common promoter elements are likely to play an important role in OR gene expression. It is possible that enhancers or LCRs interact with elements in one OR gene promoter to select that specific OR for expression. Interestingly, it was shown that the H region, which works as an LCR and is located 75 kb upstream of the *MOR28* gene cluster (Serizawa et al. 2003), also contains at least one set of homeodomain- and O/E-like sites (Hirota and Mombaerts 2004). Alternatively, *cis*-elements and protein factors that bind to these elements could bring one given OR gene promoter to a single expression site body in the nucleus (Borst 2002; Voss et al. 2006).

**Table 2.** Distribution of the sequence motifs among the OR genes

Motif	No. of OR genes (%) <sup>a</sup>	Total no. of sites <sup>b</sup>
M1	85 (42%)	110
M2	99 (50%)	129
M3	71 (36%)	79
M4	51 (26%)	64
O/E like sites	173 (87%)	382
Homeodomain sites	188 (95%)	1029

<sup>a</sup>Number of OR gene promoters (% of 198 promoters) containing motifs M1–M4, total O/E-like sites (sum of M1–M4 motifs) and homeodomain-like sites (based on the HAATTA consensus sequence).

<sup>b</sup>Total number of motif sites found in the 198 OR gene promoters.

However, it is important to note that other olfactory genes that are expressed in all mature olfactory neurons, such as *Omp*, *Gnal*, and *AcIII*, also have O/E-like sites (Fig. 3). Therefore, the mere presence of O/E-like sites in the promoter regions does not explain the mosaic pattern of OR expression in the olfactory epithelium.

In conclusion, our results indicate that intraspecies comparisons of promoter sequences are likely to be a useful strategy for identifying common regulatory motifs that may be involved in regulation of OR gene expression. A similar strategy can also be applied to other multigene families whose members are coordinately regulated, such as the pheromone receptor families (Dulac and Torello 2003).

## Methods

### 5' RLM-RACE

Total RNA was purified from C57BL/6J mice (6–8 wk old) olfactory epithelium using TRIzol reagent (Invitrogen), following the manufacturer's instructions. RLM-RACE was performed using the GeneRacer kit (Invitrogen) and 4 µg of total RNA. Twenty-five-microliter PCR reactions containing 1 µL of RLM-RACE cDNA, 0.2 mM dNTP, 1.5 mM MgCl<sub>2</sub>, 0.5 µM each forward and reverse primer (or 2 µM degenerate primers), 1.25 U of Platinum Taq DNA polymerase (Invitrogen) were heated to 95°C for 2 min, followed by 40 thermal cycles of 95°C for 1 min, 50°C for 3 min, 72°C for 2 min, and a final incubation at 72°C for 10 min.

Twenty-five-microliter nested PCR reactions were done using 1 µL of a 200-fold dilution of the primary PCR product and 25 cycles as above.

### Degenerate oligonucleotide primers

The following degenerate OR primers were used in this study:

Reverse primers

P8 (TM-VII): (GA)TTIC(TG)IA(AG)I(GC)(TA)(GA)TA  
IAT(AG)AAIGG(GA)TT

P27 (TM-VI): ACIACIGAIAG(GA)TGIGAI(GC)C(GA)CAIGT

P26R (TM-III): CAIATIGCIAC(AG)TAICG(GA)TCAIGTAIGC

Forward primer

P26 (TM-III): GCITA(CT)GA(CT)CGITA(CT)GTIGCIATITG

### Cloning and sequencing

The RACE nested PCR products were gel-purified and cloned into the pCRII vector (Invitrogen). Colonies containing OR cDNAs were selected by colony PCR using the OR degenerate primers P26 and P27, and their orientation was determined also by colony PCR using the pair of primers T7/P27 or SP6/P27. Plasmid DNA was prepared from positive colonies using Filter Plate for high-throughput separations (Multiscreen Millipore). DNA was sequenced with the ABI PRISM Big Dye Terminator V3.1 Cycle sequencing kit using T7 or SP6 primers on an ABI PRISM 3100 Genetic Analyzer (Hitachi). Four percent of the OR sequences were truncated OR RNAs, probably because the CIP reaction during RLM-RACE was not 100% efficient.

### Genomic alignment of cDNA sequences

We aligned the cDNA sequences against the mouse genome using BLAST. Only the alignments with percent identities >93% were considered. The position of each alignment was calculated, and the flanking 50-kb genomic sequences were extracted from the corresponding genomic contigs. Each sequence was realigned with its corresponding extracted genomic sequence using the

Sim4 program (Florea et al. 1998). Only the Sim4 alignments showing average percent identity >93%, entire sequence alignment >50%, and with the best score (based on the nucleotide identity over the entire alignment) were selected. A MySQL database was loaded with the alignment information.

### Clustering of cDNA sequences

The cDNA sequences were clustered based on their genomic coordinates. Sequences that share at least one same exon/intron boundary were included in the same cluster. When no exon/intron boundaries were defined, sequences with at least 30-bp overlap in one same genomic location were included in the same cluster. The Olfactory Receptor cDNA Clusters Viewer site was generated using the Generic Genome Browser (Stein et al. 2002; <http://www.gmod.org/ggb/>).

### Promoter sequence analysis

Promoter sequences were analyzed using the Gibbs Recursive Sampler (Thompson et al. 2003). A FASTA sequence file containing the 198 promoter sequences (600 bp upstream of the TSS) (Supplemental material 1) was analyzed using the eukaryotic default values for all parameters and motif lengths 8, 8, 6, 8, 8; 12, 12, 10, 12, 12; or 14, 14, 12, 14, 14. The parameters used to identify some of the motifs are shown in the Gibbs output files (Supplemental material 2). The promoter sequences were also analyzed using Consensus (Hertz and Stormo 1999) and Weeder (Pavesi et al. 2004) (in both cases, motif widths were set to 6, 8, 10, or 12). Potential TATA-box sequences were predicted using HcTata ([http://125.itba.mi.cnr.it/~webgene/wwwHC\\_tata.html](http://125.itba.mi.cnr.it/~webgene/wwwHC_tata.html)). The location of the motifs within the promoter regions was determined using SiteSeer (<http://rocky.bms.umist.ac.uk/SiteSeer/>). Motifs were searched in both strands of the input sequences.

### Preparation of nuclear extracts

Nuclear extracts were prepared using the method described by Kudrycki et al. (1993) from olfactory epithelium dissected from 30 4–7-wk-old C57BL/6J mice. The extract was first concentrated using a Microcon centrifugal filter device (Millipore), and then the buffer was exchanged with binding buffer (10 mM Tris-HCl at pH 7.9, 1 mM EDTA, 5 mM MgCl<sub>2</sub>, 50 mM KCl, 10% glycerol, 3 mM DTT, 0.3 mM PMSF) using a Micro Bio-Spin P-6 chromatography column (Bio-Rad). Aliquots were stored at –80°C. Protein concentration was determined using the Bradford assay (Bio-Rad).

### Gel shift assay

The digoxigenin (DIG) gel shift kit (Roche Applied Science) was used for gel shift assays. Binding reactions contained 2 µg of poly[d(I-C)], 0.1 µg of poly-L-lysine, 1.2 ng of labeled oligonucleotide, and 15 µg of nuclear protein extract. After a 10-min incubation on ice and a 15-min incubation at room temperature, the mixture was added with 5 µL of loading buffer (60% 0.25× TBE buffer, 40% glycerol, 0.2% bromophenol blue) and electrophoresed in 0.5× TBE on a nondenaturing 4% polyacrylamide gel in 0.5× TBE containing a 2-cm 15% acrylamide layer at the bottom to retain the unbound probe in the gel, as described by Bell et al. (1999). The gel was pre-electrophoresed for 1 h at 80 V before the samples were applied. Competition experiments were performed by incubating the binding reaction mixtures with a 100× excess of unlabeled competitor oligonucleotide for 5 min before the addition of the labeled oligonucleotide. Blotting was performed using a Bio-Rad electroblotting system, and chemiluminescence detection of the DIG-labeled DNA–protein complexes was performed using anti-digoxigenin antibody conju-

gated to alkaline phosphatase and the CSPD substrate (Roche Applied Science).

The following pairs of complementary oligonucleotides were used as double-stranded DNA probes for the gel shift reactions (motif sequences are underlined, and the OR genes from which sequences were extracted are indicated):

M1 motif (from olfr720):

5'-TCTCAGACTTTTCTCCTGGGAGACATCTCAG-3' and  
5'-CCTGAGATGTCTCCAGGAGAAAAGTCTGAG-3';

M2 motif (olfr165):

5'-TAAGATGCTAAATTCCTGGAGAAATTGTAA-3' and  
5'-TTTACAATTCTCCAGGGAATTTAGCATCTT-3';

M3 motif (olfr211):

5'-CCTGGCATCTCCACTGGGGCTTATATTCTG-3' and  
5'-ACAGAATATAAGCCCCAGTGGGAGATGCCAG-3';

M4 motif (olfr1339):

5'-CTTCAGCTTCATCTCCCTGAGGACAGGGAG-3' and  
5'-GCTCCCTGTCTCAGGGAGGATGAAGCTGAA-3'.

## Acknowledgments

We thank Hiroaki Matsunami for critically reading the manuscript. We are grateful to Sandro José de Souza and Noboro Jo Sakabe for providing bioinformatics assistance. We also thank Erica Bandeira and Luci Navarro for technical assistance. This work was supported by research grants from FAPESP and CNPq (to B.M.) and fellowships from FAPESP (to J.S.M. and P.A.F.G.).

## References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**: 403–410.
- Bell, A., Feng, X., and Reder, A. 1999. Improved band resolution, loading reliability and reduced <sup>32</sup>P contamination in mobility shift assays by retention of unbound probe. *Biotechniques* **27**: 1122–1126.
- Borst, P. 2002. Antigenic variation and allelic exclusion. *Cell* **109**: 5–8.
- Buck, L. and Axel, R. 1991. A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell* **65**: 175–187.
- Bulger, M., Bender, M.A., van Doorninck, J.H., Wertman, B., Farrell, C.M., Felsenfeld, G., Groudine, M., and Hardison, R. 2000. Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse  $\beta$ -globin gene clusters. *Proc. Natl. Acad. Sci.* **97**: 14560–14565.
- Chess, A., Simon, I., Cedar, H., and Axel, R. 1994. Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78**: 823–834.
- Dulac, C. and Torello, A.T. 2003. Molecular detection of pheromone signals in mammals: From genes to behaviour. *Nat. Rev. Neurosci.* **4**: 551–562.
- Florea, L., Hartzell, G., Zhang, Z., Rubin, G.M., and Miller, W. 1998. A computer program for aligning a cDNA sequence with a genomic DNA sequence. *Genome Res.* **8**: 967–974.
- Freitag, J., Krieger, J., Strotmann, J., and Breer, H. 1995. Two classes of olfactory receptors in *Xenopus laevis*. *Neuron* **15**: 1383–1392.
- Glusman, G., Bahar, A., Sharon, D., Pilpel, Y., White, J., and Lancet, D. 2000a. The olfactory receptor gene superfamily: Data mining, classification, and nomenclature. *Mamm. Genome* **11**: 1016–1023.
- Glusman, G., Sosinsky, A., Ben-Asher, E., Avidan, N., Sonkin, D., Bahar, A., Rosenthal, A., Clifton, S., Roe, B., Ferraz, C., et al. 2000b. Sequence, structure, and evolution of a complete human olfactory receptor gene cluster. *Genomics* **63**: 227–245.
- Godfrey, P.A., Malnic, B., and Buck, L.B. 2004. The mouse olfactory receptor gene family. *Proc. Natl. Acad. Sci.* **101**: 2156–2161.
- Hertz, G. and Stormo, G. 1999. Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* **15**: 563–577.
- Hirota, J. and Mombaerts, P. 2004. The LIM-homeodomain protein Lhx2 is required for complete development of mouse olfactory neurons. *Proc. Natl. Acad. Sci.* **101**: 8751–8755.
- Hoppe, R., Weimer, M., Beck, A., Breer, H., and Strotmann, J. 2000. Sequence analyses of the olfactory receptor gene cluster mOR37 on mouse chromosome 4. *Genomics* **66**: 284–295.
- Hoppe, R., Frank, M.E., Breer, H., and Strotmann, J. 2003. The clustered olfactory receptor gene family 262: Genomic organization, promoter elements, and interacting transcription factors. *Genome Res.* **13**: 2674–2685.
- Kolterud, A., Alenius, M., Carlsson, L., and Bohm, S. 2004. The Lim homeobox gene Lhx2 is required for olfactory sensory neuron identity. *Development* **131**: 5319–5326.
- Kozak, M. 1999. Initiation of translation in prokaryotes and eukaryotes. *Gene* **234**: 187–208.
- Kudrycki, K., Stein-Izsak, C., Behn, C., Grillo, M., Akesson, R., and Margolis, F. 1993. Olf-1 binding site: Characterization of an olfactory neuron-specific promoter motif. *Mol. Cell. Biol.* **13**: 3002–3014.
- Lane, R., Cutforth, T., Young, J., Athanasiou, M., Friedman, C., Rowen, L., Evans, G., Axel, R., Hood, L., and Trask, B.J. 2001. Genomic analysis of orthologous mouse and human olfactory receptor loci. *Proc. Natl. Acad. Sci.* **98**: 7390–7395.
- Lewcock, J.W. and Reed, R.R. 2004. A feedback mechanism regulates monoallelic odorant receptor expression. *Proc. Natl. Acad. Sci.* **101**: 1069–1074.
- Lin, H. and Grosschedl, R. 1995. Failure of B cell differentiation in mice lacking the transcription factor EBF. *Nature* **376**: 263–267.
- Malnic, B., Hirono, J., Sato, T., and Buck, L.B. 1999. Combinatorial receptor codes for odors. *Cell* **96**: 713–723.
- Mombaerts, P. 2004. Odorant receptor gene choice in olfactory sensory neurons: The one receptor–one neuron hypothesis revisited. *Curr. Opin. Neurobiol.* **14**: 31–36.
- Mombaerts, P., Wang, F., Dulac, C., Chao, S., Nemes, A., Mendelsohn, M., Edmondson, J., and Axel, R. 1996. Visualizing an olfactory sensory map. *Cell* **87**: 675–686.
- Ngai, J., Dowling, M.M., Buck, L., Axel, R., and Chess, A. 1993. The family of genes encoding odorant receptors in the channel catfish. *Cell* **72**: 657–666.
- Pavesi, G., Mereghetti, P., Mauri, G., and Pesoli, G. 2004. Weeder Web: Discovery of transcription factor binding sites in a set of sequences from co-regulated genes. *Nucleic Acids Res.* **32**: W199–W203.
- Pesole, G., Mignone, F., Gissi, C., Grillo, G., Licciulli, F., and Liuni, S. 2001. Structural and functional features of eukaryotic mRNA untranslated region. *Gene* **276**: 73–81.
- Qasba, P. and Reed, R.R. 1998. Tissue and zonal-specific expression of an olfactory receptor transgene. *J. Neurosci.* **18**: 227–236.
- Ressler, K.J., Sullivan, S.L., and Buck, L.B. 1993. A zonal organization of odorant receptor gene expression in the olfactory epithelium. *Cell* **73**: 597–609.
- Rothman, A., Feinstein, P., Hirota, J., and Mombaerts, P. 2005. The promoter of the mouse odorant receptor gene M71. *Mol. Cell. Neurosci.* **28**: 535–546.
- Serizawa, S., Ishii, T., Nakatani, H., Tsuboi, A., Nagawa, F., Asano, M., Sudo, K., Sakagami, J., Sakano, H., Ijiri, T., et al. 2000. Mutually exclusive expression of odorant receptor transgenes. *Nat. Neurosci.* **3**: 687–693.
- Serizawa, S., Miyamichi, K., Nakatani, H., Suzuki, M., Saito, M., Yoshihara, S., and Sakano, H. 2003. Negative feedback regulation ensures the one receptor–one olfactory neuron rule in the mouse. *Science* **302**: 2088–2094.
- Serizawa, S., Miyamichi, K., and Sakano, H. 2004. One neuron–one receptor rule in the mouse olfactory system. *Trends Genet.* **20**: 648–653.
- . 2005. Negative feedback regulation ensures the one neuron–one receptor rule in the mouse olfactory system. *Chem. Senses* **30**: i99–i100.
- Shykind, B.M. 2005. Regulation of odorant receptors: One allele at a time. *Hum. Mol. Genet.* **14**: R33–R39.
- Shykind, B.M., Rohani, S.C., O'Donnel, S., Nemes, A., Mendelsohn, M., Sun, Y., Axel, R., and Barnea, G. 2004. Gene switching and the stability of odorant receptor gene choice. *Cell* **117**: 801–815.
- Sosinsky, A., Glusman, G., and Lancet, D. 2000. The genomic structure of human olfactory receptor genes. *Genomics* **70**: 49–61.
- Stein, L., Mungall, C., Shu, S., Caudy, M., Mangone, M., Day, A., Nickerson, E., Stajich, J., Harris, T., Arva, A., et al. 2002. The Generic Genome Browser: A building block for a model organism system database. *Genome Res.* **12**: 1599–1610.
- Thompson, W., Rouchka, E., and Lawrence, C. 2003. Gibbs recursive sampler: Finding transcription factor binding sites. *Nucleic Acids Res.* **31**: 3580–3585.
- Vassalli, A., Rothman, A., Feinstein, P., Zapotocky, M., and Mombaerts, P. 2002. Minigenes impart odorant receptor-specific axon guidance in the olfactory bulb. *Neuron* **35**: 681–696.
- Vassar, R., Ngai, J., and Axel, R. 1993. Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. *Cell* **74**: 309–318.

- Voss, T.S., Healer, J., Marty, A.J., Duffy, M.F., Thompson, J.K., Beeson, J.G., Reeder, J.C., Crabb, B.S., and Cowman, A.F. 2006. A *var* gene promoter controls allelic exclusion of virulence genes in *Plasmodium falciparum* malaria. *Nature* **439**: 1004–1008.
- Wang, M.M., Tsai, R.Y.L., Schrader, K.A., and Reed, R.R. 1993. Genes encoding components of the olfactory signal transduction cascade contain a DNA binding site that may direct neuronal expression. *Mol. Cell. Biol.* **13**: 5805–5813.
- Wang, S., Tsai, R., and Reed, R. 1997. The characterization of the Olf-1/EBF-like HLH transcription factor family: Implications in olfactory gene regulation and neuronal development. *J. Neurosci.* **17**: 4149–4158.
- Wang, F., Nemes, A., Mendelsohn, M., and Axel, R. 1998. Odorant receptors govern the formation of a precise topographic map. *Cell* **93**: 47–60.
- Wang, S., Betz, A., and Reed, R. 2002. Cloning of a novel Olf-1/EBF-like gene, O/E-4, by degenerate oligo-based direct selection. *Mol. Cell. Neurosci.* **20**: 404–414.
- Wang, S.S., Lewcock, J.W., Feinstein, P., Mombaerts, P., and Reed, R.R. 2003. Genetic disruptions of O/E2 and O/E3 genes reveal involvement in olfactory receptor neuron projection. *Development* **131**: 1377–1388.
- Young, J.M., Friedman, C., Williams, E.M., Ross, J.A., Tonnes-Priddy, L., and Trask, B.J. 2002. Different evolutionary processes shaped the mouse and human olfactory receptor gene families. *Hum. Mol. Genet.* **11**: 535–546.
- Young, J.M., Shykind, B.M., Lane, R.P., Tonnes-Priddy, L., Ross, E.M., Walker, M., Williams, E.M., and Trask, B.J. 2003. Odorant receptor expressed sequence tags demonstrate olfactory expression of over 400 genes, extensive alternate splicing and unequal expression levels. *Genome Biol.* **4**: R71.
- Zhang, X. and Firestein, S. 2002. The olfactory receptor gene superfamily of the mouse. *Nat. Neurosci.* **5**: 124–133.

Received February 22, 2006; accepted in revised form July 6, 2006.

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)