

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE BIOCÊNCIAS

Ricardo D'Oliveira Albanus

**Estudo da evolução dos receptores químicos nos organismos *Mus musculus*, *Homo sapiens* e *Rattus norvegicus* através de ferramentas de
Biologia de Sistemas**

Porto Alegre

2011

Ricardo D'Oliveira Albanus

Estudo da evolução dos receptores químicos nos organismos *Mus musculus*, *Homo sapiens* e *Rattus norvegicus* através de ferramentas de Biologia de Sistemas

Trabalho de Conclusão de Curso, apresentado ao Instituto de Biociências da UFRGS em preenchimento parcial dos requisitos para obtenção do título de Bacharel em Ciências Biológicas.

Orientador: Prof. Dr. José Cláudio Fonseca Moreira

Porto Alegre

2011

Ricardo D'Oliveira Albanus

Estudo da evolução dos receptores químicos nos organismos *Mus musculus*, *Homo sapiens* e *Rattus norvegicus* através de ferramentas de Biologia de Sistemas

Trabalho de Conclusão de Curso, apresentado ao Instituto de Biociências da UFRGS em preenchimento parcial dos requisitos para obtenção do título de Bacharel em Ciências Biológicas.

Comissão Examinadora:

Prof. Dr. Diego Bonatto
Centro de Biotecnologia - UFRGS

Profa. Dra. Rita Maria Cunha de Almeida
Instituto de Física – UFRGS

Prof. Dr. José Cláudio Fonseca Moreira
Departamento de Bioquímica – UFRGS

Porto Alegre, 8 de Julho de 2011

Aos meus pais e todos os outros companheiros de jornada.
Vocês sabem quem vocês são.

*“Lights will guide you home
and ignite your bones.”*

– Coldplay

Resumo

A quimiorrecepção (QR) é uma modalidade sensorial muito importante em todos os animais. Estudos recentes permitiram a identificação das famílias de genes QR, divididas em receptores olfatórios (OL), vomeronasais (VN), gustatórios (TS) e de traços de aminas (TAAR). Utilizando dados do Gene Ontology Project (GO), nós conseguimos determinar os genes envolvidos em cada modalidade sensorial química nos organismos *Mus musculus*, *Homo sapiens* e *Rattus norvegicus*. Os GO IDs utilizados foram 0004984 e 0007608, 0050909 e 0008527, 0016503 e 0019236, respectivamente, para os receptores OL, TS e VN (os TAAR foram retirados das análises por falta de dados disponíveis). Adquirimos os parâmetros das redes funcionais formadas pelas proteínas codificadas por esses genes através do banco de dados STRING-DB e Biomart. Os critérios utilizados para nossas análises foram conectividade [$k(i)$], clusterização [$c(i)$] e índice estimado de plasticidade (EPI) desses receptores. Dividimos o conjunto de dados em proteínas envolvidas diretamente na QR e proteínas envolvidas na transdução do sinal e posteriores cascatas químicas, denominados de Grupo dos Receptores (GR) e dos Acessórios (GA). Comparando as médias de EPIs do GR e GA através de ANOVA simples e teste Tuckey, foi observado que as médias do GR eram significativamente maiores que as do GA ($p < 0,001$). Esses testes foram repetidos em cada subgrupo receptor com seu respectivo subgrupo acessório e foram observados os mesmos resultados gerais, além de ser encontrado nos três organismos uma similaridade entre os OL e TS2, VN e TS1 ($p < 0,05$). Foi analisada a entropia da distribuição dos valores de $k(i)$ e $c(i)$ e foi encontrado que estes eram significativamente maiores ($p < 0,001$) no GA. Isso se reflete na distribuição de forma mais uniforme desses valores no GA e de forma mais concentrada no GR. Os resultados sugerem que a seleção natural é menos estrigente nos receptores do que no maquinário de transdução de sinal, permitindo uma variação mais dinâmica desses. Essa capacidade é muito importante para a adaptação dos organismos aos diversos nichos ecológicos em que estão inseridos. Estes dados estão em acordo com os atuais modelos propostos para a evolução dos receptores químicos.

Palavras-chaves: Biologia de Sistemas, Quimiorrecepção. Evolução

Abstract

Chemoreception (CR) is a very important sensory modality in all animals. Recent studies allowed the characterization of CR gene families, divided in olfactory receptors (OL), vomeronasal (VN), gustatory (TS) and trace amine (TAAR). Using data gathered from The Gene Ontology Project (GO), we were able to amass the genes associated with each type of sensory perception. The GO ID's used for this work were 0004984 and 0007608, 0050909 and 0008527, 0016503 and 0019236, for olfactory, gustatory and vomeronasal ontology groups, respectively. We sorted these genes in groups by receptor modality, with one further division of the Taste Group into the Taste-1 and Taste-2 receptors groups. We acquired the network parameters of the proteins translated from these genes groups using data from STRING Database. We used as criterions for our analysis the connectivity and clusterization indexes (*i.e.* $k(i)$ and $c(i)$, respectively). Also used in our analysis was the Evolutionary Plasticity Index (EPI). We subsequently divided the whole dataset into two more groups: those of proteins directly involved in the chemical binding of the stimuli, named Receptor Group (RG); and those involved into signal transduction and further chemical pathways, named Accessory Group (AG). We compared the mean EPIs of all the RGs and AGs using One-Way ANOVA and Tukey test and found that the AGs' means were significantly lower than the RGs' ($p < 0,001$). We also found that the OL and TS2, VN and TS1 EPI means were similar in all three organisms ($p < 0,05$). We compared the overall distribution of the $c(i)$ and $k(i)$ on and observed that all of the RGs presented very low values, whereas the AGs presented a more uniform distribution. The sole exception in this analysis was the Olfactory RGs' $c(i)$, which all have a more extreme oriented distribution, with proteins with very high and very low values, but no intermediary ones. Also with ANOVA, we observed that the Entropy of AGs' $k(i)$ distribution was significantly higher than the RGs' ($p < 0,05$). Our findings suggest that the evolutionary plasticity in the RGs is higher than that on the AGs. This means that the selective pressures among the chemoreceptors are weaker than that on the transduction machinery, resulting in a greater diversity for these receptors. Put together, these results support the Birth-and-Death Evolution model proposed for chemosensory receptors.

Keywords: System Biology, Chemoreception, Evolution

Lista de figuras e tabelas

Figura 1: Exemplo de uma rede.....	17
Figura 2: Médias dos EPIs dos grupos de receptores e acessórios.....	20
Figura 3: Distribuição dos valores de conectividade das proteínas dos Grupos de Acessórios e Receptores.....	22
Figura 4: Distribuição dos valores de clusterização das proteínas dos Grupos de Acessórios e Receptores.....	23
Figura 5: Diferenças na entropia das distribuições de $k(i)$ e $c(i)$ entre os grupos.....	24
Figura 6: Médias dos EPIs dos subgrupos de receptores e acessórios.....	24
Figura 7: Comparação entre as médias dos EPIs dos subgrupos.....	25
Tabela 1: Grupos do <i>Gene Ontology</i> utilizados nas análises.....	18
Tabela 2: Divisão dos grupos e subgrupos de receptores químicos.....	18
Tabela 3: Número de genes incluídos em cada subgrupo de análise.....	19

Lista de abreviações e siglas

ANOVA – Análise de Variância

BaDE – *Birth-and-Death Evolution*

$c(i)$ – Clusterização

EPI – Índice de Plasticidade Evolutiva (*Evolutionary Plasticity Index*)

GA – Grupo dos Acessórios

GO – *Gene Ontology Project*

GR – Grupo dos Receptores

$k(i)$ – Conectividade

MHC – Complexo principal de histocompatibilidade (*Major histocompatibility complex*)

OL – Receptores olfatórios

S – Entropia

TAAR – Receptores de traços de aminas

TS – Receptores gustatórios

VN – Receptores vomeronasais

Sumário

Resumo	5
<i>Abstract</i>	6
Lista de figuras e tabelas	7
Lista de abreviações e siglas.....	8
1 Introdução.....	11
1.2 Os receptores químicos	11
1.3 O modelo <i>Birth-and-Death Evolution</i>	13
1.4 A Biologia de Sistemas	13
2 Objetivos.....	15
2.1 Objetivos gerais	15
2.2 Objetivos específicos	15
3 Materiais e Métodos	16
3.1 A rede proteica e sua modelagem	16
3.2 A escolha dos genes e proteínas	17
3.3 A análise dos dados.....	18
4 Resultados.....	20
4.1 Análises entre os grupos:	20
4.1.1 Os EPIs do Grupo de Acessórios são menores que os do Grupo de Receptores	20
4.1.2 As distribuições dos valores $k(i)$ e $c(i)$ são diferentes entre GA e GR.....	20
4.2 Análises entre os subgrupos:.....	21
4.2.1 As médias individuais de EPIs dos subgrupos receptores são maiores que as dos acessórios	21
4.2.2 As médias de EPI são homogêneas entre os OL e TS2, assim como entre os TS1 e VN	25
4.2.3 A distribuição de $k(i)$ e $c(i)$ dos TS1 e OL é diferente dos demais subgrupos....	25
5 Discussão	26
5.1 As variações na plasticidade dos grupos e subgrupos	26
5.2 Os parâmetros funcionais de redes dos quimiorreceptores.....	28
6 Conclusões.....	30
6.1 Conclusão geral.....	30

6.2 Conclusões específicas.....	30
6.2.1 Os parâmetros de redes são distintos entre os quimiorreceptores e o maquinário de transdução de sinal	30
6.2.2 Os parâmetros de redes são distintos entre os próprios quimiorreceptores.....	30
7 Referências Bibliográficas.....	31

1 Introdução

1.1 A quimiorrecepção

A quimiorrecepção é uma das modalidades sensoriais mais importantes nos organismos. Seu surgimento ocorre cedo na evolução, estando presente na árvore evolutiva desde as bactérias mais primitivas (BERG, 1975). Os quimiorreceptores estão envolvidos direta ou indiretamente nas principais atividades dos organismos, auxiliando na procura de alimento, busca de parceiros sexuais, reconhecimento de territórios e percepção de perigos.

Estudos recentes de sequências genômicas completas permitiram a identificação e quantificação da maioria dos genes codificadores desses receptores (*e.g.* ADLER *et al*, 2000; NELSON *et al*, 2001; RODRIGUEZ *et al*, 2002; SHI *et al*, 2005; YOUNG *et al*, 2005). Os trabalhos publicados sobre o tema demonstram que eles ocorrem em grande número e que cada genoma possui uma considerável quantidade de pseudogenes associados. Além disso, existe uma significativa variação no número de cópias entre as espécies, de maneira que podemos encontrar diferenças na quantidade de receptores inclusive entre representantes de um mesmo gênero (NEI *et al*, 2008).

1.2 Os receptores químicos

Nos vertebrados, os receptores químicos são codificados por seis principais famílias multigênicas, nomeadas de acordo com a modalidade sensorial a qual estão associadas. São elas: receptores olfatórios (OR); receptores associados a traços de aminas (TAAR); receptores vomeronasais, subdivididos em tipos I e II (V1R e V2R, respectivamente); e receptores gustatórios, também separados em tipos I e II (T1R e T2R, respectivamente). As proteínas codificadas por esses genes são do tipo que possuem sete regiões α -hélice transmembrana e estão acopladas a uma proteína G (SHI *et al*, 2009). A expressão desses receptores foi estudada em algumas espécies de vertebrados no intuito de estabelecer quais órgãos e tecidos são responsáveis pela detecção de cada modalidade de estímulo químico. Os resultados encontrados são bastante uniformes, possibilitando uma generalização destes dados para todos os vertebrados.

Os OR e TAAR são expressos na região sensível da cavidade nasal, chamada de epitélio olfatório principal. Essa estrutura projeta aferências neuronais diretamente ao bulbo olfatório principal, localizado no córtex cerebral (HASIN-BRUMSHTEIN *et al*, 2009; HASHIGUCHI *et al*, 2007). Os T1R e T2R ocorrem nos botões gustatórios da língua, onde estão em contato direto com o alimento ingerido pelo animal (ADLER *et al*, 2000; NELSON *et al*, 2001). Essa disposição no início do processo alimentar permite que a ingestão de algumas substâncias nocivas seja interrompida antes de causar danos ao organismo. Os V1R e V2R ocorrem nas células do órgão vomeronasal, uma estrutura localizada entre o vômero e a cavidade nasal, que projeta aferências neuronais para o bulbo olfatório secundário e para o sistema límbico, modulando diretamente os comportamentos dos animais nos quais ocorre. Embora os seres humanos possuam no genoma algumas cópias dos receptores VIR, o seu órgão vomeronasal não é funcional e o gene que codifica o canal iônico *Trpc2*, responsável pela transdução desses sinais, foi perdido (SWANEY *et al*, 2009). Alguns desses receptores podem ser encontrados em outros tecidos, como nos testículos, cérebro, placenta e coração. A funcionalidade desta expressão ectópica, entretanto, continua desconhecida (FELDMESSER *et al*, 2006).

Os tipos de moléculas percebidas variam com a classe de receptor, podendo ser hidrossolúveis ou voláteis. Tanto os receptores vomeronasais quanto olfatórios interagem com substâncias voláteis, mas o primeiro grupo foi aparentemente selecionado para a detecção de substâncias envolvidas na comunicação intraespecífica (*e.g.* feromônios), e os últimos, para perceber os odores mais genéricos do ambiente. Essa segregação funcional não é uma regra, como no exemplo dos animais do gênero *Ornithorhynchus*, que durante o ato de submersão – situação em que fecham com membranas seus olhos, ouvidos e nariz – utilizam o seu bico eletrorreceptivo para detecção de alimento e o seu amplo repertório de receptores vomeronasais para a percepção genérica (NEI *et al*, 2008).

Os TAARs são considerados uma segunda classe de receptores olfatórios, capazes de detectar algumas aminas voláteis. Inicialmente se acreditava que eram receptores para algumas espécies de nitrogênio presentes em baixas concentrações no sistema nervoso central, que podiam estar envolvidas em alguns distúrbios psiquiátricos, mas agora se sabe que sua função é outra. Estudos demonstram que alguns TAARs de camundongos reconhecem aminas presente na urina, sugerindo que estão ligados na discriminação de pistas sociais (HASHIGUCHI *et al*, 2007). Os receptores gustatórios são responsáveis pela de-

tecção dos sabores doce, azedo, amargo, salgado e *umami* (que significa “delicioso” em japonês, e corresponde ao sabor do L-glutamato). Apenas os sabores salgado e azedo são detectados por canais iônicos. Os sabores doces e *umami* são percebidos pelos receptores T1R, e os amargos, que estão normalmente associados a compostos tóxicos, pelos T2R (ADLER *et al*, 2000; NELSON *et al*, 2001).

1.3 O modelo *Birth-and-Death Evolution*

Existem várias teorias atuais para explicar as variações dramáticas entre os quimiorreceptores dos organismos, tais como a adaptação a diferentes ambientes e processos de especiação. Devido ao limitado número de espécies estudadas e a nossa parca compreensão dos processos bioquímicos envolvidos nessas modalidades sensoriais, alguns pontos permanecem sem esclarecimentos concretos, tais como a variação drástica, tanto intra quanto interespecificamente, do número desses genes em vários grupos de animais, bem como o grande índice de pseudogenes associados.

Uma das teorias existentes é o modelo *Birth-and-Death Evolution* (BaDE). Ela foi inicialmente proposta para tentar explicar o surgimento das famílias gênicas codificadoras das subunidades do MHC e das imunoglobulinas no sistema imune, mas pôde ser aplicada para elucidar a ocorrência de outras famílias gênicas dotadas de notável variabilidade entre seus membros, como os receptores químicos (NEI *et al*, 1997). O modelo postula que, em algumas famílias, os novos genes são criados por repetidas duplicações, podendo alguns destes permanecer por longos períodos de tempo no genoma. Outros genes, entretanto, são logo removidos ou inativados por mutações deletérias, de forma diferente da evolução concertada (*concerted evolution*), onde as variações ocorrem de uniformemente em todos os membros (NEI *et al*, 2005).

1.4 A Biologia de Sistemas

Historicamente, o reducionismo permeia as diretrizes da pesquisa científica. Sistemas inteiros, como células vivas, são desmembrados até o nível de suas partes fundamentais (*e.g.* proteínas, lipídios e outras biomoléculas). Esses componentes são então estudados em detalhe, na tentativa de estabelecer uma coerência para o funcionamento do todo. É, portanto, um novo paradigma para a biologia moderna entender a dinâmica das complexas teias de interações que determinam a estrutura e função dos sistemas biológicos.

O desenvolvimento de ferramentas de coletas de alta densidade de dados genômicos e metabólicos (*microarrays*, *yeast two-hybrid method*, etc.) permitiu, nas duas últimas décadas, acesso a todos os componentes de uma célula a qualquer dado instante, tornando possível vislumbrar a intrincada dinâmica das redes genômicas e proteicas. Para analisar este grande volume de dados é necessária a elaboração de modelos que compreendam as mudanças topológicas e propriedades dinâmicas oriundas dessas interações tão múltiplas, de forma a criar uma visão quantitativa e, portanto, comparável das complexas redes que regulam e determinam o comportamento celular (HARRINGTON *et al*, 2008).

A Biologia de Sistemas surge a partir da compreensão de que as características estruturais das redes de interações moleculares de uma célula são compartilhadas em grande grau com outros sistemas complexos já estudados (HARTWELL *et al*, 1999; YAMADA *et al*, 2009). Esta universalidade inesperada indica que leis similares podem reger a formação e evolução das redes de alta complexidade na natureza, permitindo que utilizemos o conhecimento adquirido no estudo de sistemas não biológicos bem mapeados na caracterização das relações intrincadas contidas na organização das funções celulares (BARABASI *et al*, 2004).

2 Objetivos

2.1 Objetivos gerais

Este trabalho tem por objetivo analisar, através de uma abordagem de Biologia de Sistemas, a evolução dos receptores químicos nas espécies *Mus musculus*, *Homo sapiens* e *Rattus norvegicus*, tentando consolidar esta metodologia como mais uma ferramenta no estudo da evolução genômica, em acordo com as teorias evolutivas já propostas para estes tipos de receptores.

2.2 Objetivos específicos

1. Determinar os parâmetros funcionais das redes biológicas compostas pelos genes de quimiorrecepção;
2. Determinar se existem diferenças nos parâmetros de rede e plasticidade dos quimiorreceptores e do maquinário responsável pela transdução de sinal;
3. Determinar se existem diferenças nos parâmetros de rede e plasticidade entre os próprios subgrupos de quimiorreceptores;
4. Verificar se é possível estabelecer ligação entre os dados levantados e os modelos de evolução já propostos para esses receptores, tentando consolidar a Biologia de Sistemas como mais uma ferramenta no estudo da evolução genômica.

3 Materiais e Métodos

3.1 A rede proteica e sua modelagem

A rede funcional de um proteoma é inferida através das interações que ocorrem entre seus membros. É válido destacar que, por interações proteicas, não se entende somente aquelas do tipo físicas diretas (*e.g.* fosforilações, dimerizações, etc.), mas também as associações funcionais, como a participação em um mesmo processo celular ou rota metabólica. Na modelagem da rede, cada proteína é representada por um nó, e cada interação, por uma ligação não-direcionada entre os nós correspondentes (Figura 1). Os principais índices utilizados para o estudo de redes são:

Conectividade $[k(i)]$ – É dado pelo número de conexões que um dado nó i realiza na rede. No exemplo da Figura 1, as proteínas 1, 2 e 3 teriam valor $k(i)$ igual a 5, 3 e 4, respectivamente.

Clusterização $[c(i)]$ – É dado pela interatividade entre os vizinhos de um dado nó i . Esse valor é calculado pela razão entre o número de ligações efetivas e todas as possíveis (*i.e.* se todos os vizinhos de i estivessem conectados), de acordo com a fórmula $c(i) = \frac{2n_i}{k_i(k_i-1)}$, onde n_i é o número de conexões dos vizinhos de i entre eles mesmos. No exemplo da Figura 1, os valores de $c(i)$ seriam 0.5, 0 e 0, respectivamente, para as proteínas 1, 2 e 3.

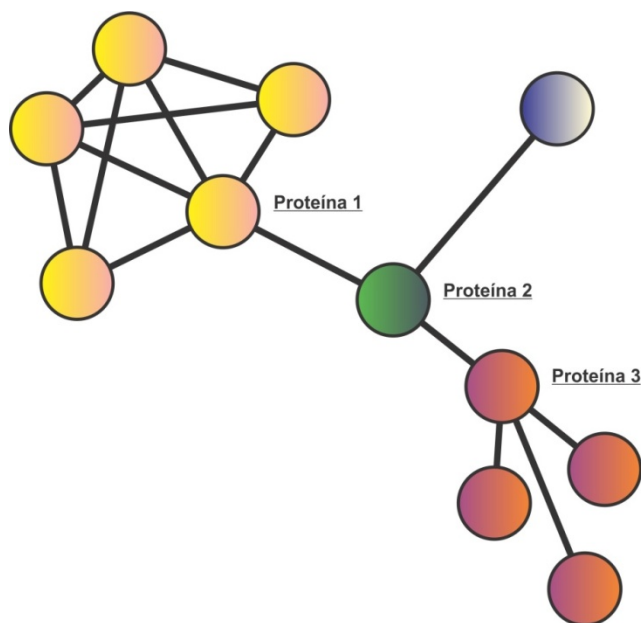


Figura 1: Exemplo de uma rede. Os nós representam proteínas e as ligações representam interações funcionais.

3.2 A escolha dos genes e proteínas

Para nosso estudo, apenas os genes dos organismos *Mus musculus*, *Homo sapiens* e *Rattus norvegicus* foram utilizados. As três espécies foram escolhidas pela robusteza e quantidade de dados disponíveis sobre seus genomas, permitindo-nos uma análise mais concreta e com menor possibilidade de inferências errôneas. A escolha dos receptores químicos foi realizada utilizando os genes dos grupos de ontologia disponibilizados nos bancos de dados do *Gene Ontology Consortium* (ASHBURNER *et al*, 2000) (Tabela 1). A família dos TAAR foi retirado das análises devido a pouca quantidade de dados disponíveis.

Os genes foram classificados em grupos por modalidade sensorial, com uma divisão adicional do Grupo Gustatório (TS) em TS1 e TS2, para distinguir entre as duas classes dessa família de receptores. Por falta de dados disponíveis, foram considerados como uma só as duas famílias de VN. Posterior a essa classificação, cada grupo de ontologia foi separado em mais duas categorias funcionais: Grupo dos Receptores (GR) e Grupo dos Acessórios (GA). O GR consistia nos genes das proteínas diretamente envolvidas na ligação do estímulo químico, e o GA, em todos os demais genes envolvidos no maquinário da transdução de sinal e outras rotas bioquímicas. A seriação final ficou

definida em dois grandes grupos, o dos Acessórios e dos Receptores, e seus subgrupos – olfatórios (OL), gustatórios (TS1 e TS2) e vomeronasais (VN) (Tabela 2).

Receptores	GO groups
Receptores olfatórios	GO:0004984 – MF <i>Olfactory receptor activity</i> GO:0007608 – BP <i>Sensory perception of smell</i>
Receptores gustatórios	GO:0008527 – MF <i>Taste receptor activity</i> GO:0050909 – BP <i>Sensory perception of taste</i>
Receptores vomeronasais	GO:0016503 – MF <i>Pheromone receptor activity</i> GO:0019236 – BP <i>Response to Pheromone</i>

Tabela 1: Grupos do *Gene Ontology* utilizados nas análises.

Grupo	Subgrupos (por espécie e modalidade)
Grupo dos Receptores	<i>M. musculus</i> OL, TS1, TS2, VN <i>H. sapiens</i> OL, TS1, TS2, VN <i>R. norvegicus</i> OL, TS1, TS2, VN
Grupo dos Acessórios	<i>M. musculus</i> OL Ac, TS Ac, VN Ac <i>H. sapiens</i> OL Ac, TS Ac, VN Ac <i>R. norvegicus</i> OL Ac, TS Ac, VN Ac

Tabela 2: Divisão dos grupos e subgrupos de receptores químicos.

Os parâmetros de rede das proteínas codificadas por esses genes foram levantadas através banco de dados *STRING – Known and Predicted Protein-Protein Interactions*, versão 8.3 (JENSEN *et al*, 2009), utilizando os ENSEMBL IDs correspondentes. Para adquirir o ENSEMBL ID dos genes, foi realizada uma busca cruzada nos bancos de dados do *HUGO Genome Name Consortium* (HGNC), *Mouse Genome Informatics* (MGI) e *Rat Genome Database* (RGD). As coletas de dados foram realizadas inicialmente através da ferramenta BioMart e, posteriormente, entre os próprios bancos. Os genes cujos ENSEMBL IDs apresentaram divergências entre os bancos foram curados manualmente e aqueles que, ainda assim, não puderam ser definidos foram retirados das análises (Tabela 3). Para a aquisição dos parâmetros das redes funcionais, utilizamos como índice de confiança o score combinado de 0,7. Complementarmente, foram somente levantados os dados de interações geradas por experimentos e pelos diversos bancos de dados curados disponíveis na internet. Essa pode ser considerada uma análise de média-alta stringência.

3.3 A análise dos dados

Foram utilizados para nosso estudo os índices de conectividade $[k(i)]$ e clusterização $[c(i)]$ dos receptores. Juntamente com esses dados, foi também calculado o índice de plasticidade evolutiva (EPI) dos ortólogos de cada receptor, calculado através da fórmula (1), onde H_α representa a diversidade de ortólogos em diferentes espécies e D_α , sua abundância (DALMOLIN *et al*, 2011). A busca dos dados de ortologia das proteínas foi feita no banco de dados *String-DB*, utilizando o ENSEMBL ID destas.

$$(1) EPI = 1 - \frac{H_\alpha}{\sqrt{D_\alpha}}$$

Tendo em mãos esses índices, procedeu-se para a análise estatística dos grupos de dados, a fim de buscar correlações significativas entre as diferentes famílias de receptores. Foi calculada a distribuição dos valores, bem como suas médias e desvios-padrão. Todas as comparações entre médias neste trabalho foram realizadas utilizando ANOVA simples com teste de Tuckey. As análises dos valores de $k(i)$ e $c(i)$ foram realizadas através do cálculo da entropia (S) da distribuição destes valores, que determina o grau de desordem de sua dispersão. O valor de S é calculado utilizando a fórmula (2), onde p representa a probabilidade de ocorrência dos valores.

$$(2) S = - \sum p \ln p$$

Espécie	Subgrupo	Número de Genes
<i>Mus musculus</i>	OL	1107 (17)
	TS1	3 (18)
	TS2	34 (18)
	VN	96 (1)
<i>Homo sapiens</i>	OL	347 (18)
	TS1	3 (14)
	TS2	17 (14)
	VN	3 (0)
<i>Rattus norvegicus</i>	OL	1202 (24)
	TS1	3 (24)
	TS2	35 (24)
	VN	91 (1)

Tabela 3: Número de genes incluídos em cada subgrupo de análise. Os valores entre parênteses representam os genes do Grupo dos Acessórios.

4 Resultados

4.1 Análises entre os grupos:

4.1.1 Os EPIs do Grupo de Acessórios são menores que os do Grupo de Receptores

Foi utilizada uma ANOVA simples associada ao teste Tuckey para identificar possíveis diferenças entre o conjunto das médias dos EPIs dos Grupos Acessórios e Receptores. O resultado encontrado foi que as médias do GA e GR eram significativamente diferentes ($p < 0,001$), sendo as do GA menores que as do GR (Figura 2).

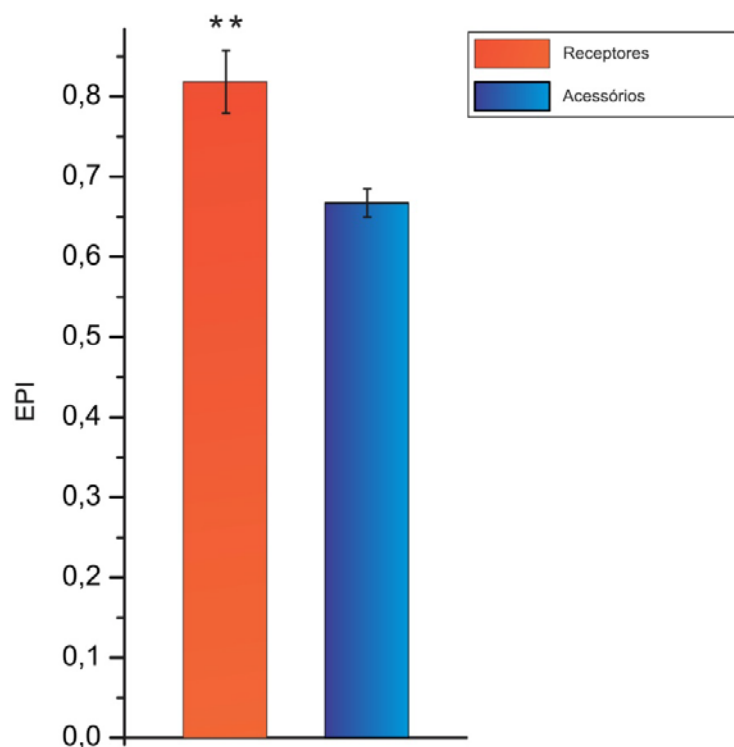


Figura 2: Médias dos EPIs dos grupos de receptores e acessórios. Asteriscos duplos indicam nível de significância $p < 0,001$; as barras indicam um desvio-padrão acima e abaixo da média.

4.1.2 As distribuições dos valores $k(i)$ e $c(i)$ são diferentes entre GA e GR

A distribuição dos valores de $k(i)$ e $c(i)$ foi significativamente diferente entre os grupos. No GA elas se apresentaram de forma aproximadamente normal, com a maioria

dos valores próximos de zero. No GR foram encontrados dois comportamentos de dispersão: bimodal nos receptores olfatórios e unimodal nos demais (Figuras 3 e 4).

Para verificar matematicamente essas diferenças, utilizou-se como parâmetro de análise a entropia dessas distribuições. Foi realizada uma ANOVA simples com teste Tuckey para verificar se havia diferenças significativas entre o conjunto de valores. O resultado encontrado foi que, em nível de significância $p < 0,001$, os valores de entropia diferiam significativamente entre os grupos (Figura 5).

4.2 Análises entre os subgrupos:

4.2.1 As médias individuais de EPIs dos subgrupos receptores são maiores que as dos acessórios

À exceção do subgrupo VN de *Homo sapiens*, que não apresenta sua contraparte acessória, a média dos EPIs dos receptores é significativamente maior ($p < 0,001$ e $p < 0,05$) do que a do maquinário de transdução de sinal (Figura 6)

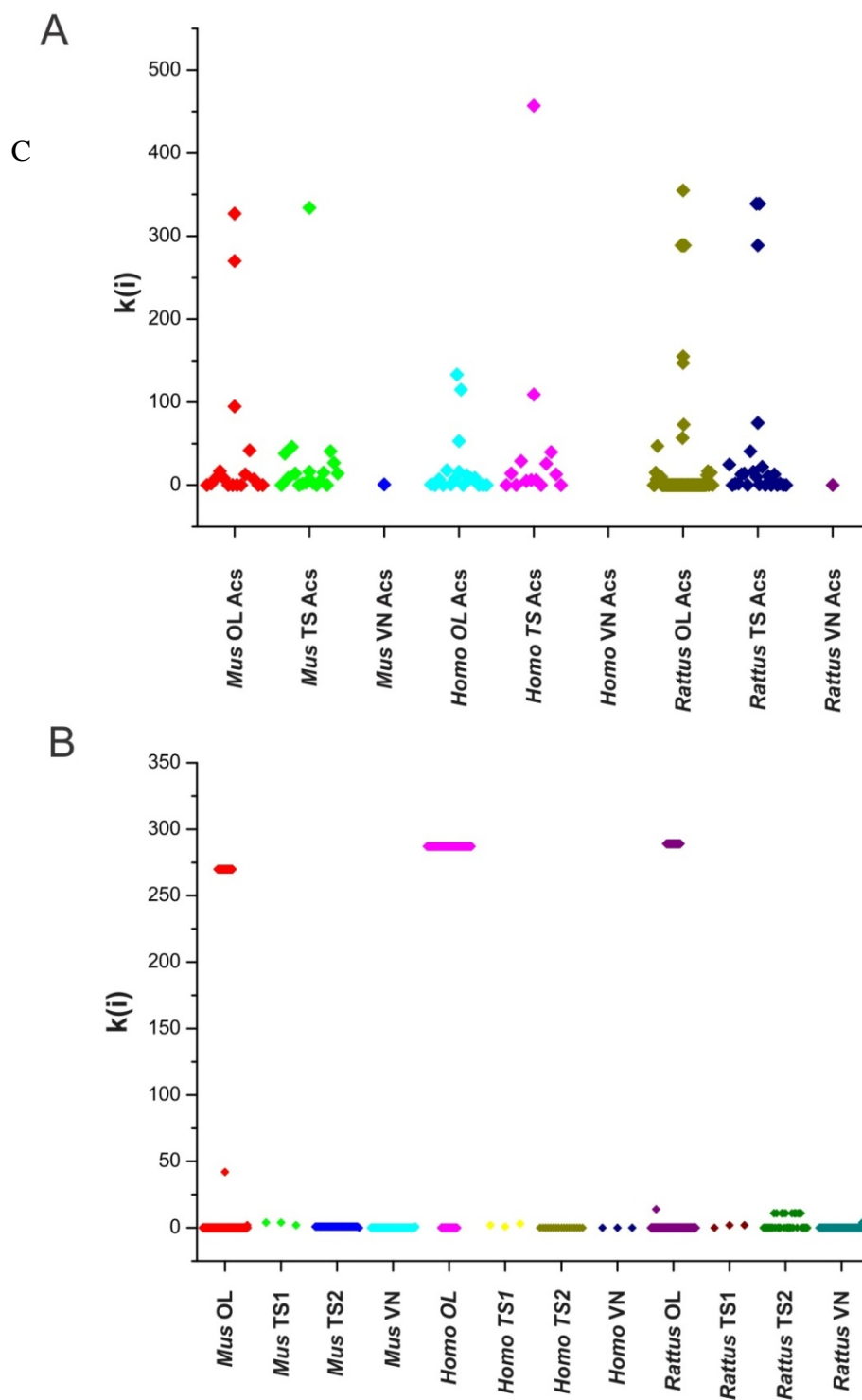


Figura 3: Distribuição dos valores de conectividade das proteínas dos Grupos de Acessórios (A) e Receptores (B).

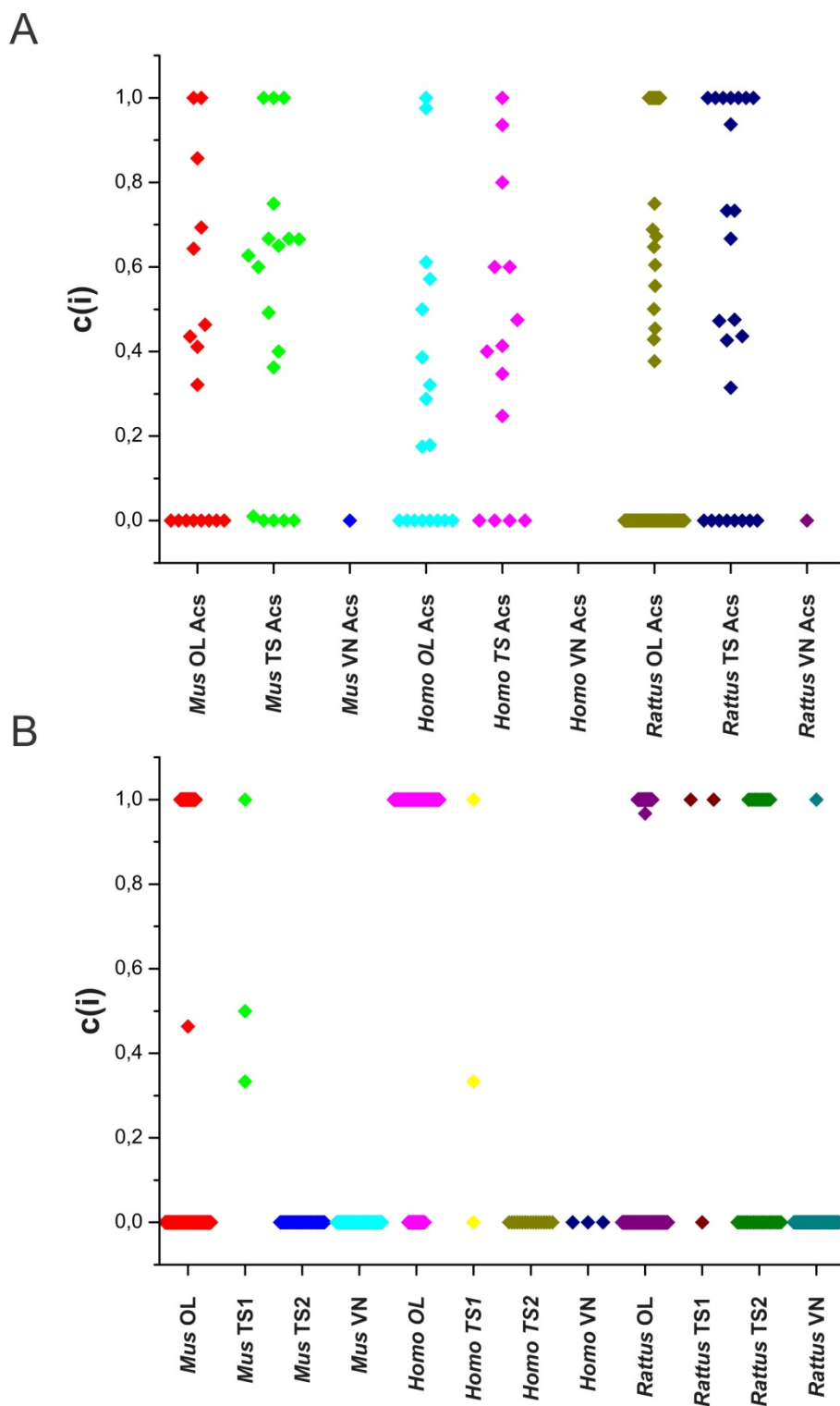


Figura 4: Distribuição dos valores de clusterização das proteínas dos Grupos de Acessórios (A) e Receptores (B).

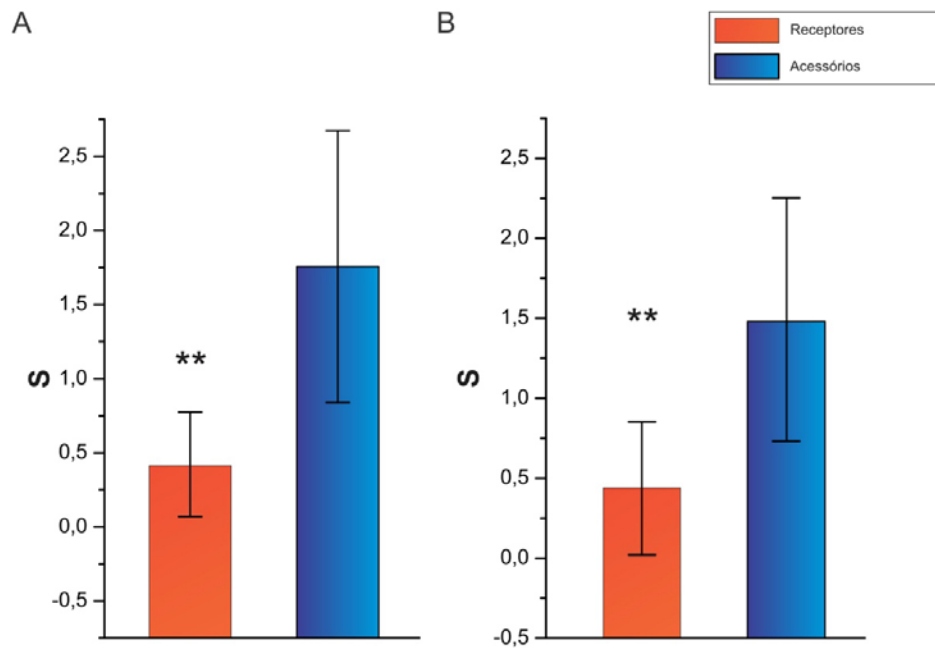


Figura 5: Diferenças na entropia das distribuições de $k(i)$ (A) e $c(i)$ (B) entre os grupos. Os duplos asteriscos indicam significância $p < 0,001$; as barras indicam um desvio-padrão acima e abaixo da média.

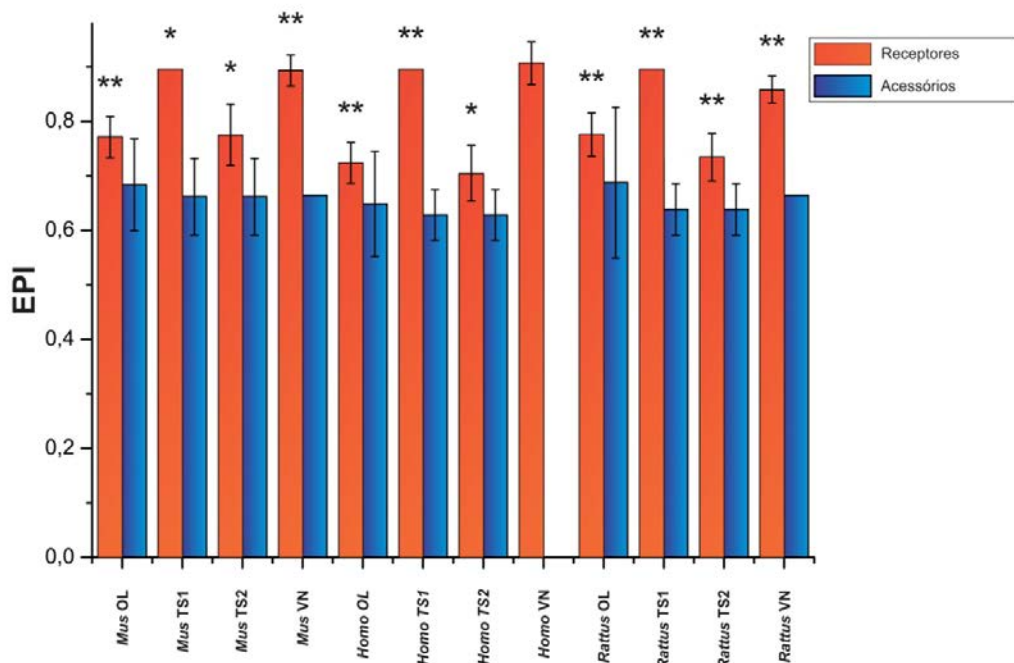


Figura 6: Médias dos EPIs dos subgrupos de receptores e acessórios. Asteriscos duplos indicam nível de significância $p < 0,001$ e simples, $p < 0,05$; as barras indicam um desvio-padrão acima e abaixo da média.

4.2.2 As médias de EPI são homogêneas entre os OL e TS2, assim como entre os TS1 e VN

Comparando as médias de EPI entre os diferentes subgrupos de cada espécie, utilizando ANOVA e teste Tuckey, foi encontrado que os valores de OL e TS2, assim como de TS1 e VN eram estatisticamente similares entre si ($p < 0,05$) em todas as espécies estudadas (Figura 7).

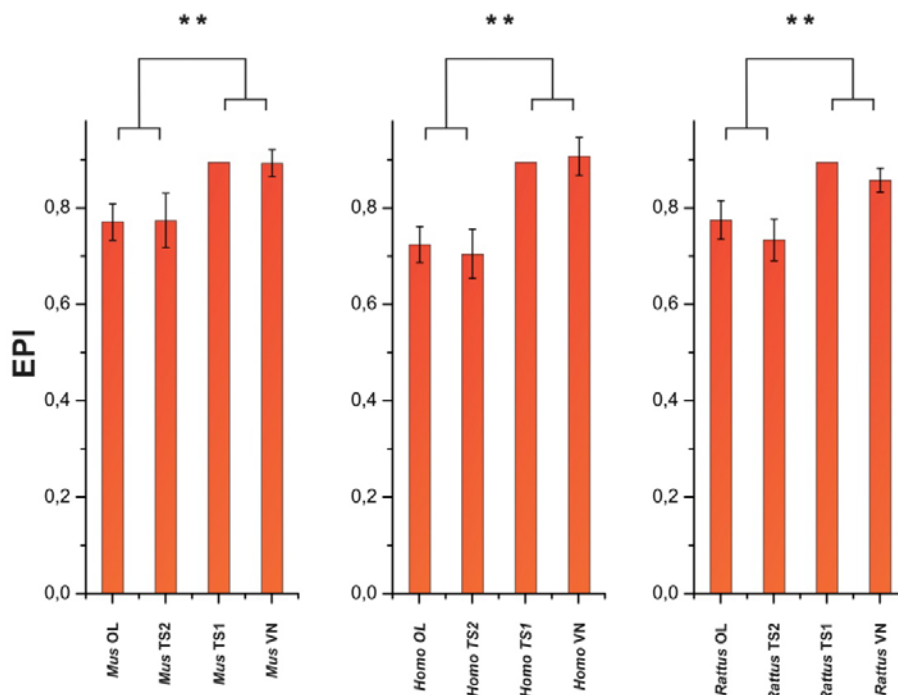


Figura 7: Comparação entre as médias dos EPIs dos subgrupos. Os colchetes horizontais indicam os subgrupos similares entre si ($p < 0,05$) e os asteriscos duplos indicam os diferentes ($p < 0,001$); as barras indicam um desvio-padrão acima e abaixo da média.

4.2.3 A distribuição de $k(i)$ e $c(i)$ dos TS1 e OL é diferente dos demais subgrupos

Analisando a distribuição dos valores de conectividade e clusterização dos subgrupos, foi encontrado que os TS1 e OL apresentam uma distribuição bimodal em todas as espécies. À exceção dos TS2 de *Rattus norvegicus*, todos os outros subgrupos apresentam valores de $c(i)$ e $k(i)$ próximos ou iguais a zero (Figuras 3A e 3C).

5 Discussão

5.1 As variações na plasticidade dos grupos e subgrupos

As complexas interações ecológicas entre as espécies nos diversos habitats estão constantemente desafiando sua sobrevivência. Plantas renovam seu já extenso repertório de produtos químicos defensivos de geração em geração, forçando aqueles organismos que delas dependem para sua sobrevivência a manterem igualmente atualizados seus mecanismos de detecção e detoxificação dessas substâncias (HARTMANN, 2007). Os constantes processos de especiação também tornam necessário que as diferentes espécies consigam detectar com segurança parceiros compatíveis, evitando desperdiçar energia em encontros que produzam descendência menos apta por fenômenos deletérios de hibridização ou endogamia.

Os receptores químicos são a primeira linha de contato dos organismos com os desafios que o seu ambiente proporciona. É, portanto, necessário que haja uma plasticidade em sua evolução, possibilitando às espécies se adaptarem dinamicamente a esses desafios. Em acordo com o modelo BaDE (NEI *et al*, 2005), nossos dados sugerem que os genes que codificam esses receptores estão menos sujeitos à seleção natural do que aqueles que codificam o maquinário de transdução de sinal. Essa plasticidade se reflete em uma variabilidade elevada, que permite aos quimiorreceptores conferirem essa adaptabilidade dinâmica aos organismos frente aos diferentes contextos ecológicos em que estão inseridos.

O índice estimado de plasticidade (EPI) é baseado na diversidade e abundância dos ortólogos de uma proteína ao longo da árvore evolutiva (DALMOLIN *et al*, 2011). No caso dos receptores TS1, encontramos um valor de EPI inesperadamente maior do que os outros receptores. Esse achado vem como uma surpresa, já que, aparentemente, os genes TS1 são bastante conservados em sequência e em número nos vertebrados – com poucas exceções, todos possuem apenas três receptores (NEI *et al*, 2008).

Analisando as proteínas do grupo de ortólogos dos receptores de TS1, foram encontrados membros incluídos em diversas funções não relacionadas com detecção de sabor (*e.g. Bride of sevenless*, um gene homeótico). Essa variabilidade funcional do grupo de ortologia pode ser a razão pela qual o EPI dos TS1 é mais alto do que o esperado para

essa classe de receptores, o que significa que o “espaço” ocupado por esses genes nas redes de interações seja capaz de preencher outros nichos funcionais além da recepção química dos sabores doces e *umami*.

Os mecanismos que permitem que percebamos uma gama tão grande de sabores com poucos receptores ainda são largamente desconhecidos (NELSON *et al*, 2001). Uma possibilidade alternativa para o alto EPI dos TS1 é o fato de que eles estão envolvidos na preferência alimentar, característica esta com alta variabilidade entre indivíduos e que, aparentemente, não está tão sujeita a seleção natural.

Outra situação que, apesar de compatível com os valores de EPI esperados, merece ser tratada em separado das demais é a dos receptores VN. A plasticidade encontrada para eles foi bastante elevada, o que está em concordância com as variações entre esses genes nos vertebrados, já apontadas em outros estudos (YOUNG *et al*, 2010). Entretanto, ressaltamos que o STRING classificou os grupos de ortologia desses receptores em sua maioria como *Non supervised Orthologous Groups* (NOGs), o que não nos permite fazer inferências com alto grau de segurança, já que estes grupos sofrem constantes modificações em sua ordenação.

Dentre os vários quimiorreceptores, os OL e os TS2 são os maiores responsáveis por trazer informações acerca do ambiente em que os organismos vivem. É de se esperar, portanto, que suas plasticidades sejam similares, refletindo essa função de percepção externa. Em acordo, eles possuem médias de EPI bastante consistentes entre si, um indício de que são regidos por princípios evolutivos semelhantes nos organismos.

Similarmente, os TS1 e VN foram considerados semelhantes em todos os organismos de nosso estudo. Embora pudéssemos dispensar esse achado em função da pouca robustez de dados disponíveis para os VN e a grande variabilidade do grupo de ortologia dos receptores gustatórios, é importante ressaltar que o intervalo de confiança encontrado nas análises estatísticas é grande, tornando necessário que esse aspecto seja estudado com mais profundidade no futuro.

Uma forma de interpretação desses dados é a seguinte: o grupo formado pelos OL e TS2 está mais associado com fatores relacionados diretamente a sobrevivência dos organismos do que os TS1 e VN, que determinam aspectos mais relacionados às preferências

alimentares e sociais. Esperaríamos, portanto, que a plasticidade dos primeiros fosse menor, pois a seleção natural é maior nos atributos que aumentam a aptidão evolutiva dos indivíduos.

5.2 Os parâmetros funcionais de redes dos quimiorreceptores

Estudos sugerem que a evolução se processa na periferia das redes biológicas, pois ali estão as proteínas menos estabelecidas funcionalmente e, portanto, menos sujeitas à seleção natural (ERWIN *et al*, 2009; WOLF, 2009). Os valores de conectividade e clusterização são atributos que nos permitem deduzir a construção de uma rede e determinar os seus limites, não só por identificar e agrupar nós e módulos funcionais, mas também por inferir a função e importância dos componentes na organização do sistema (KOONIN *et al*, 2010).

Os perfis encontrados de conectividade e clusterização para os receptores químicos foram significativamente diferentes dos encontrados para maquinário de transdução de sinal. Os quimiorreceptores apresentaram, de maneira geral, baixos valores de $k(i)$ e $c(i)$, sugerindo que, diferente dos membros do grupo acessório, esses receptores sejam proteínas de periferia, com módulos pouco estabelecidos e pouca comunicação entre seus membros. O grupo dos receptores olfatórios, entretanto, apresentou uma intensa clusterização entre seus membros, que formam um grande módulo funcional na rede. Embora os bancos de dados não sejam claros sobre a natureza dessas interações, esse achado poderia ser explicado por alguns estudos recentes, que indicam que a expressão dessas proteínas no epitélio olfatório principal é do tipo *one neuron-one receptor* – onde apenas um receptor é expresso por neurônio (SERIZAWA *et al*, 2004). O mecanismo que explicaria essa expressão ainda é desconhecido, mas acredita-se que envolva uma inibição ativa exercida pelo primeiro receptor transcrito ou traduzido sobre os demais.

Outro grupo que apresentou um índice alto de modularidade foi o dos TS1. Esses dados estão em acordo com a literatura, que postula que tanto os receptores de doce e *umami* devem necessariamente se associar ao terceiro receptor (que, aparentemente, não detecta nenhum tipo de sabor), para ativarem a proteína G e darem continuidade à cascata de transdução de sinal (NELSON *et al*, 2001).

No caso dos VN, foram encontrados valores próximos e iguais a zero de $k(i)$ e $c(i)$, o que está em acordo com a posição periférica que ocupam na rede biológica. Ressaltamos, contudo, que o STRING não indicou interações funcionais entre a maioria dos VN e a proteína *Trpc2*, o canal iônico responsável pela transdução do sinal feromonal (YOUNG *et al.*, 2010). Como esses receptores são pouco estudados, é possível que essas interações tenham sido negligenciadas nos bancos de dados e os valores de $k(i)$ e $c(i)$ subestimados.

A abordagem para evolução dos receptores químicos realizada neste estudo foi coerente com o modelo evolutivo proposto para essas famílias gênicas. Os parâmetros funcionais que encontramos para as diferentes classes de quimiorreceptores nos permitiram a modelagem de uma rede compatível com os dados da literatura. Podemos considerar, portanto, que as ferramentas da Biologia de Sistemas são uma complementação bastante útil às análises de evolução genômica, mesmo quando a quantidade de dados disponíveis não é tão grande.

6 Conclusões

6.1 Conclusão geral

Embora nem todos os dados encontrados estivessem perfeitamente em acordo com as informações disponíveis na literatura, podemos afirmar com segurança que a Biologia de Sistemas demonstrou ser uma ferramenta muito útil neste estudo. Ela foi capaz de identificar diferenças entre os quimiorreceptores e o maquinário de transdução de sinal, além de indicar possíveis objetos para estudos futuros, como os VN e os TS1. Esses dados contribuíram para gerar um panorama bastante coerente dos mecanismos evolutivos que regem a evolução da quimiorrecepção nos organismos estudados.

6.2 Conclusões específicas

6.2.1 Os parâmetros de redes são distintos entre os quimiorreceptores e o maquinário de transdução de sinal

Nosso estudo levantou os parâmetros funcionais das redes de proteínas dos quimiorreceptores e seu maquinário de transdução de sinal. Encontramos diferenças significativas entre os dois grupos, notavelmente entre os EPIs, o que sugere que os receptores ocupam uma posição mais periférica na sua rede funcional e, portanto, sofrem menos pressão evolutiva.

6.2.2 Os parâmetros de redes são distintos entre os próprios quimiorreceptores

Outro ponto abordado em nosso estudo foram as diferenças entre os próprios subgrupos de receptores. Verificamos que existem diferenças nos perfis de conectividade e clusterização das diferentes famílias. Os OL e TS1 apresentam altos índices $k(i)$ e $c(i)$, que estão relacionados às prováveis relações de inibição ativa de expressão/tradução entre os primeiros e a formação de complexos para a transdução de sinal nos últimos. Os EPIs dos OL e TS2 foram considerados estatisticamente semelhantes, assim como os dos VN e TS1, sugerindo que esses receptores possuem funções adaptativas próximas.

7 Referências Bibliográficas

ADLER E., Hoon, M. A., Mueller, K. L., Chandrashekar, J., Ryba, N. J., e Zuker, C. S. A novel family of mammalian taste receptors. *Cell*, v. 100, n. 6, p. 693-702, Mar, 2000.

ASHBURNER M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., e Sherlock, G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*, v. 25, n. 1, p. 25-29, May, 2000.

BARABASI A. L. & Oltvai, Z. N. Network biology: understanding the cell's functional organization. *Nat Rev Genet*, v. 5, n. 2, p. 101-113, Feb, 2004.

BERG H. C. Chemotaxis in bacteria. *Annu Rev Biophys Bioeng*, v. 4, n. 00, p. 119-136, 1975.

DALMOLIN R. J., Castro, M. A., Rybarczyk, J. F., Souza, L. H., de Almeida, R. M., e Moreira, J. C. Evolutionary plasticity determination by orthologous groups distribution. *Biol Direct*, v. 6, n. 1, p. 22-May, 2011.

ERWIN D. H. & Davidson, E. H. The evolution of hierarchical gene regulatory networks. *Nat Rev Genet*, v. 10, n. 2, p. 141-148, Feb, 2009.

FELDMESSER E., Olender, T., Khen, M., Yanai, I., Ophir, R., e Lancet, D. Widespread ectopic expression of olfactory receptor genes. *BMC Genomics*, v. 7, n. 121-2006.

HARRINGTON E. D., Jensen, L. J., e Bork, P. Predicting biological networks from genomic data. *FEBS Lett*, v. 582, n. 8, p. 1251-1258, Apr, 2008.

HARTMANN T. From waste products to ecochemicals: fifty years research of plant secondary metabolism. *Phytochemistry*, v. 68, n. 22-24, p. 2831-2846, Nov, 2007.

HASHIGUCHI Y. & Nishida, M. Evolution of trace amine associated receptor (TAAR) gene family in vertebrates: lineage-specific expansions and degradations of a second class

of vertebrate chemosensory receptors expressed in the olfactory epithelium. *Mol Biol Evol*, v. 24, n. 9, p. 2099-2107, Sep, 2007.

HARTWELL L. H., Hopfield, J. J., Leibler, S., e Murray, A. W. From molecular to modular cell biology. *Nature*, v. 402, n. 6761 Suppl, p. C47-C52, Dec, 1999.

HASIN-BRUMSHTEIN Y., Lancet, D., e Olender, T. Human olfaction: from genomic variation to phenotypic diversity. *Trends Genet*, v. 25, n. 4, p. 178-184, Apr, 2009.

JENSEN L. J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., Bork, P., e von, M. C. STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res*, v. 37, n. Database issue, p. D412-D416, Jan, 2009.

KOONIN E. V. & Wolf, Y. I. Constraints and plasticity in genome and molecular-phenome evolution. *Nat Rev Genet*, v. 11, n. 7, p. 487-498, Jul, 2010.

NEI M. & Rooney, A. P. Concerted and birth-and-death evolution of multigene families. *Annu Rev Genet*, v. 39, n. 121-152, 2005.

NEI M., Gu, X., e Sitnikova, T. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc Natl Acad Sci USA*, v. 94, n. 15, p. 7799-7806, Jul, 1997.

NEI M., Niimura, Y., e Nozawa, M. The evolution of animal chemosensory receptor gene repertoires: roles of chance and necessity. *Nat Rev Genet*, v. 9, n. 12, p. 951-963, Dec, 2008.

NELSON G., Hoon, M. A., Chandrashekar, J., Zhang, Y., Ryba, N. J., e Zuker, C. S. Mammalian sweet taste receptors. *Cell*, v. 106, n. 3, p. 381-390, Aug, 2001.

RODRIGUEZ I., Del, P. K., Rothman, A., Ishii, T., e Mombaerts, P. Multiple new and isolated families within the mouse superfamily of V1r vomeronasal receptors. *Nat Neurosci*, v. 5, n. 2, p. 134-140, Feb, 2002.

SERIZAWA S., Miyamichi, K., e Sakano, H. One neuron-one receptor rule in the mouse olfactory system. *Trends Genet*, v. 20, n. 12, p. 648-653, Dec, 2004.

SHI P. & Zhang, J. Extraordinary diversity of chemosensory receptor gene repertoires among vertebrates. *Results Probl Cell Differ*, v. 47, n. 1-23, 2009.

SWANEY W. T. & Keverne, E. B. The evolution of pheromonal communication. *Behav Brain Res*, v. 200, n. 2, p. 239-247, Jun, 2009.

WOLF Y. I., Novichkov, P. S., Karev, G. P., Koonin, E. V., e Lipman, D. J. The universal distribution of evolutionary rates of genes and distinct characteristics of eukaryotic genes of different apparent ages. *Proc Natl Acad Sci USA*, v. 106, n. 18, p. 7273-7280, May, 2009.

YAMADA T. & Bork, P. Evolution of biomolecular networks: lessons from metabolic and protein interactions. *Nat Rev Mol Cell Biol*, v. 10, n. 11, p. 791-803, Nov, 2009.

YOUNG J. M., Kambere, M., Trask, B. J., e Lane, R. P. Divergent V1R repertoires in five species: Amplification in rodents, decimation in primates, and a surprisingly small repertoire in dogs. *Genome Res*, v. 15, n. 2, p. 231-240, Feb, 2005.

YOUNG J. M., Massa, H. F., Hsu, L., e Trask, B. J. Extreme variability among mammalian V1R gene families. *Genome Res*, v. 20, n. 1, p. 10-18, Jan, 2010.