

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE BIOCÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR**

**EVOLUÇÃO DOS GENES DA REDE *OXT* - *AVP* - *PRL*:
ASPECTOS MOLECULARES, FISIOLÓGICOS E COMPORTAMENTAIS EM
MAMÍFEROS**

PAMELA LAIZ PARÉ DA ROSA

Tese submetida ao Programa de Pós-Graduação
em Genética e Biologia Molecular da UFRGS
como requisito parcial para a obtenção do Grau
de Doutor em Ciências.

Orientadora: Prof^a. Dr^a. Maria Cátira Bortolini

Co-orientador: Prof. Dr. Fabricio Rodrigues dos Santos

Porto Alegre, fevereiro de 2018

INSTITUIÇÕES E FONTES FINANCIADORAS

Os estudos foram realizados no Laboratório de Evolução Humana e Molecular (LEHM) do Departamento de Genética da Universidade Federal do Rio Grande do Sul e no Laboratório de Biodiversidade e Evolução Molecular (LBEM) do Departamento de Biologia Geral da Universidade Federal de Minas Gerais.

O financiamento foi subsidiado por recursos do PPGBM, da FAPEMIG e, a bolsa de doutorado foi concedida pelo CNPq.

“We are all virtually identical twins”.

Craig Venter 2002.

Dedico este trabalho e esses anos de estudo aos meus pais, que lutaram incansavelmente para meu crescimento em todos os âmbitos.

AGRADECIMENTOS

Primeiramente, agradeço a minha orientadora, a profa. Maria Cátira Bortolini, por seu inestimável apoio e suporte nessa jornada, tornando os dias mais alegres ao partilhar de café, biscoitos de chocolate e ótimas histórias. Muito obrigada pela confiança ao permitir o desenvolvimento de parte de meu doutorado em colaboração com o prof. Fabrício Rodrigues dos Santos na UFMG, ao qual também sou muito grata por todo apoio, disponibilização de recursos e suporte de sua equipe.

Àqueles que disponibilizaram recursos biológicos imprescindíveis para a execução da presente Tese: Prof. Alcides Pissinatti do Centro de Primatologia do Rio de Janeiro, que doou um número razoável de amostras de tecido de primatas ao LEHM – UFRGS. À curadora do Museu de Ciências e Tecnologia da PUC – MG, Cláudia Costa que, prontamente disponibilizou seu acervo de tecidos de Marsupiais para a realização especificamente desse trabalho e que continuarão a ser exploradas em pesquisas no LBEM – UFMG.

Ao Elmo, pela dedicação ao Programa PPGBM e prontidão nas tarefas do dia-a-dia, tornando tudo muito mais simples. Assim como aos colegas que partilharam de disciplinas, atividades, e, portanto, auxiliando em nossa construção pessoal e profissional.

Mais especificamente, aos colegas e professores que em algum momento ofereceram incentivo e / ou que serviram de exemplo a ser seguido nessa caminhada: Salzano, Enrique, Marcus, Luiz, Rafael, Danae, Caio, Virginia, Tábita, Gislene. Em especial à Vanessa Paixão-Côrtes e à Luciana Tovo, pessoas com quem tive a oportunidade de trabalhar mais diretamente e que foram muito generosas em partilhar conhecimento e habilidades.

Aos colegas do LEHM que têm caminhado junto: Lucas, Vanessa, Bibiana, Guilherme, Aline, Álex, Thiago, e, em especial ao Pedro com quem pude partilhar o vocabulário fronteiriço logo de chegada, além do gosto pela oxitocina.

Aos colegas do LBEM: Maycon, Flávia, Dominique, Davidson, Tomaz, Fernanda, Camila, José Eustáquio, Pedro, Rafael, e ao seu querido Pós doc sênior Prof. Romeu. Mas em especial à Larissa que, cheia de motivação e disponibilidade, me incluiu em seu laboratório e se tornou uma grande amiga, e ao gaúcho Jean com quem pude partilhar chimarrão e me sentir em casa em Belo Horizonte.

Àqueles de longa data ou atuais com quem já partilhei trabalho, amizade e é claro muitos cafés: Bruna, Gerson, Camila, Priscila, e Carla;

A meus familiares por toda compreensão, apoio e carinho;

Aos meus pais Vilmar e Janete por toda estrutura familiar e aconchego necessários para seguir em diante;

Aos meus queridos irmãos Paola e Bernardo, com quem partilho um amor e cumplicidade sem tamanho;

Ao meu companheiro Eduardo que tem crescido e amadurecido comigo desde a maioridade, e agora juntamente com nossa pequena Bianca;

Aos órgãos que financiaram essa pesquisa e a todos aqueles que a tornaram possível.

SUMÁRIO

CAPÍTULO I	INTRODUÇÃO.....	12
1.1	CONSIDERAÇÕES INICIAIS.....	13
1.2	ABORDAGEM DE GENES CANDIDATOS.....	21
1.3	SISTEMA OXITOCINÉRGICO.....	22
1.3.1	Os nonapeptídeos: oxitocina (OXT) e vasopressina (AVP).....	22
1.3.2	Receptores: O receptor da oxitocina (OXTR) e os receptores da vasopressina (AVPR1a, AVPR1b e AVPR2).....	24
1.4	SISTEMA DA PROLACTINA.....	26
1.4.1	Prolactina.....	26
1.4.2	O receptor da Prolactina (PRLR).....	27
1.4.3	O peptídeo liberador de PRL, PRLH.....	30
1.5	AMOSTRAGEM TAXONÔMICA.....	31
1.5.1	Ordem Primata.....	31
1.5.2	Marsupiais.....	35
CAPÍTULO II	JUSTIFICATIVA E OBJETIVOS.....	37
CAPÍTULO III	OXYTOCIN AND ARGININE VASOPRESSIN RECEPTOR EVOLUTION: IMPLICATIONS FOR ADAPTIVE NOVELTIES IN PLACENTAL MAMMALS.....	39
CAPÍTULO IV	MOLECULAR EVOLUTION OF THE PROLACTIN RECEPTOR (PRLR) AND IMPLICATIONS FOR ADAPTIVE TRAITS IN PRIMATES (MANUSCRITO EM PREPARAÇÃO: (PRIMEIRA VERSÃO).....	41
CAPÍTULO V	DADOS ADICIONAIS: GENE DO RECEPTOR DA PROLACTINA (PRLR) EVOLUÇÃO MOLECULAR EM MARSUPIAIS.....	75
CAPÍTULO VI	DADOS ADICIONAIS: VARIABILIDADE DO HORMÔNIO LIBERADOR DE PROLACTINA (PRLH) EM PRIMATAS DO NOVO MUNDO.	80
CONCLUSÃO	84
ANEXO I	99
ANEXO II	101
ANEXO III	103
ANEXO IV	105
ANEXO V	107

LISTA DE ABREVIATURAS, SÍMBOLOS E UNIDADES

OXT – Oxitocina
OXTR – Receptor da oxitocina
Pro⁸OXT – Oxitocina, Prolina 8
Val³Pro⁸OXT – Oxitocina, Valina 3- Prolina 8
AVP – Arginina vasopressina
AVPR1a – Receptor de arginina vasopressina 1a
GPCRs – Receptores acoplados a proteínas G
BARK – Quinase do Receptor Beta Adrenérgico
ESR1 – Receptor do Estrogênio 1
FOS – Proto-Oncogene FOS
HCRT – Neuropeptídeo precursor de hipocretina
PRL –Prolactina
PRLH – Hormônio liberador de prolactina
TRH – Hormônio liberador de tireotrofina
AVPR2 – Receptor de arginina vasopressina 2
DigiCGA – Abordagem digital de gene candidato
PVN – Núcleo paraventricular
SON - Núcleo supra - óptico
AVPR1b – Receptor de arginina vasopressina 1b
Ala⁸OXT – Oxitocina, alanina 8
Thr⁸OXT – Oxitocina, treonina 8
Phe²OXT – Oxitocina, fenilalanina 2
ACTH – Hormônio adrenocorticotrófico
GH – Hormônio do crescimento
PLs – Lactogênicos placentários
SFKs – Tirosina quinases não-receptoras
JAK – Janus quinases
STAT – Transdutor de sinal e Ativador de Transcrição
PI3K – Fosfatidilinositol-3-Quinase
AKT – Proteína kinase B
MAPK – Proteínas quinase ativadas por mitógenos
GPR10 – Receptor do peptídeo liberador de prolactina

RESUMO

A busca pelo repertório genético por trás de características comportamentais e reprodutivas de espécies de primatas tem desafiado nosso grupo de pesquisa. A premissa principal nesse tipo de estudo baseia-se na hipótese de que um traço fenotípico (seja ele fisiológico, comportamental, etc) compartilhado por um grupo taxonômico inteiro deve ser determinado por um repertório genético comum a esses táxons. Considerando a complexidade de muitas dessas características, temos ampliado nossos estudos para vários genes da rede OXT – AVP – PRL, usando abordagens *in silico*, *in vitro*, e *in vivo*. Na presente Tese, o conjunto gênico de estudo foi selecionado por uma metodologia baseada na ontologia biológica, com critério de seleção para características como comportamento materno, amamentação e aspectos reprodutivos, tais como comportamento de acasalamento e de corte. Essa seleção resultou em 12 genes candidatos para o estudo: *AVP*, *AVPR1A*, *AVPR1B*, *ESR1*, *FOS*, *HCRT*, *OXT*, *OXTR*, *PRL*, *PRLH*, *PRLR* e *TRH*. Exploramos aqui esse conjunto gênico da rede OXT – AVP – PRL através da mineração de dados, buscando por seus ortólogos nas várias espécies de primatas e de outros mamíferos, bem como através de novos dados de sequências da região codificadora dos genes *PRLR* e *PRLH*, em um conjunto de espécies de primatas do Novo Mundo (NWM, conforme sigla em inglês). Adicionalmente, sequências das regiões codificadoras de *PRLR* foram também obtidos em espécies de marsupiais. Com o objetivo de elucidar os padrões evolutivos dos genes de interesse, utilizamos análises como os testes NsSites e Branch Sites do pacote PAML, assim como vários testes populacionais clássicos, para diferentes conjuntos amostrais. Além disso, foi feita a predição da estrutura secundária das proteínas-alvo, utilizando metodologia específica do programa PSIPRED, bem como o PONDER-FIT, para a caracterização de aminoácidos intrinsecamente desordenados. Nossos resultados das análises *in silico* sugerem que os genes da família de receptores da vasopressina (*AVPR1A*, *AVPR1B*, e *AVPR2*) apresentam um padrão compatível com seleção positiva em mamíferos placentários. Alguns dos sítios com sinal de seleção apresentam motivos lineares (SLiMS) preditos no receptor *AVPR2*, que podem ter facilitado a emergência de novidades adaptativas, conforme foi sugerido para a espécie do rato-canguru *Dipodomys ordii*, que habita regiões áridas. As análises dos dados originais da região codificadora do gene *PRLR* em 17 espécies de NWM revelaram vários sítios presentes na forma longa do receptor com alta probabilidade de estarem sob seleção positiva, sendo que alguns deles (posições 507, 532 e 572) estão associados com o parto gemelar, uma característica das espécies de Callitrichidae. Adicionalmente verificamos no ramo dos Siimiformes um motivo linear de interação que reconhece domínios SH3 (*Src*

Homology 3). Os domínios SH3 e os seus locais de ligação foram descritos para centenas de proteínas; eles fornecem à célula um meio particularmente conveniente e adaptável de interação específica proteína-proteína, que pode ser de importância funcional. Esse trabalho como um todo contribuiu para o conhecimento do repertório genético relacionado à complexa rede de mecanismos neuroendócrinos associados à emergência de traços adaptativos, tanto fisiológicos quanto comportamentais em diferentes clados de mamíferos.

Palavras chaves: Evolução molecular, cuidado parental, primatas, marsupiais, OXT – AVP – PRL.

ABSTRACT

The search for the genetic repertoire behind behavioral and reproductive features of Primate species has challenged our research group. The principal premise in this kind of study is based on the hypothesis that a phenotypic trait (either physiological, behavioral, etc.) shared by an entire taxonomic group should be determined by a genetic repertoire common to these taxa. Considering the complexity of many of these features, we have expanded our studies for several genes of the OXT - AVP - PRL network, using *in silico*, *in vitro*, and *in vivo* approaches. In the present Thesis, the set of genes for the study was selected by a methodology based on biological ontology, with features like maternal behavior, breastfeeding, and reproductive aspects, such as mating and courtship behavior. This selection resulted in 12 candidate genes for the study: *AVP*, *AVPR1A*, *AVPR1B*, *ESR1*, *FOS*, *HCRT*, *OXT*, *OXTR*, *PRL*, *PRLH*, *PRLR*, and *TRH*. We explored here this gene set of the OXT – AVP – PRL network through data mining, searching for their orthologues in many Primate species and other mammals, as well as through new sequence data from the *PRLR* and *PRLH* coding region in a set of New World Monkey (NWM) species. Additionally, sequences from the *PRLR* coding regions were also obtained in marsupial species. To elucidate evolutionary patterns of the genes of interest, we used the NsSites and Branch Sites tests from PAML package, as well as several classic population tests, for different sample sets. In addition, we predicted the secondary structure of target proteins, using a specific methodology of the PSIPRED program, as well as PONDER-FIT for prediction of intrinsically disordered amino acids. Our *in silico* results suggest that the genes of the vasopressin receptor family (*AVPR1A*, *AVPR1B*, and *AVPR2*) present a pattern compatible with positive selection in placental mammals. Some of the sites with selection signals have linear motifs (SLiMS) predicted in the AVPR2 receptor, which may have facilitated the emergence of adaptive novelties, as was suggested for the kangaroo rat *Dipodomys ordii*, which inhabits arid regions. Analyses of the original *PRLR* coding region data on 17 NWM species revealed several sites present in the long form of the receptor with a high probability of being under positive selection, some of them (positions 507, 532 and 572) being associated with twin births, a characteristic of Callitrichidae species. Additionally, we verified in the Siimiformes branch a linear interaction motif that recognizes SH3 domains (Src Homology 3). The SH3 domains and their ligands were described for hundreds of proteins; they provide a particularly convenient and adaptable medium of specific protein-protein interaction to the cell, which can be of functional importance. This work as a whole contributed to the knowledge of the genetic repertoire

connected to the complex network of neuroendocrine mechanisms associated to the emergence of physiological and behavioral adaptive traits in different mammalian clades.

Keywords: Molecular evolution, parental care, primates, marsupials, OXT – AVP – PRL.

CAPÍTULO I *INTRODUÇÃO*

1.1 CONSIDERAÇÕES INICIAIS

Charles Darwin brindou a humanidade com sua obra “*On the Origin of Species*” publicada em 1859, introduzindo de maneira científica o conceito de evolução biológica. Tal obra em conjunto com estudos realizados pelo monge Gregor Mendel compõem a base para a Teoria Sintética da Evolução, também denominada de Neodarwinismo que emergiu no início do século XX. Esse período prolífico para a Ciência influenciou o primo de Charles Darwin, Francis Galton que desenvolveu estudos pioneiros de genética do comportamento humano, como a primeira publicação na área, intitulada: “*Hereditary talent and character*” em 1865 (GREEN, 2017). Galton também popularizou o conceito de “*Nature-Nurture*”, que destaca a influência do ambiente sobre o comportamento, no qual o termo *Nature* qualifica as características intrínsecas aos seres humanos, as quais são determinadas geneticamente, e *Nurture*, a influência de fatores externos (NEILL, 2004). Subsequentemente a esses conceitos do início do século XX, veio a percepção de que habilidades específicas, tais como arte e política, por exemplo, eram conspícuas em algumas famílias mais especificamente.

O surgimento da genética do comportamento em si foi dependente do desenvolvimento das disciplinas de Genética, Etologia e Psicologia. Como estratégia de estudo em humanos, a utilização de gêmeos monozigóticos e dizigóticos foi de grande valia, assim como a inclusão de casos de adoção por envolver os diferentes fatores, genéticos e ambientais. Uma grande contribuição foi realizada através desses estudos para o entendimento sobre a desordem de autismo, doença na qual a responsabilidade já foi atribuída ao comportamento dos pais, sugerindo que filhos não responsivos eram exclusivamente um resultado de uma suposta indiferença da mãe (hipótese da “mãe geladeira”). Esse tabu foi quebrado com a inclusão de gêmeos monozigóticos nos estudos, demonstrando que o risco de ambos exibirem o fenótipo de autismo era muito alto (60%), diferentemente do que ocorria quando os gêmeos eram dizigóticos (PLOMIN *et al.*, 2000).

Sendo assim, a genética do comportamento pode ser definida como: “*A ciência que utiliza os genes como ferramentas para auxiliar na compreensão do que nos faz comportar da maneira como comportamos*” (informação verbal)¹

¹ Definição realizada por Matt McGue no Curso “*Introduction to Human Behavioral Genetics*” fornecido pela Universidade de Minnesota e disponível *online*.

Tendo em vista a herança de caracteres complexos como o comportamento humano, nosso grupo de pesquisa (Laboratório de Evolução Humana e Molecular-LEHM) tem se dedicado a compreender as bases genéticas envolvidas no estabelecimento de alguns comportamentos que caracterizam certos grupos taxonômicos, como cuidado com a prole e comportamento de cuidado paterno mais especificamente. Utilizando-se de abordagens de genes candidatos (ver itens específicos mais abaixo nessa Tese), algumas famílias gênicas têm sido o foco de estudo de nosso grupo, destacando-se os achados pioneiros contidos na Dissertação de Mestrado do discente Carlos Menton Vieira, em que as variações moleculares no gene do receptor da oxitocina (*OXTR*) começaram a ser desvendadas em Platyrrhini (ou macacos do Novo Mundo - NWM; VIEIRA, 2012). O foco no gene do ligante (neuro-hormônio oxitocina; *OXT*), veio a partir do trabalho publicado por Lee *et al.*, (2011), em que cinco espécies de macacos do Novo Mundo, das 6 que foram investigadas pelos autores, apresentavam uma transição T->C no gene *OXT*, que resultava numa substituição não sinônima na cadeia de aminoácidos (Leu-> Pro). Esses resultados indicaram que mais estudos seriam necessários para um maior entendimento sobre a exata extensão da variabilidade do gene *OXT*.

Subsequentemente, nosso grupo de pesquisa e outros autores utilizaram dados sobre as variações do nonapeptídeo OXT e de seu receptor OXTR que revelaram ser táxon-específicas em Platyrrhini, associadas a fenótipos complexos tais como monogamia social, partos gemelares e machos (pai, irmão e mesmo não aparentados com o infante) que cuidam diretamente dos filhotes (anexo I; VARGAS-PINILLA *et al.*, 2015). Sinais de seleção positiva bem como coevolução entre variantes de OXT e OXTR foram descritas em nosso trabalho, assim como sua associação com partos gemelares. Tive a oportunidade de participar do estudo de Vargas-Pinilla *et al.* (2015), contribuindo com análises de dados e descrição dos resultados (ver anexo I). Como um todo, adicionalmente ao trabalho de Lee *et al.* (2011) esses estudos quebraram o paradigma da conservação da mesma sequência de aminoácidos de OXT em todos mamíferos placentários. Em outro recente estudo do grupo, investigamos a variação interespecífica no número predito de elementos de resposta à progesterona na região promotora OXTR, considerando 32 espécies de primatas. Ao correlacionar os achados genéticos com comportamento, sistemas sociais e parâmetros ecológicos / história de vida verificamos que os elementos de resposta à progesterona estavam apenas presentes em macacos do Novo Mundo e que seu número estava significativamente correlacionado com a presença de cuidados paternos nesse clado (VARGAS-PINILLA *et al.*, 2017; ver anexo II).

Dando continuidade aos estudos sobre a variabilidade de OXT e OXTR em primatas, realizamos a caracterização *in vitro* e *in vivo* de duas variantes táxon-específicas de OXT. Esse

trabalho resultou numa segunda publicação na área, na qual demonstramos que as variantes denominadas de Pro⁸OXT (encontrada em todo clado Cebidae) e Val³Pro⁸OXT (encontrada no gênero *Saguinus*) são agonistas igualmente eficientes, quando comparadas com a forma comum de OXT (Leu⁸OXT), na mediação de rotas dependentes da proteína G, mas que apresentam uma capacidade reduzida para o recrutamento de β -arrestinas, diminuindo assim ou mesmo impedindo a internalização de OXTR, levando a um possível impacto na dessensibilização de todo o sistema (ver anexo II; PARREIRAS-E-SILVA *et al.*, 2017). Val³Pro⁸OXT também mostrou recrutamento reduzido de β -arrestina após a interação com o receptor de AVP, AVPR1a. A capacidade de Pro⁸OXT e Val³Pro⁸OXT para estimular o cuidado paterno em ratos também foi mostrada no nosso estudo (PARREIRAS-E-SILVA *et al.*, 2017). Esses resultados apresentaram, pela primeira vez, um modelo natural para o conceito de agonismo com seletividade funcional (COSTA-NETO *et al.*, 2016), com prováveis implicações evolutivas (PARREIRAS-E-SILVA *et al.*, 2017).

Os achados envolvendo a família OXT/AVP revelaram variações específicas em AVPR1b, cuja informação para espécies de Platyrrhini não era conhecida. Variantes táxon-específicas, com possível implicação evolutiva e funcional, também foram descritas (ver anexo IV; Fam *et al.*, sob revisão). Dado o papel desempenhado por AVPR1b foi proposto que mudanças descritas para Platyrrhini, (classificação descrita por Perelman *et al.*, 2011) fazem parte de um repertório genético maior conectado à complexa rede de mecanismos neuroendócrinos associados ao eixo hipotálamo-pituitária-adrenal. Essas mudanças estão associadas ao surgimento de comportamentos vantajosos como cuidados paternos diretos em espécies de primatas socialmente monogâmicas, as quais também são caracterizadas por um pequeno tamanho corporal e partos gemelares.

Tanto no contexto histórico quanto científico, envolvendo as bases genéticas do comportamento, o processo de domesticação animal não poderia passar despercebido e será citado aqui, pois envolve abordagens e genes que dizem respeito a esta Tese.

Quando o *Homo sapiens* anatomicamente moderno ainda era uma espécie nômade e seus hábitos de vida e dieta eram de caçador-coletor, teve início a manipulação de algumas espécies selvagens para cruzamentos preferenciais entre indivíduos com características de interesse na domesticação. Assim, comportamentos específicos tais como docilidade, bem como características físicas desejadas começaram a ser selecionadas, através de cruzamentos seletivos, em certas espécies animais (SKOGLUND *et al.*, 2011). Essa seleção artificial tem resultado na emergência de características marcantes em relativamente curto espaço de tempo nas espécies domesticadas (os cães, suas raças com diferentes morfologias e tipos

comportamentais são um bom exemplo disso). Muitos estudos buscam desvendar as marcas no genoma desse processo seletivo. Nesse sentido, o nosso grupo de pesquisa também tem contribuído com achados na área, já que sugeriu pela primeira vez que o sistema OXT-AVP poderia também ter sido “selecionado” durante o processo de domesticação animal. Mais especificamente, foi proposto que os dois receptores de AVP, AVPR1b e AVPR2 estavam sob seleção positiva nos animais domesticados, em comparação com seus relativos selvagens. Por outro lado, o receptor AVPR1a foi encontrado sob relaxamento da pressão seletiva (ver anexo V desta Tese; Fam *et al.*, 2018, no prelo). Esses resultados obtidos com a minha participação, forneceram novos subsídios sobre a natureza do processo de domesticação com relação ao sistema OXT-AVP, o qual tem, como já salientado, um papel importante na mediação do comportamento, incluindo controle do estresse e agressividade, traços fundamentais no contexto da domesticação animal.

Além de OXT, AVP e seus receptores, outros genes da mesma rede ou de rotas correlacionadas também devem exercer papel fundamental em fenótipos de interesse em nossos estudos, tais como os genes das arrestinas/beta-arrestinas, visto que essas proteínas são elementos fundamentais no processo de dessensibilização das GPCRs (receptores acoplados à proteína-G), o que causa a redução específica da resposta celular a estímulos como hormônios, neurotransmissores ou sinais sensoriais. A arrestina beta 1 é uma proteína citosólica e age como um cofator na dessensibilização de receptores beta-adrenérgicos mediada pela Quinase do receptor beta adrenérgico (BARK). Além do sistema nervoso central, ela é expressa em níveis altos nos leucócitos sanguíneos periféricos, levando à hipótese de que o sistema BARK/beta-arrestina possui um papel de destaque na regulação mediada por receptor de funções imunes. Transcritos diferentes gerados por *splicing* alternativo codificando diferentes isoformas da arrestina beta 1 já foram descritos, mas suas funções continuam elusivas (COSTA-NETO *et al.*, 2016). A construção da rede funcional do sistema OXT-AVP, que indica quais genes fazem parte da mesma, também foi uma contribuição do Mestrado do discente Carlos Menton Vieira. Tal construção foi baseada na ontologia e buscou genes especificamente relacionados a fenótipos reprodutivos tais como comportamento materno e paterno, amamentação, comportamento de acasalamento e de corte. Resultou na identificação de 12 genes: *AVP*, *AVPR1a*, *AVPR1b*, *ESR1*, *FOS*, *HCRT*, *OXT*, *OXTR*, *PRL*, *PRLH*, *PRLR* e *TRH*; Figura 1; Tabela 1). Merecem destaque, devido aos objetivos da presente Tese, os genes do sistema oxitocinérgico (*OXT*, *OXTR*, *AVP*, *AVPR1a* e *AVPR1b*) e aqueles do sistema da prolactina (*PRL*, *PRLH* e *PRLR*).

Outro membro da família de receptores de OXT foi estudado aqui, AVPR2, o qual tem função fisiológica, de modo que não aparece na construção da rede acima. O estado de conservação desses 12 genes foi estudado em 36 espécies de vertebrados por meio da comparação das proteínas traduzidas com o seu homólogo humano. Foram identificados dois grupos gênicos de padrão de conservação significativamente diferentes: um grupo mais conservado, associado a OXT, e um grupo mais diversificado, relacionado à prolactina (PRL). Em termos de identidade, também há valores bastante diferenciados dentro de ramos específicos. Essa rede foi então utilizada para a seleção posterior de genes candidatos a serem explorados em futuros estudos do grupo. Como veremos nos itens abaixo, essa Tese é fruto dessa abordagem, considerando também a experiência e conhecimento sobre alguns genes da rede, como é o caso de OXT, AVP e seus receptores.

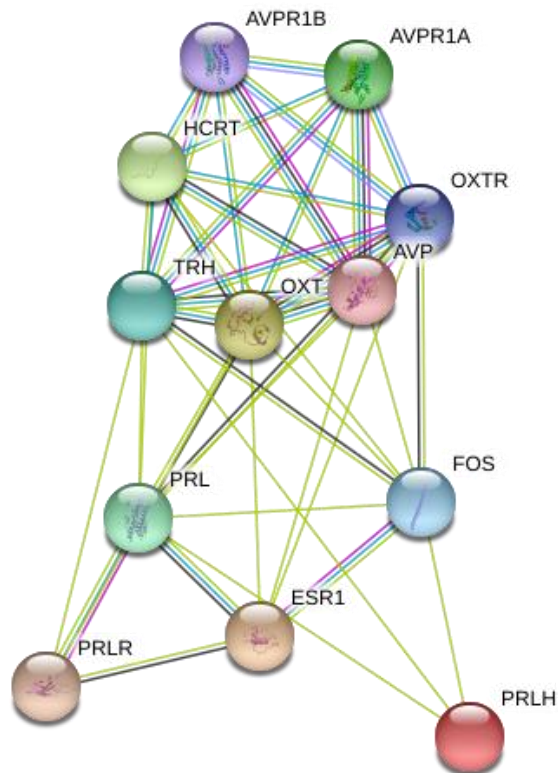


Figura 1- Rede funcional do sistema OXT-AVP-PRL gerada através de análise de cluster (VIEIRA, 2012). Imagem obtida no STRING: *functional protein association networks* (<https://string-db.org/>).

Tabela 1 – Descrição detalhada dos 12 genes da rede funcional da oxitocina e adicionalmente de *AVPR2*.

GENE	Função (Uniprot)	Éxons	pb	Cod. (nb)	Proteína (aa)	SNP (Cod; OMIM)	OMIM	C R O	Camundongo nocaut (MGI)	Estrutura (pb)
<i>AVP</i> <i>Arginina vasopressina</i>	A vasopressina tem uma ação antidiurética direta nos rins, mas também provoca vasoconstrição dos vasos periféricos.	3	2,169	619	164	rs121964882, rs121964884, rs121964886, rs121964888, rs121964890, rs121964892, rs121964893, rs121964883, rs121964885, rs121964887, rs121964889, rs121964891, rs74315383, rs28934878	125700 Diabetes insipidus, neurohipofisial	20	-	E1=170; E2=202; E3=247
<i>AVPR1a</i> <i>Receptor de arginina vasopressina 1a</i>	A atividade desse receptor é mediada pelas proteínas G, que ativam um sistema de segundo mensageiro fosfatidil-inositol-cálcio. Está envolvido em comportamentos sociais, incluindo filiação e apego.	2	7,574	1,525	418	-	-	12	Esquizofrenia	E1= 1.076; E2= 2.221
<i>AVPR1b</i> <i>Receptor de arginina vasopressina 1b</i>	A atividade desse receptor é mediada pelas proteínas G, que ativam um sistema de segundo mensageiro fosfatidil-inositol-cálcio.	2	7,664	2,237	424	-	-	1	Comportamental/ neuroológico	E1=1.400; E2=832
<i>AVPR2</i> <i>Receptor de arginina vasopressina 2</i>	A atividade desse receptor é mediada pelas proteínas G que ativam a adenilato-ciclase. Envolvidos na reabsorção renal de água.	3	4,636	1,622	371	rs104894748, rs104894750, rs104894751, rs104894753, rs104894755, rs104894757, rs104894759, rs104894761, rs104894749, rs28935496, rs104894752, rs104894754, rs104894756, rs104894758, rs104894760,	304800 - Diabetes insipidus, nefrogênica, 300539 - Síndrome nefrogênica de antidiurese inapropriada	X	-	E1=96; E2=885; E3=641

<p>ESRI <i>Receptor do estrogênio 1</i></p>	<p>Os hormônios esteroides e os seus receptores estão envolvidos na regulação da expressão de genes eucarióticos e afetam a proliferação celular e diferenciação do tecido alvo. Pode ativar a atividade transcricional de TFF1.</p>	10	472.929	6.466	595	rs104893956	133430 Resistência a estrogênio, 114480 Câncer de mama, 608446 - Suscetibilidade a infarto de	6	-	E1=169; E2=131; E3=522; E4=191; E5=117; E6=336; E7=139; E8=134; E9=184; E10=4.543
<p>FOS <i>Proto-Oncogene</i></p>	<p>Fosfoproteína nuclear que forma um complexo ligado, mas não covalentemente com o fator de transcrição JUN/AP-1. Nos heterodímeros, FOS e JUN/AP-1 cada região básica parece interagir com locais simétricos em meio aos sítios no DNA. Por ativação de TGF-beta, forma um complexo SMAD3/SMAD4/JUN/FOS complexos multiméricos no local AP1/SMAD-ligação para regular a sinalização de TGF-beta-mediada. Tem uma função fundamental na regulação do desenvolvimento de células destinadas a formar e manter o esqueleto. Acredita-se que tem um papel importante na transdução de sinal, proliferação, e diferenciação celular. Em células de crescimento, ativa a síntese de fosfolipídios, possivelmente ativando CDS1 e PI4K2A. Essa atividade requer Tyr-desfosforilação e associação com o retículo endoplasmático.</p>	4	3.461	2.158	380	-	-	14	-	E1 =350; E2=252; E3=108; E4=1.448
<p>HCRT <i>Neuropeptídeo precursor de hypocretina (orexina)</i></p>	<p>Os neuropeptídeos que desempenham um papel importante na regulação da ingestão de alimentos e de sono-vigília, possivelmente através da coordenação das respostas comportamentais e fisiológicas complexas dessas funções homeostáticas complementares. Um papel mais amplo na regulação homeostática do metabolismo energético, a função autônoma, o equilíbrio hormonal e a regulação dos fluidos corporais, também é sugerida. Orexina-A se liga a ambos OX1R e OX2R com uma elevada afinidade, enquanto que a orexina-B liga-se apenas a OX2R com uma afinidade elevada semelhante.</p>	2	1.393	577	131	rs13046884	161400- Narcolepsia tipo 1	17	Comportamental/ neurológico	E1=108; E2=469

OXT	<i>Oxitocina</i>	A oxitocina provoca a contração do músculo liso do útero e da glândula mamária.	3	898	513	125	-	-	20	-	E1=156 E2=202 E3=155
OXR	<i>Receptor da oxitocina</i>	A atividade desse receptor é mediada pelas proteínas G, que ativam um sistema de segundo mensageiro fosfatidilinositol-cálcico.	4	19221	4.364	389	-	-	3	-	E1=387 E2=96 E3=1.064 E4=2.817
PRL	<i>Prolactina</i>	Prolactina atua principalmente na glândula mamária ao promover a lactação	5	15608	1.354	227	-	-	6	-	E1=547 E2=176 E3=108 E4=180 E5=343
PRH	<i>Hormônio liberador de prolactina</i>	Estimulador de liberação de prolactina (PRL) e regulador da expressão de prolactina através do seu receptor GPR10. Pode estimular Lactotrofos diretamente a secretar PRL.	2	602	264	87	-	-	2	-	E1=100; E2=164
PRLR	<i>Receptor da prolactina</i>	Esse é um receptor para o hormônio prolactina da pituitária anterior. As isoformas 4 e 6 são incapazes de transduzir sinalização.	9	181.393	11.680	622	-	-	5	-	E1=322; E2= 62; E3= 113; E4= 133; E5=170; E6= 170;; E7=142; E8= 100; E9= 70; E10=10,406
TRH	<i>Hormônio liberador de tirotrófina</i>	Funciona como um regulador da síntese de TSH na glândula pituitária anterior e como neurotransmissor/neuromodulador no sistema nervoso central e periférico. Pode promover o alongamento do eixo do cabelo, prolonga a fase de crescimento do ciclo do cabelo (anágena) e antagoniza a sua rescisão por TGFB2. Também pode aumentar a proliferação e inibir a apoptose da matriz de queratinócitos do cabelo.	2	3,668	1.978	242	-	-	3	-	E1=519; E2=219; E3=1.240

pb = pares de base; Cod= região codificadora; aa = aminoácidos; SNP= *Single nucleotide polymorphism*; OMIM= *Online Mendelian Inheritance in Man*; CRO= cromossomo; E1= Éxon 1; E2= Éxon 2; E3= Éxon 3; E4= Éxon 4; E5= Éxon 5; E6= Éxon 6; E7=Éxon 7; E8= Éxon 8; E9= Éxon 9; E10= Éxon 10; Dados coletados do Ensemble (<http://www.ensembl.org/>); Gene Cards (<http://www.genecards.org/>); e OMIM (<http://www.http://omim.org/>).

1.2 ABORDAGEM DE GENES CANDIDATOS

Investigações com genes candidatos são naturalmente mais econômicas quanto comparadas a estudos genômicos ou exômicos, pois apresentam uma abordagem direta, na medida que partem da investigação de um ou mais genes de uma determinada rede, com funções e implicações fenotípicas previamente já bem estabelecidas. Um dos maiores desafios dessa abordagem é desvendar a funcionalidade de uma dada variante, bem como sua conexão com o traço sob investigação. Para isso, é fundamental buscar compreender os processos fisiológicos, bioquímicos e/ou vias metabólicas envolvidos com a(s) variante(s) (ZHU & ZHAO, 2007). Vale lembrar que a expressão de traços complexos depende de um “*repertório genético*”, e esse, pode apresentar um padrão de herança tanto poligênico quanto monogênico (NADEAU, 2001). Enfim, espera-se que um conjunto de genes, redes e rotas estejam envolvidos, cada qual com seu efeito, atuando sobre a manifestação de um determinado fenótipo.

A abordagem de gene candidato pode ser definida conforme segue abaixo:

The candidate-gene approach can be defined as the study of the genetic influences on a complex trait by: generating hypotheses about, and identifying candidate genes that might have a role in, the aetiology of the disease; identifying variants in or near those genes that might either cause a change in the protein or its expression, or be in linkage disequilibrium with functional changes; genotyping the variants in a population; and by using statistical methods to determine whether there is a correlation between those variants and the phenotype (TABOR et al., 2002).

Como pode ser visto, essa definição usando a estratégia de genes candidatos foi delineada para contemplar a variabilidade normal e patológica dentro da nossa espécie. Porém, também pode ser ampliada em nível taxonômico, já que fenótipos específicos, incluindo de natureza complexa, também caracterizam espécies, gêneros, etc.

Nos últimos anos, o progresso no conhecimento sobre os genomas de mamíferos tem disponibilizado um amplo conjunto de informações o que propiciou um avanço na identificação de genes candidatos (chamadas abordagens *in silico*). Essas abordagens, por sua vez, também podem ser denominadas mais especificamente de “*Digital Candidate Gene Approach*” (DigiCGA; DUNCAN & KELLER, 2011). Essas são baseadas na ontologia biológica e envolvem também diferentes e complexas abordagens de análises tais como: cadeias de Markov, análises de cluster, aprendizagem de máquina, dentre outras (ZHU & ZHAO, 2007 e referências lá contidas). Tal abordagem vem sendo utilizada com sucesso por nosso grupo de pesquisa para identificar variantes táxon-específicas associadas a fenótipos táxon-específicos,

como comentado já no item 1.1 Considerações Iniciais, e como será visto em maiores detalhes ao longo desta Tese.

1.3 SISTEMA OXITOCINÉRGICO

1.3.1 Os nonapeptídeos: oxitocina (OXT) e vasopressina (AVP)

Nas últimas duas décadas, vem sendo obtidos um volume maior de informações sobre a oxitocina (OXT; Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly; LEE *et al.*, 2009), bem como sobre seu parálogo a vasopressina (AVP; Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly), e suas conexões com traços comportamentais. Vários estudos demonstraram que esses dois hormônios estariam conectados com doenças psiquiátricas tais como o autismo e a esquizofrenia, o que indica que também atuam como neurotransmissores (VAN ANDERS *et al.*, 2011; HIGASHIDA *et al.*, 2012; MCCALL & SINGER, 2012).

Ambos nonapeptídeos assumem uma conformação cíclica em sua versão funcional devido à formação de pontes bissulfeto entre os resíduos de cisteína 1 e 6 (BARBERIS *et al.*, 1998). O papel fisiológico de OXT no organismo data de 1906, em que foi descrito o efeito da hipófise na contração uterina, ao mostrar a persistência de uma resposta do miométrio à sua administração endovenosa (DALE, 1906). O nome OXT deriva do grego e significa “nascimento rápido”, fazendo relação a seu efeito na indução do parto. Posteriormente foi sugerido seu papel na ejeção do leite materno durante a amamentação. A proteína AVP, proximamente relacionada à OXT, atua na homeostase de fluídos corporais tem seu nome derivando do latim, fazendo referência a seu efeito vasopressor (YOUNG & FLANAGAN-CATO, 2012).

Todos os membros de neuropeptídeos da família OXT e AVP apresentam um anel N-terminal de 6 resíduos seguido de uma cauda flexível C-terminal de três resíduos (KOEHBACH *et al.*, 2013), possibilitando a interação de OXT e AVP com seus receptores (HOYLE, 1998). Esses neuropeptídios, que diferem somente em duas posições na sua sequência de aminoácidos, estão presentes nos mamíferos placentários, com algumas variações (KOEHBACH *et al.*, 2013).

Considerando a sequência de aminoácidos desse nonapeptídeo a posição 8 é determinante para distingui-los se pertencentes à família da oxitocina ou da vasopressina, respectivamente (GIMPL & FAHRENHOLZ, 2001). A sequência do peptídeo AVP também é moderadamente conservada, sendo descritas algumas poucas variações não sinônimas, mais especificamente em *Sus scrofa* e *Monodelphis domestica* (Arg->Lys 8; FERGUSON, 1969; WALLIS, 2012). Ambos os genes codificantes para OXT e AVP estão localizados no braço curto do cromossomo 20 e são transcritos em direções opostas, sendo compostos por 3 éxons e 2 íntrons (Figura 2; LEE *et al.*, 2009).

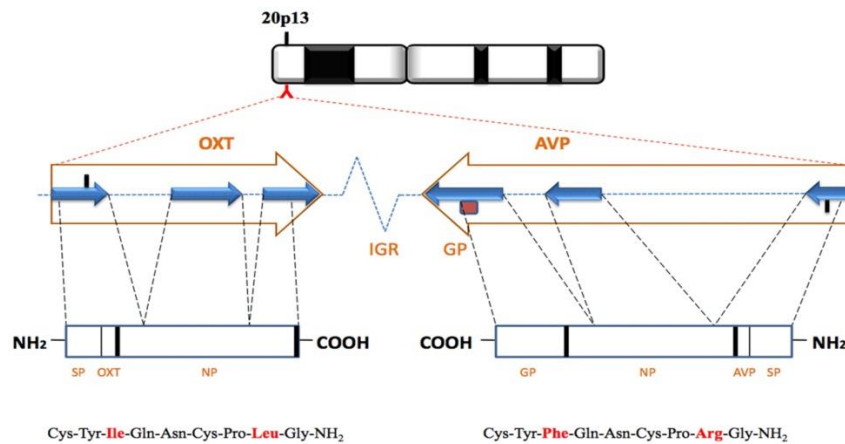


FIGURA 2 - Diagrama esquemático dos genes OXT e AVP (setas grandes), pré-pró-hormônios (caixas) e os nonapeptídeos (sequência de aminoácidos abaixo). Acima é exibida a localização cromossômica em humanos. As setas azuis pequenas representam os éxons e as linhas pontilhadas representam os íntrons. IGR, região intergênica; SP, peptídeo sinal; NP, neurofina; GP, glicopeptídeo. Os aminoácidos que diferem entre os peptídeos estão representados em vermelho. Fonte: Lee *et al.* (2009).

A principal fonte de OXT e AVP é o hipotálamo, de onde são liberados através dos neurônios magnocelulares dos Núcleos Supra-Óptico e Paraventricular (PVN e SON) e pelos neurônios parvocelulares do PVN. A sua atuação no sistema endócrino como hormônio e/ou neuromodulador é dependente do tipo de neurônio onde é expresso e liberado. Quando expressos nos neurônios magnocelulares, OXT e AVP são liberados na corrente sanguínea e atuam como hormônios, e quando derivam dos neurônios parvocelulares são direcionados a diferentes áreas do cérebro e atuam como neuromoduladores (GABOR *et al.*, 2012). No entanto, outras regiões do cérebro, bem como órgãos periféricos, também são fontes de OXT e AVP (GWEE *et al.*, 2009).

Invertebrados possuem um nonapeptídeo pertencente à família da oxitocina/vasopressina. Enquanto que todos os vertebrados possuem um nonapeptídeo

pertencente à família da oxitocina (isotocina/mesotocina/oxitocina) e da vasopressina (vasotocina/vasopressina), com exceção da lampreia que possui somente uma forma de vasopressina (GWEE *et al.*, 2009; GRUBER, 2014). A isotocina é comum nos peixes ósseos e a mesotocina aparece nos répteis e aves, além de alguns marsupiais.

Acredita-se que esses nonapeptídeos se originaram à cerca de 600 milhões de anos (ACHER *et al.*, 1995; GRUBER, 2014). Assim que, todas as formas de nonapeptídeos dos vertebrados teriam surgido por duplicação em tandem de seu ancestral vasotocina (GRUBER, 2014). A vasotocina está presente na lampréia, um peixe sem mandíbula que descende de uma linhagem irmã dos peixes mandibulados e demais vertebrados terrestres (GIMPL & FAHRENHOLZ, 2001), assim como no anfioxo, um cordado basal utilizado na compreensão de aspectos evolutivos dos vertebrados (GWEE *et al.*, 2009). Resultados de sintenia indicam uma conservação na lampreia, no tubarão, no celacanto e nos tetrápodes, tendo ocorrido uma inversão gênica provavelmente no genoma do ancestral comum entre primatas e roedores, enquanto que, os peixes teleósteos sofreram outros rearranjos, que foram provavelmente facilitados por eventos de duplicação no genoma do seu ancestral comum (CHRISTOFFELS *et al.*, 2004).

A sequência de aminoácidos descrita no início desse item para a oxitocina e a vasopressina são as formas mais comumente encontradas nos mamíferos placentários, embora variações tenham sido descritas (GIMPL & FAHRENHOLZ, 2001; STOOP, 2012), incluindo aquelas que caracterizam variantes de OXT em clados de macacos do Novo Mundo (Pro⁸OXT, Val³Pro⁸OXT Ala⁸OXT, Thr⁸OXT, Phe²OXT) descritas por nosso grupo de pesquisa (anexo I, VARGAS-PINILLA *et al.*, 2015) e outros autores (LEE *et al.*, 2011; REN *et al.*, 2015). Curiosamente, os grupos de macacos caracterizados por essas variantes apresentam a mesma sequência de AVP.

1.3.2 Receptores: O receptor da oxitocina (OXTR) e os receptores da vasopressina (AVPR1a, AVPR1b e AVPR2).

A mediação de comportamentos sociais complexos, bem como de ações fisiológicas dos nonapeptídeos OXT e AVP se dá através de sua interação com seus receptores OXTR e AVPR1a, AVPR1b e AVPR2, respectivamente. São membros da classe 1 da família dos receptores acoplados as proteínas G, que compõem a superfície de todos os tipos celulares, sendo responsáveis por transmitir sinais para dentro da célula, e atuando em processos

fisiológicos chave que se constituem num dos maiores alvos de fármacos da atualidade (SIEHLER & MILLIGAN, 2011). Esses receptores são conhecidos por apresentar 7 domínios transmembrana e estão acoplados com proteínas G para realizar a sinalização intracelular. Além disso, caracteristicamente, são observadas 4 alças extracelulares e 4 alças intracelulares em que a formação de uma ponte bissulfeto entre os resíduos de cisteína dos domínios extracelulares 2 e 3 são importantes para a conformação e estabilização dessas regiões transmembrana (BOCKAERT & PIN, 1999).

O gene codificante para OXTR está localizado no cromossomo 3 em humanos, sendo composto por 4 éxons, em que somente os dois últimos são codificantes. O receptor é expresso no sistema nervoso central e em órgãos periféricos, tais como: glândulas mamárias e útero (GIMPL & FAHRENHOLZ, 2001). Os genes codificantes para os receptores da vasopressina (AVPR1a, AVPR1b e AVPR2) estão localizados em diferentes cromossomos: 12, 1 e X, respectivamente. Os receptores da vasopressina possuem a mesma estrutura, 4 éxons com apenas 2 codificantes para a proteína (ENSEMBLE, <http://www.ensembl.org>). Os receptores AVPR1a e AVPR1b são expressos no sistema nervoso central e estão envolvidos na modulação do comportamento, além de mediar importantes ações fisiológicas (GABOR *et al.*, 2012). A vasoconstrição é mediada por AVPR1a quando expresso nas paredes dos vasos sanguíneos, enquanto que a liberação do hormônio adrenocorticotrófico (ACTH) é mediada via AVPR1b, sendo desse modo elemento chave na regulação do estresse. O AVPR2 é o único receptor a ser expresso somente nos rins e é importante para o balanço hídrico do organismo (KOSHIMIZU *et al.*, 2012).

O cenário sobre a história evolutiva dessa família de receptores, por outro lado, foi proposto por Lagman *et al.*, (2013), que através de análises filogenéticas e de sintenia indicaram que AVPR1a, AVPR1b e OXTR se originaram de um ancestral comum e que AVPR2 teve uma origem paralela (LAGMAN *et al.*, 2013). Já Yamashita & Kitano, (2013) não encontraram uma clara divergência entre os receptores.

Sendo assim, nesta Tese pretendemos investigar os padrões evolutivos dos receptores OXTR, AVPR1a, AVPR1b, e AVPR2.

1.4 SISTEMA DA PROLACTINA

1.4.1 Prolactina

A prolactina (PRL) pertence à família somatotropina/prolactina e umas de suas principais funções descritas é promover a lactação (BOLE-FEYSOT *et al.*, 1998). Atua num amplo espectro de funções, incluindo a mediação do comportamento materno (LUCAS *et al.*, 1998), e paterno (FLEMING *et al.*, 2002; PATIL *et al.*, 2014). Além disso, atua mais especificamente na reprodução, gravidez e lactação, crescimento e desenvolvimento, metabolismo, modulação imunológica, transporte de eletrólitos, regulação do integumento e oncogênese (BEN-JONATHAN & HUGO, 2015). A prolactina age como um sinalizador sendo produzida principalmente por células do anterior da pituitária, chamadas de lactotrofos, células as quais podem sofrer hiperplasia durante a amamentação (KUCKA *et al.*, 2013). Contudo, sabe-se que a prolactina produzida durante a amamentação também ocorre de maneira extra hipotalâmica, em ambos os compartimentos, estromal e glandular das mamas, onde provavelmente possui funções importantes, mas a descrição dessas atividades ainda não está bem clara (HORSEMAN & GREGERSON, 2014). Enquanto que, no tecido adiposo, sua expressão é três vezes maior que nos tecidos glandulares mamários (MARANO & BEN-JONATHAN, 2014). A produção extra hipotalâmica de PRL também ocorre nos folículos capilares em momentos específicos: quando o bulbo capilar está em plena atividade (fase anágena) e na fase de repouso capilar (fase telógena; MARANO E BEN-JONATHAN, 2014).

Esse hormônio provavelmente surgiu há mais de 400 milhões de anos, assim que, pode ser observado em insetos, peixes ósseos e tetrápodes (ADÁN *et al.*, 2014 e referências lá contidas). Embora em invertebrados o papel fisiológico de PRL não seja bem estabelecido, nos vertebrados, sua associação mais comum é com funções reprodutivas pós-parto, de modo que estaria relacionada a uma vasta gama de estratégias reprodutivas no Reino Animal (HORSEMAN & GREGERSON, 2014). Aparentemente, PRL possui um gene ancestral comum com o hormônio do crescimento (GH, *Growth Hormone* conforme a sigla em inglês), bem como com os lactogênicos placentários (PLs; COOKE *et al.*, 1981). No entanto, PLs têm uma história recente na cronologia evolutiva quando comparado aos demais, surgindo no ramo dos mamíferos placentários. Os ciclóstomos (lampreias etc), por sua vez, possuem uma outra forma do hormônio, a somatolactina (AMEMIYA *et al.*, 1999).

O gene codificante para a prolactina (*PRL*) está localizado no braço curto do cromossomo 6 em humanos e está constituído por 5 éxons e 4 íntrons, onde o primeiro éxon codifica somente parte do peptídeo sinal e os outros 4 éxons codificam o hormônio PRL (AKEN *et al.*, 2016). Quando comparada com os outros hormônios acima listados, a prolactina é a mais extensa, composta de 199 aminoácidos e algumas características únicas, como a glicosilação. Todos hormônios (PRL, GH, e PLs) são sintetizados como pré-pró-hormônios e estão constituídos de um peptídeo sinal e um neuropeptídeo maduro (UniProt Consortium, 2010), mas transcritos de diferentes tamanhos são produzidos através de processamento alternativo do RNA mensageiro (TRUONG *et al.*, 1984).

Estudos com camundongos nocautes para *PRL* têm produzido fêmeas com deficiência no desenvolvimento de glândulas mamárias e totalmente inférteis, enquanto que os machos em sua maioria são férteis. Através da interação com vias de sinalização e fatores de transcrição, tal como EIF5, tem sido demonstrado que PRL também atua na diferenciação e proliferação epitelial (HORSEMAN & GREGERSON, 2014).

Diferentes abordagens vêm sendo usadas para investigar o papel de PRL além das funções reprodutivas, já que estágios reprodutivos particulares, como a lactação, impõem muitas exigências à fêmea. Por exemplo, estudos que buscam desvendar susceptibilidades genéticas a doenças psiquiátricas em humanos têm demonstrado que a prolactina, juntamente com a somastatina e o receptor da dopamina D2, desempenham um papel importante no mecanismo de resiliência após o estresse, bem como após o tratamento com antidepressivos (FARON-GÓRECKA *et al.*, 2013). Outros autores mostraram que drogas administradas em mulheres que realizam tratamento psicológico acarretam um aumento dos níveis de PRL (hiperprolactinemia), e esse aumento pode ter consequências fisiológicas que incluem a amenorreia e oligomenorreia (BARGIOTA *et al.*, 2013). Por outro lado, hormônios incluindo PRL protegem o cérebro feminino de danos durante a gestação, pois atuam minimizando a resposta imunológica que causa inflamação (MONASTERIO *et al.*, 2013).

1.4.2 O receptor da Prolactina (*PRLR*)

A ação de PRL depende de sua interação com seu receptor, *PRLR*, o qual pertence à família tipo I dos receptores de citocinas (PATIL *et al.*, 2014). O gene que codifica *PRLR* está localizado no braço curto do cromossomo 5 em humanos e está constituído por 10 éxons, sendo que os dois primeiros não são codificantes para a proteína (Figura 3). Essa família de proteínas

está composta por uma região extracelular (EC), uma região transmembrana (TM) e uma região intracelular (IC), atuando através da formação de homodímeros (GOFFIN & KELLY, 1997). Importantes grupos funcionais desse receptor são éxon-específicos: como os resíduos de cisteína localizadas nos éxons 4 e 5, onde também é codificado o primeiro domínio de fibronectina (FBN) tipo III, o motivo WSxWS encontrado no éxon 7, onde também é codificado o domínio FBN tipo III. A região transmembrana é codificada inteiramente pelo éxon 8, enquanto que o motivo rico em prolina está codificado no éxon 9. Já no éxon 10 está contido o box 2 e vários resíduos essenciais para a transdução de sinal (ORMANDY *et al.*, 1998).

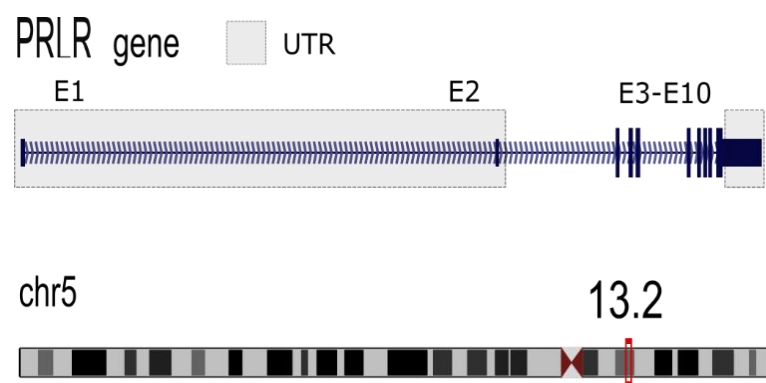


Figura 3 - Localização cromossômica do receptor da prolactina e esquema do gene codificante PRLR. Chr: cromossomo; E: Éxon; UTR: região não traduzida.

Segundo as informações disponíveis no UniProt, existem 9 isoformas de PRLR descritas, que são produzidas através do processamento alternativo do RNA, das quais as isoformas 4 e 6 são incapazes de traduzir a sinalização via interação com PRL. Essas isoformas diferem principalmente quanto ao tamanho de sua região intracelular, sendo caracterizadas basicamente como: isoforma curta, intermediária e longa (BOLE-FEYSOT *et al.*, 1998 e referências lá contidas; Figura 4). Conforme pode ser observado na figura 4, a forma longa expressa o receptor inteiramente e, portanto, possui características únicas. Por exemplo, é a única capaz de estimular a transcrição de genes da proteína do leite (LESUEUR *et al.*, 1991).

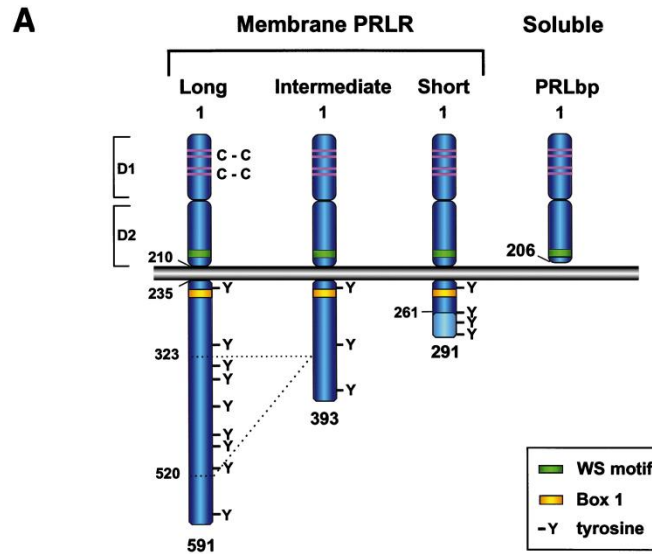


Figura 4- Esquema mostrando a classificação geral das isoformas de PRLR (longa, intermediária e curta) e suas principais distinções (BOLE-FEYSOT *et al.*, 1998).

Funcionalmente, PRLR não possui capacidade enzimática intrínseca e está constitutivamente associado à Janus quinase 2 (JAK2) e a uma proteína da família das Src tirosina-quinases (SFKs). Essas interações são necessárias para que ocorra a fosforilação de PRLR, assim como de outras proteínas da via (CLEVINGER & MEDAGLIA, 1994; DUSANTER-FOURT *et al.*, 1994). A forma longa de PRLR ativa uma ampla cascata de sinalização em que JAK2/STAT é a maior delas, além das vias das Src *quinases*, (PI3K) /AKT, e MAPK (Figura 5). A via JAK2/STAT atua na regulação de muitos aspectos críticos do crescimento celular, sobrevivência e diferenciação sendo que erros nessa via são frequentemente observados em muitos tumores humanos, refletindo a sua importância para a manutenção e integridade celular. A proteína JAK ativada leva à fosforilação de resíduos de tirosina nos receptores de citocinas com subsequente recrutamento e fosforilação de proteínas STAT, que resultam na sinalização para a transcrição de outras importantes proteínas (Mitchell & John, 2005). Já as proteínas da família SRC são quinases que também modulam a sinalização de outras proteínas, sendo compostas por uma região N-terminal seguida pelos domínios SH3 e SH2 e uma curta região C-terminal (BOGGON & ECK, 2004).

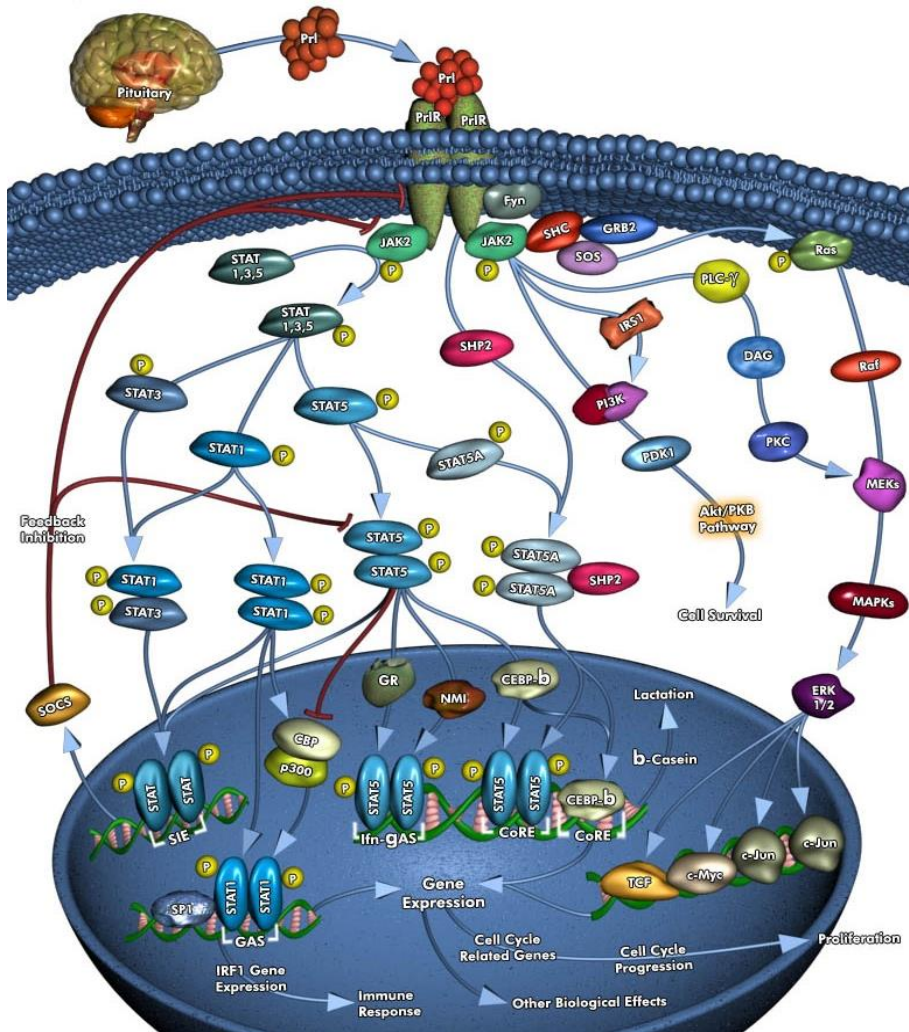


Figura 5 – Sinalização da prolactina através de diferentes vias obtida no site da Qiagen (<https://www.qiagen.com>).

1.4.3 O peptídeo liberador de PRL, PRLH

O peptídeo PRLH estimula a liberação de PRL através de sua interação com o receptor GPR10 (*G-protein coupled receptor 10*), regulando então, a expressão de PRL (LANGMEAD *et al.*, 2000). No entanto, o maior controlador hipotalâmico de liberação de PRL é a dopamina (TAYLOR & SAMSON, 2001).

O gene que codifica PRLH está localizado no braço longo do cromossomo 2 em humanos e é composto por 2 éxons, produzindo uma proteína de 87 aminoácidos em sua forma canônica (ZERBINO *et al.*, 2017). Atua na lactação, na regulação de ingestão alimentar, processamento de sinais de dor, assim como homeostase metabólica, resposta ao estresse, regulação cardiovascular, secreção de gonadotrofinas, secreção de GH (hormônio do crescimento) e regulação do sono (SUNET *al.*, 2005). Entre as funções também mediadas por

PRLH, inclui-se o seu envolvimento no controle de secreção de OXT, assim como do hormônio liberador de corticotropina e somatostatina (IBATA *et al.*, 2000). Maiores quantidades de PRLH são expressas na medula e no hipotálamo, contudo, esse hormônio é expresso em vários tecidos indicando uma ampla gama de funções (FUJII *et al.*, 1999).

1.5 AMOSTRAGEM TAXONÔMICA

Aqui serão investigadas diferentes espécies de mamíferos considerando vários genes da rede estendida de OXT-AVP, denominada na presente Tese de rede OXT-AVP-PRL. Como já mencionado, a abordagem de genes candidatos se faz presente, bem como outras que envolvem, por exemplo, a geração de dados originais de sequências a partir da identificação de regiões de potencial interesse funcional e evolutivo, reveladas por análises iniciais em bancos de dados públicos. Conforme mostra a Figura 6, um número considerável de espécies de mamíferos foi acessado. Contudo, o conjunto de táxons variou para os diferentes genes estudados. Vale destacar, no entanto, que a ordem Primates, foco constante de estudos dos quais participei e que já se encontram publicados, no prelo, ou sob revisão (ver anexos I, II, III e IV, V), terão destaque na presente Tese.

1.5.1 Ordem Primates

Os primatas se originaram no final do Cretáceo, juntamente com uma grande quantidade de outras ordens de mamíferos, que diversificaram após a extinção em massa do período de transição Cretáceo-Paleógeno, que ficou conhecida como “*K-Pg extinction*” ocorrida ao redor de 65 milhões de anos atrás (POZZI *et al.*, 2014).

Conforme mostrado na figura 7, os Strepsirrhini (Lemuriformes e Lorisiformes) constituem os ramos mais basais dos primatas, os quais apresentam como distinção dos outros primatas a presença de uma narina com uma dobra de membrana mucosa que a comunica com as gengivas e lábio superior. Ainda outras características marcantes são a presença de um órgão vômero-nasal para detecção de feromônios, útero bicornado, e uma placenta epiteliocorial, e dentes em formato de pente. O Clado Strepsirrhini, apresenta como representantes notáveis os

lêmures de Madagascar, os quais fazem parte da infraordem Lemuriformes (WILSON & REEDER 2005). Haplorrhini é o clado irmão de Strepsirrhini agrupando uma diversidade muito maior de animais, que inclui os grandes ramos dos primatas do Novo Mundo (NWM) ou Platyrrhini, que ocorrem no continente americano (ARISTIDE *et al.*, 2015). Os OWM (primatas do Velho Mundo, ou Catarrhini), constituem outro grande clado de Haplorrhini, assim como os Tarsiiformes, que são representados por poucas espécies do Velho Mundo, incluindo *Carlito syrichta*, espécie das Filipinas que analisamos nessa Tese. Como distinção de Strepsirrhini, o clado dos Haplorrhini apresenta um *rostrum* relativamente curto, formação óssea separando a órbita da fossa temporal, e narinas individualizadas com relação ao lábio superior e gengivas, o que lhes permite exibir uma gama maior de expressões faciais. Concomitantemente, também é possível observar diferentes comportamentos relacionados a sua estrutura social. Tais peculiaridades em primatas do Novo Mundo têm sido bem documentadas na espécie *Callithrix jacchus*, que apresentam comportamento monogâmico, além de intenso cuidado parental por parte dos machos. Assim que, ao cuidar da prole, o macho pode facilitar e usufruir de um curto intervalo reprodutivo. Contudo, filhotes relativamente imaturos resultam num intenso cuidado parental, aumentando o estabelecimento de vínculos afetivos por longos períodos, podendo ser observado ainda comportamento monogâmico nessas espécies (STOREY & ZIEGLER, 2016).

Além dos primatas, dados originais de outro grupo particular de mamíferos, os marsupiais, foram obtidos para a presente Tese, de modo que no item abaixo características gerais sobre os mesmos serão apresentadas.

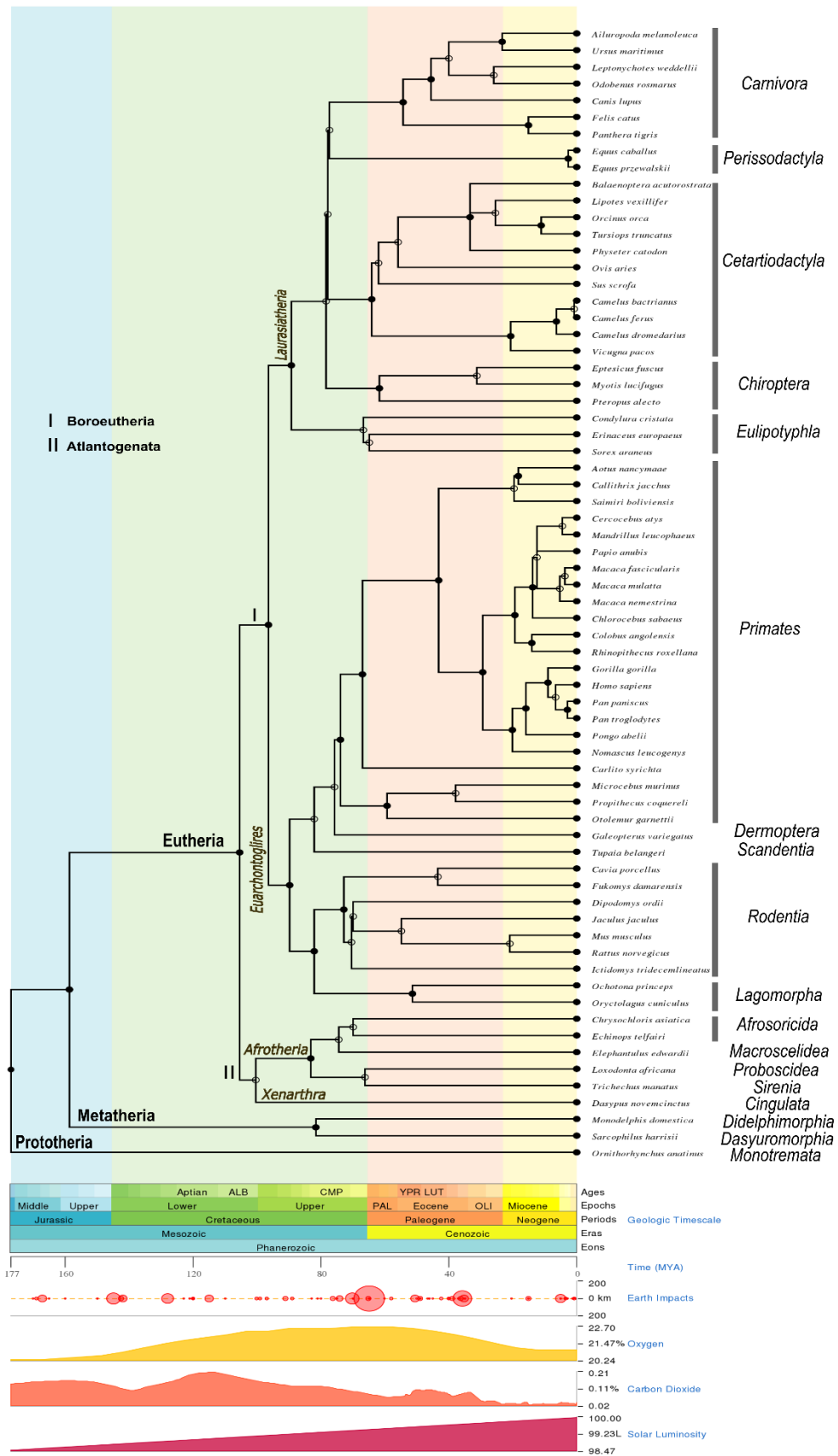


Figura 6 – Amostragem de diferentes espécies de mamíferos que podem atualmente ser recuperadas através dos bancos de dados genômicos. Figura adaptada do *Time Tree of Life* (<http://www.timetree.org/>).

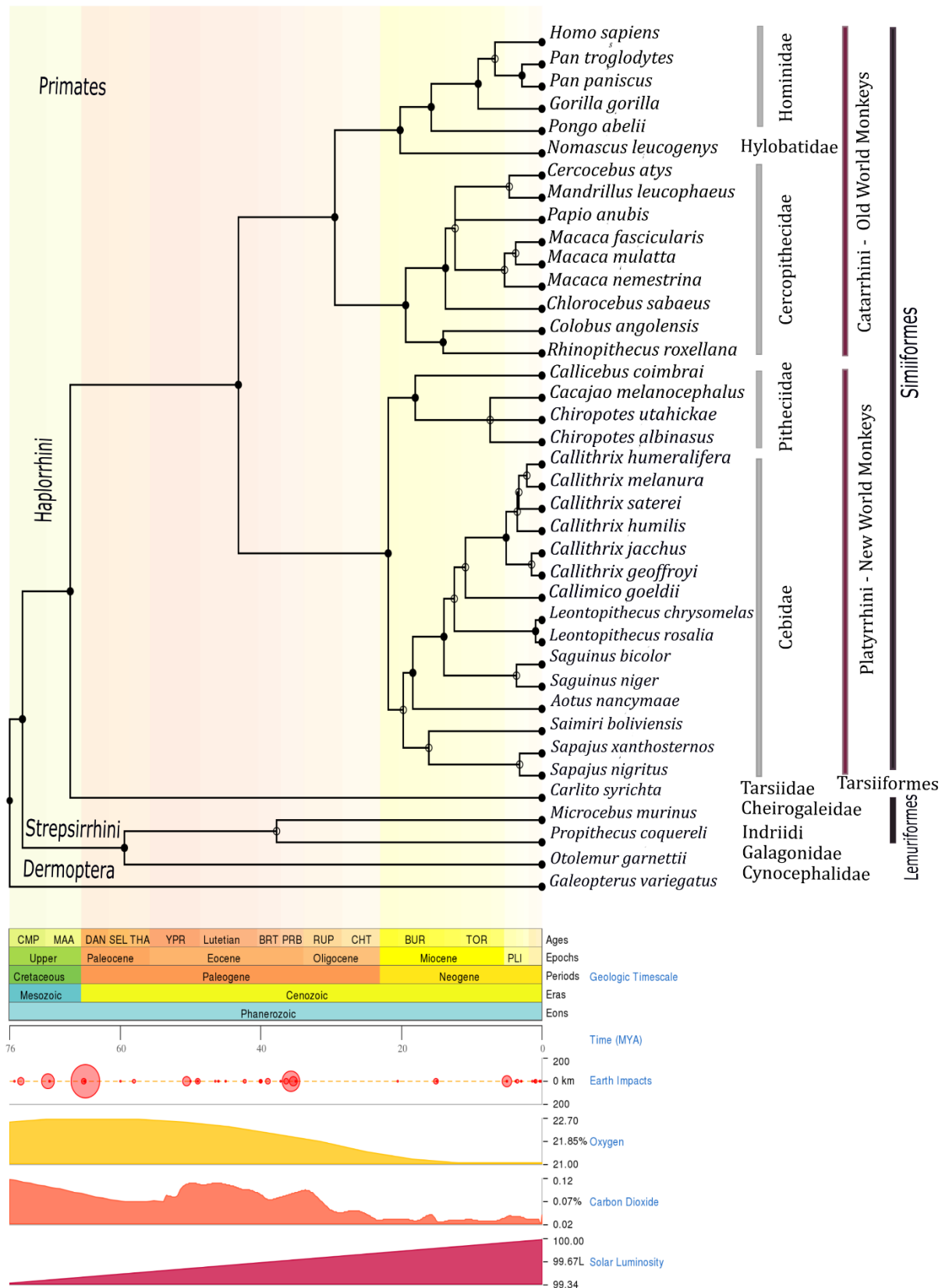


Figura 7 – Filogenia dos primatas obtida através do *Time Tree of Life* e com identificação taxonômica revisada em Perelman *et al.*, (2011).

1.5.2 Marsupiais

Os marsupiais (Metatheria) fazem parte da subclasse de mamíferos Theria, que agrupa também os mamíferos placentários (Eutheria; WILLIAMSON *et al.*, 2014). O registro fóssil mais antigo e confiável de Theria foi descoberto no nordeste da China: uma espécie similar a um musaranho que foi batizada com o nome de *Juramaia sinensis*. Sua datação contribuiu para a indicação de que os clados Metatheria e Eutheria teriam divergido há cerca de 160 milhões de anos (LUO *et al.*, 2011).

Foram relativamente abundantes e muito diversos durante o Cretáceo, mas no final desse período, há cerca de 65 milhões, ocorreu a extinção em massa *K-Pg* (POZZI *et al.*, 2014), fazendo desaparecer diversas espécies de animais, incluindo muitas do clado Metatheria. Os marsupiais tiveram uma grande radiação mais recente na América do Sul e na Austrália, posterior à extinção *K-Pg*, sendo que a conexão entre as espécies da América do Sul e Austrália é atualmente alvo de muitos debates (AMRINE-MADSEN *et al.*, 2003).

Em termos de densidade de táxons foi observada uma grande redução recente nos continentes Americano e Australiano, que ocorreu provavelmente devido à chegada (na América do Sul) e introdução (na Austrália) de mamíferos placentários (<http://animaldiversity.org/>). Na América do Sul, são encontradas principalmente duas famílias de Marsupiais, Didelphidae, que ocorre também na América do Norte e Caenolestidae, encontrada somente na América do Sul (MAY-COLLADO *et al.*, 2015). Didelphidae é a família mais abundante que compreende cerca de 95 espécies. Em termos de reprodução os membros dessa família são considerados poligínicos, ou seja, os machos competem pelas fêmeas e emitem ruídos ao circundá-las, mas esse comportamento não é considerado de corte (ASTÚA *et al.*, 2015). A família Didelphidae possui ainda um único representante considerado semiaquático, a espécie *Chironectes minimus* que é conhecida vulgarmente como cuíca-d'água. Ela é encontrada nos rios e lagos do Sul do México até a Argentina, sendo uma espécie promíscua que exhibe comportamentos pré-copulatórios, desenvolvendo forte relação social e, aparentemente, apresenta cuidado parental. A outra família, Caenolestidae, por sua vez, contém cerca de 5 espécies, as quais são de pequeno porte e similares a um musaranho. Curiosamente, nas fêmeas dessa família a bolsa está ausente (<https://www.ed.ac.uk/>).

As características reprodutivas variam consideravelmente, sendo muitas das espécies solitárias por todo o ano, ficando juntos somente para o acasalamento. Esse padrão reflete, provavelmente, uma estrutura social de acasalamento promíscua (<http://animaldiversity.org/>). Contudo, é possível observar características reprodutivas mais sociais em algumas espécies,

como *Macropus parryi* que pode abranger grupos de 50 a 80 indivíduos. A formação de pares, ou seja, comportamento monogâmico, também pode ser observado na espécie *Petauroides volans* que compartilha a toca até que o filhote deixe o marsúpio. Nessa espécie o macho também pode exibir comportamento bígamo, embora não exibam cuidado parental. Essa transição entre monogâmico e poligâmico estaria mais relacionada com a disponibilidade de alimento, sendo que em áreas com pouca disponibilidade a monogamia é considerada a norma, mas quando há bastante recursos, os machos exibem o comportamento bígamo (NORTON, 1988).

CAPÍTULO II *JUSTIFICATIVA E OBJETIVOS*

Nosso grupo de pesquisa tem feito contribuições importantes ao conhecimento de variações específicas relacionadas ao cuidado parental no sistema *OXT* e seu receptor *OXTR*, assim como de outros genes da mesma rede *AVPR1b* (ver anexos I e IV, VARGAS PINILLA *et al.*, 2015; FAM *et al.*, sob revisão). Nossos esforços em demonstrar o efeito dessas variantes *in vivo* e *in vitro* foram efetivos no caso de *OXT* (PARREIRAS-E-SILVA *et al.*, 2017). Sabe-se que características complexas fazem parte de uma rede maior de interações que podem vir a ser reveladas, ou seja, há um repertório genético no qual vários genes de uma mesma rota podem apresentar maior ou menor relevância na sua conexão com fenótipos adaptativos. Por isso, expandimos nossos estudos para além dos genes do sistema *OXT-AVP*, incluindo mais genes candidatos na investigação.

O presente trabalho teve como objetivos basicamente o que segue:

- Descrever os padrões evolutivos para as regiões codificantes dos genes *OXTR*, *AVPR1a*, *AVPR1b*, *AVPR2*, *PRL*, *PRLR* e *PRLH*; se correspondentes a neutralidade, a seleção purificadora, ou direcional/seleção positiva;
- Elucidar se as mudanças de aminoácidos das proteínas estariam relacionadas com características táxon-específicas e seu potencial impacto na história evolutiva das espécies investigadas.

CAPÍTULO III ***OXYTOCIN AND ARGININE VASOPRESSIN***
RECEPTOR EVOLUTION: IMPLICATIONS FOR ADAPTIVE
NOVELTIES IN PLACENTAL MAMMALS



Oxytocin and arginine vasopressin receptor evolution: implications for adaptive novelties in placental mammals

Pamela Paré^{1,*}, Vanessa R. Paixão-Côrtes^{2,*}, Luciana Tovo-Rodrigues³, Pedro Vargas-Pinilla¹, Lucas Henriques Viscardi¹, Francisco Mauro Salzano¹, Luiz E. Henkes⁴ and Maria Catira Bortolini¹

¹Programa de Pós-Graduação em Genética e Biologia Molecular, Departamento de Genética, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil.

²Programa de Pós-Graduação em Genética e Biodiversidade, Instituto de Biologia, Universidade Federal da Bahia (UFBA), Salvador, BA, Brazil.

³Laboratório de Fisiologia da Reprodução Animal, Universidade Federal de Santa Catarina (UFSC), Curitibanos, SC, Brazil.

⁴Programa de Pós-Graduação em Epidemiologia, Universidade Federal de Pelotas (UFPEL), Pelotas, RS, Brazil.

Abstract

Oxytocin receptor (*OXTR*) and arginine vasopressin receptors (*AVPR1a*, *AVPR1b*, and *AVPR2*) are paralogous genes that emerged through duplication events; along the evolutionary timeline, owing to speciation, numerous orthologues emerged as well. In order to elucidate the evolutionary forces that shaped these four genes in placental mammals and to reveal specific aspects of their protein structures, 35 species were selected. Specifically, we investigated their molecular evolutionary history and intrinsic protein disorder content, and identified the presence of short linear interaction motifs. *OXTR* seems to be under evolutionary constraint in placental mammals, whereas *AVPR1a*, *AVPR1b*, and *AVPR2* exhibit higher evolutionary rates, suggesting that they have been under relaxed or experienced positive selection. In addition, we describe here, for the first time, that the *OXTR*, *AVPR1a*, *AVPR1b*, and *AVPR2* mammalian orthologues preserve their disorder content, while this condition varies among the paralogues. Finally, our results reveal the presence of short linear interaction motifs, indicating possible functional adaptations related to physiological and/or behavioral taxa-specific traits.

Keywords: Oxytocin receptor, Arginine vasopressin receptors, molecular evolution, protein disorder, interaction motifs.

Received: December 18, 2015; Accepted: February 28, 2016.

Introduction

Genome and tandem duplication events have an important role in biological evolution (Van de Peer *et al.*, 2009). These processes give rise to paralogous genes, which can evolve by speciation along the evolutionary timeline, thus giving rise to orthologous genes (Fitch, 1970; Gabaldón and Koonin, 2013). As result of these processes, so-called “gene families” emerge, whose members may retain similar or identical functions, but might also diverge extensively, resulting in adaptive novelties (Neduva and Russell, 2005; Huang and Sarai, 2012). The emergence and differentiation of the paralogous neuroendocrine non-peptides oxytocin (OXT) and vasopressin (AVP; Gwee *et*

al., 2009), as well of their paralogous receptors (*OXTR* and *AVPR1a*, *AVPR1b*, *AVPR2*, respectively) illustrate this phenomenon (Yamaguchi *et al.*, 2012; Lagman *et al.*, 2013).

The OXT peptide is comprised of a nine amino acid sequence (Lee *et al.*, 2009), differing in only two amino acids from its paralogue, AVP. These nonapeptides, produced in their highest quantities in the brain, mediate both similar and distinct functions through their interactions with their native receptors (*OXTR*; *AVPR1a*, *AVPR1b*, and *AVPR2*), which are produced in various organs and tissues (Barberis *et al.*, 1998). Some level of cross-reaction among OXT and AVP with their non-native receptors occurs as well, but with distinct affinities (Zingg and Laporte, 2003; Slusarz *et al.*, 2013). For instance, the synthesis of *OXTR* in the uterus and mammary glands guarantees uterine contraction and milk ejection in placental mammals (Kimura *et al.*, 1998; Gimpl and Fahrenholz, 2001), whereas *AVPR1a*

Send correspondence to Maria Catira Bortolini. Departamento de Genética, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil. E-mail: maria.bortolini@ufrgs.br

*These authors contributed equally to this work.

**CAPÍTULO IV *MOLECULAR EVOLUTION OF THE PROLACTIN
RECEPTOR (PRLR) AND IMPLICATIONS FOR ADAPTIVE TRAITS IN
PRIMATES (manuscrito em preparação: primeira versão)***

Molecular evolution of the prolactin receptor (PRLR) and implications for adaptive traits in Primates.

¹Paré P, ¹Vargas-Pinilla P, ¹Viscardi LH, ¹Reales G, ¹Fam B, ²Paixão-Côrtes VR, ³Santos FR, ¹Bortolini MC.

¹Laboratório de Evolução Humana e Molecular, Programa de Pós-Graduação em Genética e Biologia Molecular, Departamento de Genética, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil.

²Laboratório de Biologia Evolutiva e Genômica (LABEG), Programa de Pós-Graduação Biodiversidade e Evolução, Instituto de Biologia, Universidade Federal da Bahia (UFBA), Salvador, BA, Brazil.

³Laboratório de Biodiversidade e Evolução Molecular, Departamento de Biologia Geral da Universidade Federal de Minas Gerais (UFMG), Belo-Horizonte, MG, Brazil.

Correspondence

Maria Cátira Bortolini, Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil.
E-mail: maria.bortolini@ufrgs.br

Abstract

The prolactin-prolactin receptor (PRL-PRLR) complex is a pivotal system for physiological and behavioral features in the entire animal kingdom. It shows remarkable complexity regarding splicing variants and, consequently, signaling mechanisms. Specifically, PRLR Exon 8 (corresponding to the protein intracellular region) has been previously highlighted for its high variability and signaling ability for specific pathways such as JAK/STAT. In this study, we used the PRLR Exon 8 sequence data to evaluate its likely connection with reproduction and behavioral characteristics in Primates. Seventeen primate species were sequenced and analyzed together sequences from 32 species retrieved from public databases (N = 49). We found great variability in the PRLR Exon 8 in Primates, with a signal of positive selection ($\omega = 1.9$). A total of six substitutions were suggested with a highest posterior probability to be under positive selection, with one significant site, Leu⁵⁰⁷ (BEB = 0.952). For some of these sites (507, 532 and 572) we found association with adaptive traits as the number of offspring and sexual maturity in females. Additionally, variations in the Simiiformes branch with adaptive potential, as the emergence of SH3 ligand, suggest further capability for Src kinase modulation of PRLR signaling. Different from the published literature related to PRL, no positive selection signal was found in any of the primate branches suggesting distinct evolutionary strategies employed by PRL-PRLR systems during Primates' diversification.

Keywords: New World Monkeys, PRLR; Long Isoform, Molecular adaptive changes.

Introduction

Prolactin (PRL) is a neurohormone (~200 amino acids) secreted by the anterior pituitary which acts in a broad spectrum of phenotypes and is associated with essential physiological, behavioral, and reproductive traits (AKEN *et al.*, 2016). PRL plays a more specific role in pregnancy and lactation, growth and development, metabolism, immunological modulation, electrolyte transport, integument regulation, and oncogenesis (BEN-JONATHAN & HUGO, 2015). Its production occurs mainly in the cells of the anterior pituitary gland, which are called lactotrophs, although it acts in both autocrine and paracrine systems (KUCKA *et al.*, 2013).

In mammals, PRL has a pivotal role in lactation and intensive maternal care, which is essential for the occurrence of a health relationship between the mother and her offspring (BRIDGES, 2015). A pioneering study in this area reported higher PRL plasma levels in common marmoset (*Callithrix jacchus*) mothers with single suckling infants than non-lactating mothers, and even significantly higher in mothers with suckling twins (MCNEILLY *et al.*, 1981). The same work also found no implications of elevated plasma concentration of PRL in postpartum ovulation delay, unlike what was observed in women and rhesus females (*Macaca mulata*; MCNEILLY *et al.*, 1981). Highest levels of PRL were also observed in fathers exhibiting paternal behavior, as demonstrated in *Callithrix jacchus*, which showed higher PRL concentration in the hypothalamus of experienced fathers than non-fathers, also showing that PRL level increased after birth (SCHRADIN *et al.*, 2003; STOREY & ZIEGLER, 2016).

All the functional performances of PRL depends on its interaction with the dimeric PRLR protein complex, composed of two receptor subunits (BOLE-FEYSOT *et al.*, 1998). PRLR belongs to the class 1 of cytokine receptor family, comprising of an extracellular region (EC), a single transmembrane region (TM), and an intracellular region (IC), with its activity mediated by ligand-induced homodimerization (GOFFIN & KELLY, 1997). PRLR lacks intrinsic enzymatic activity being constitutively associated with kinases, as Janus kinase 2 (JAK2) and Src family of protein tyrosine kinases (SFKs) in its intracellular region. These interactions are required for phosphorylation of PRLR and of the other proteins downstream, which includes Jak2/STAT as the main signaling pathway, Ras-Raf-MAPK, and PI-3K (CLEVINGER & MEDAGLIA, 1994; DUSANTER-FOURT *et al.*, 1994). The JAK2/STAT pathway regulates many critical aspects of cell growth, survival and differentiation, and errors in this pathway are often observed in many human tumors, reflecting their importance for cell maintenance and integrity. Activated JAK protein leads to phosphorylation of tyrosine residues at PRLR with

subsequent recruitment and phosphorylation of STAT proteins resulting in downstream signaling molecules (MITCHELL & JOHN, 2005). Src kinases also regulate signal transduction through downstream molecules signaling, and are composed of an N-terminal region followed by the SH3 and SH2 domains and a short C-terminal region (BOGGON & ECK, 2004; PARSONS & PARSONS, 2004).

Three PRLR isoforms (short, intermediate and long), resulting from alternative splicing have been described, but some of them are unable to translate signaling via interaction with PRL (BOLE-FEYSOT *et al.*, 1998). On the other hand, it is known that for efficient activation of all related signaling pathways, the complete expression of the PRLR is required (CLEVENGER *et al.*, 2003). For instance, the long isoform (which express the three domains of PRLR, EC, TM, and IC) has been reported as the only capable of stimulating the transcription of milk protein genes (LESUEUR *et al.*, 1991). Additional studies, considering the PRLR genetic variations within species of economic interest, in turn, revealed alleles associated with phenotypic changes of prolificacy-related traits (*e.g.*, litter size, number of teats, puppy survival, and milk production/composition (PUTNOVA *et al.*, 2002; RENS *et al.*, 2003; TOMÁS *et al.*, 2006; Lü *et al.*, 2011).

Experiments with PRLR knockout female mice showed severe deficits in maternal care behaviors when compared to PRL knockout mice, suggesting a more critical role for the PRLR receptor than for the neurohormone regarding the modulation of those responses (HORSEMAN *et al.*, 1997; LUCAS *et al.*, 1998). Other investigations showed that an increased level of PRLR mRNA was positively correlated with paternal care in hamsters and fishes (KHONG *et al.*, 2009). Interestingly, stronger expression of PRLR in scented glands of muskrat (*Ondatra zibethicus*) male was observed during the breeding season (CLEVENGER *et al.*, 2003). These findings encouraged Babb and colleagues to characterize intraspecific PRLR coding variation in *Aotus azarai*, an NWM monogamous species (BABB *et al.*, 2014). It is also known that *A. azarai* males become direct caregivers when their infants have more than a week of life (BABB *et al.*, 2014, and references therein). In summary, Babb *et al.*, (2014) revealed that PRLR exhibits extensive nucleotide, codon, and amino acid changes, being the IC domain the most variable portion, considering *Aotus azarai* variation and both approaches (intra- and interspecific). Although PRLR forms may be driven by selection for different behaviors and physiologies from modulations of the pleiotropic prolactin pathway, they did not find evidence that differences in PRLR among primate taxa were correlated with taxon-level differences in paternal care.

Identifying the genetic basis of complex reproductive phenotypes has been one of the greatest challenges of modern science since it is expected that a wide range of neuroendocrine genes and pathways, as well as environmental and epigenetic factors, have some (greater or lesser) connection with the expression of these traits. Considering also an evolutionary context, one more level of complexity is posed since the traits under investigation can have potential adaptive relevance. Notwithstanding the challenges, recent studies published by us, and other research groups regarding the neuroendocrine hormone oxytocin (OXT)-vasopressin (AVP) system, could be cited as successful examples. More specifically, approaches concerning association analysis between taxon-specific genetic changes in OXT, and individual adaptive behaviors as direct paternal, intense maternal care, and social monogamy in Platyrrhini or New World monkeys (NWMs), were found, especially in Callitrichidae (REN *et al.*, 2015; VARGAS-PINILLA *et al.*, 2015; FRENCH *et al.*, 2016; PARREIRAS-&-SILVA *et al.*, 2017). The particularities regarding the Callitrichidae clade include small body size and frequent occurrence of twin births, commonly followed by females lactate and pregnant, simultaneously, due to an early postpartum estrus (STOREY & ZIEGLER, 2016). As a consequence, additional effort is required to ensure their survival, from their father, siblings, and also unrelated individuals. Some studies showed that the survival rate of the Callitrichidae infants is related to the intense and direct (*e.g.*, feeding, carrying and grooming) care of both parents. It was also observed that no female had success in raising their offspring without the help of at least one male (CLEVELAND & SNOWDON, 1984). The functionality of these NWM OXT variants and their potential for modulating behavior in male mice has also been recently demonstrated by our research group (VARGAS-PINILLA *et al.*, 2015; PARREIRAS-&-SILVA *et al.*, 2017).

Here we sequenced the exon 8 portion of PRLR, corresponding to the highly variable IC region, in 17 NWM species to explore and expand our knowledge about the genetic repertoire behind behavior-related adaptive traits in different mammal clades. Thus, we analyzed our NWM original data set with other data recovered from databases for Primates and one Dermoptera species, as an outgroup. Our goal was to delineate a general scenario about PRLR IC molecular evolution, diversity, and its possible connections with the emergence of adaptive phenotypes in Primates.

Material and methods

DNA Samples and Sequencing

We sampled 48 Primate and one Dermoptera species, including original sequence data from 17 Platyrrhini (NWM) species, and public data from genomic databases and the literature, as further detailed below.

Blood samples from 15 Primates were provided by Rio de Janeiro's Primatology Center (CPRJ), geographically located between 22°27'S-22°32'S and 42°50'W-42°56'W. This collection of biological material from conservation units for research is registered in the SISBIO system under the number 27951-2. SISBIO is a Brazilian official remote-access system that allows researchers to request authorizations for biological material collection and research at conservation units. Additionally, *Callithrix penicillata* (n = 5), *Callicebus nigrifrons* (n = 3) and *Callithrix geoffroyi* (n = 6) were provided by the Taxonomic Collections Center of the Universidade Federal de Minas Gerais (UFMG). See Table S1 for full species list.

Extraction of genomic DNA from blood and tissue samples were performed in two ways: DNA from blood samples were purified using the QIAamp® DNA Mini Kit (QIAGEN GmbH, Hilden, Germany), according to the protocol recommended by the manufacturer; DNA purification from tissue samples followed the Phenol: Chloroform protocol (SAMBROOK & RUSSELL, 2006).

This project was released for execution by the Ethics Committee on the Use of Animals (CEUA) of the Universidade Federal do Rio Grande do Sul, on November 25, 2014.

PCR and Sequencing

We designed the primers based on conserved regions within Primates by using the UCSC Genome Browser database (<http://genome.ucsc.edu/>) and considering a marmoset (*Callithrix jacchus*) as the reference genome. The Oligo Analyzer Software (www.genelink.com) was used for primer quality checking. Three sets of primers were designed to cover the entire PRLR Exon 8 with 1021 base pairs:

Set 1): PRLR_E8F1 CCTGGCACACAAATGTCTGAAG, and PRLR_E81R GGATTCTCTGGCTTCTCAATGAC, TM = 61°C.

Set 2): PRLRE8_F2 CTGGATCCTGACACTGACTC, and PRLRE8_R2 AGCTTTCTACCTCCCTCTG; TM = 61°C.

Set 3): PRLR_E8F3 GTGTGTGAGCTGGCTGTG, and PRLR_E8R3 GTGTGTGAGCTGGCTGTG; TM = 61°C).

Moreover, a combination between PRLRE8_F2 CTGGATCCTGACACTGACTC and PRLR-Ex8-R GTCCCTCAAGAATACTAAGCAG from Babb *et al.*, (2014) was also performed. The general PCR protocol implemented here contained in the reaction mixture 1 µl of genomic DNA (about 20 ng), 0.5 µl of each primer (final concentration), five µl of Master Mix (Qiagen, Hilden, Germany), and three µl of water in the final volume of 10 µl. The thermocycling conditions were as follows: 5 min at 95 °C, followed by 35 cycles of 30 sec at 94 °C, 40 sec at the optimal annealing temperature, and 1 min at 72 °C, and a final extension of 7 min at 72 °C. PCR products were purified using exonuclease-alkaline phosphatase (Thermo-Fisher Scientific, Waltham, MA, USA) and then sequenced using Applied Biosystem 3130xl or 3730 Genetic Analyzer Sequencers.

Data mining

The canonical sequences for PRLR – IC (E8) of 31 Primates, and one Dermoptera species (Table S1) were downloaded using the TaxOnTree tool (<http://maxixe.icb.ufmg.br/tot/#x>), automatically coloring a phylogenetic tree according to their taxonomical information obtained from NCBI taxonomy (www.ncbi.nlm.nih.gov/taxonomy). The download of the sequences was performed using “exclude isoforms” option, to avoid misleading. A script was used to automatically download the nucleotide sequence through the complete cds provided by TaxOnTree (<https://github.com/tetsufmbio/collab>). Moreover, a search in the online databases was performed, including further sequences from GenBank (BENSON *et al.*, 2014) and UCSC Genome Browser database (KENT *et al.*, 2002). The exome of *Homo neanderthalensis* was also checked making use of the Max Planck Institute database (CASTELLANO *et al.*, 2014). Some few sequences were also obtained from the literature (BABB *et al.*, 2014).

Analyses

Sequence analysis

The individual sequences had their quality manually checked using Codon Code Aligner 4 (Codon Code Corporation, Centerville, MA, USA). After that, consensus sequences were generated for each, and multiple alignments were performed in MEGA 7.0 version (KUMAR,

STECHEER *et al.*, 2016), where the sequences were also further visualized and manually checked. The haplotype information was obtained with PHASE algorithm (STEPHENS & DONNELLY 2003) available in DnaSP software (ROZAS, 2009).

The amino acid changes were classified considering their Grantham score (GS) as conservative (GS 0–50), moderately conservative (51–100), moderately radical (101–150), and radical (> 151 ; GRANTHAM 1974, LI *et al.*, 1984).

Positive selection analyses

We estimated the rates of nonsynonymous (dN) and synonymous (dS) substitutions ($\omega = dN/dS$), to explore the putative effect of natural selection on *PRLR* (IC) gene. In general, ω values are widely used as an indicator of natural selection acting on protein-coding sequences. Thus, $\omega < 1$ ($dS > dN$) indicates negative selection, $\omega \cong 1$ indicates neutral or relaxed selection, and $\omega > 1$ ($dN > dS$) indicates positive selection. We estimated ω values following two approaches, implemented in the codeml program of PAML package (YANG, 2007). For both approaches, phylogenetic species trees are required, and we obtained them from the *Time Tree of Life* web server (<http://www.timetree.org>), and the topology was further checked considering the classification proposed by Perelman *et al.*, (2011), and further revised by Schneider & Sampaio, (2015).

In the first approach, we used site models (NsSites), which estimate ω on sites across the alignment, allowing ω to vary between sites (NIELSEN & YANG, 1998). Two types of models, assuming neutrality (M1a and M7, $0 < \omega \leq 1$) or allowing positive selection (M2 and M8, $\omega \geq 1$) were applied to our data. Their likelihood function was estimated and compared pairwise (M1a vs. M2a, $df = 2$; M7 vs. M8, $df = 2$), using a Likelihood Ratio Test (LRT) implemented in HyPhy v2.2.4 package (POND *et al.*, 2005). These models differ in their assumptions and parameters (see YANG, 2007) for more detailed information). In the second approach, we used branch-site models. These models allow ω to vary both among sites and branches, and constitute a powerful tool to detect episodic selective pressure acting in specific lineages. Model M2a_rel was used as the neutral model, and Clade C as the alternative, positive selection model (WEADICK & CHANG, 2012). Given the high coverage and quality of the sequences included in the analysis, we considered all sites for ω calculation (clean data option = 0; we manually removed from the analysis intraspecific variable sites that we detected in our primate diversity data). When models of positive selection differed significantly in their log-likelihood from

neutral models, a post hoc test Bayes Empirical Bayes was implemented to determine which specific sites are potentially under positive selection (YANG *et al.*, 2005).

Phenotypic traits association and functionality

Behavioral and reproductive traits as gestation period, interbirth interval, number of offspring, age of sexual maturity, social structure, dominant sex, dispersive sex, mating system, and weaning age were compiled from the Zoology Museum of the University of Michigan (<http://animaldiversity.ummz.umich.edu>; Table S2) and the literature (KAMILAR & COOPER, 2013). In order to explore statistical associations between phenotypic traits and specific amino acid sites putatively evolving under positive selection, we performed classical statistical tests (i.e., Kruskal-Wallis and Fisher's Exact Test) using the Superwise package v0.1 (available at <https://github.com/GRealesM/superwise>) in R environment (R CORE TEAM, 2017).

The Superwise function uses two sets of data: phenotypic data and amino acid data; then it performs pairwise trait-amino acid analyses for each column in both datasets. Kruskal-Wallis was run when phenotypes were numerical, and Fisher Exact Test when categorical. Then *P*-values were calculated for each test and corrected by Bonferroni.

These positively selected sites were also explored regarding their putative changes in protein signaling since adaptive modifications in the protein sequences may lead to both conformational and interaction changes. To functionally investigate these putative interaction changes in the protein disordered regions, we used the Eukaryotic Linear Motif (ELM) web server (DINKEL *et al.*, 2016) as an exploratory tool for motif prediction. ELM predicts Short Linear Motifs (SLiMs), common stretches of protein sequence in intrinsically disordered regions (IDRs), which consist of approximately 3–11 contiguous amino acids with functional importance and essential for protein-protein interaction (NEDUVA & RUSSELL, 2005; DINKEL *et al.*, 2014).

Intra-specific diversity

At the intraspecific level, we performed neutrality tests (HKA, McDonald–Kreitman, Tajima's D, and Fu & Li) using DNAsp 6 software (ROZAS *et al.*, 2017), and population structure (Fst) in Arlequin 3.5 (EXCOFFIER & LISCHER, 2010) to detect any deviations from neutral expectations. Additionally, aDNA sequence information (haplotype) and their

frequency obtained in DNAsp, a Median – Joining Network was built using PopART (BANDELT *et al.*, 1999).

Results

Primates interspecific specificities

We sequenced the entire PRLR – IC (E8; 1014 bp) for 17 NWM species. This data set was also analyzed together with information from 31 other primate species and *Galeopterus variegatus* (Table S1). We found a signal of positive selection within Exon 8 of PRLR, since the alternative model (M8), in the NsSites likelihood ratio test, was significantly better than the neutral one (M7), and admits positive selection ($p = 0.02$; Table 1). As indicated by the Bayesian analysis (BEB), one site was specifically found with significant posterior probability to be under positive selection ($\omega = 1.9$; BEB = 0.952; Figure 1a). Moreover, according to the test, 7% of the sites from PRLR - IC were under positive selection with the highest posterior probability coming from the site class of $\omega > 1$ ($\omega = 1.9$; Table 1), although they do not reach significance. We found that the changes observed in position 507 are associated with sexual maturity in females (Kruskal-Wallis test, $p < 0.001$, Figure 1a, Table S2).

Regarding other sites showing a high posterior probability to be under positive selection (>70%), another association between Asp⁵³²Asn and number of offspring was found through Kruskal-Wallis test ($p < 0.001$). The Asn amino acid was concomitantly observed in Callitrichidae species expressing the twin births phenotype (Figure 1a; Table S3; present in *Callithrix*, *Saguinus*, and *Leontopithecus*), except for *Callimico goeldii* that does not fit to this pattern. Similarly, Asp⁵⁷²Ser also occurs in species delivering twins (Figure 1a, Table S2), as cited above, and was also observed in association with sexual maturity, which has lower values in Callitrichidae species containing Ser amino acid (Table S2, Figure 1a). However, in this position, *Leontopithecus rosalia* and *Callimico goeldii*, does not show this pattern.

As a whole, some radical changes in PRLR (IC) were found: as the Cys³⁰⁷Tyr (GS = 194) which emerges in Siimiformes; While Tyr⁵⁰⁹Asp (GS= 160) was found exclusively in *Aotus nancymaae*. As shown in Figure 1b, all these sites came from PRLR – long form. While, at Tyr⁴¹⁴ conserved residue within the Box 2 motif, there is a fixed Phe⁴¹⁴ in *Carlito syrichta* (Haplorrhini: Tarsiiformes), Strepsirrhini, and *Galeopterus variegatus* (Order Dermoptera), which is a functional domain within PRLR (IC). Meanwhile, other putative sites for protein

phosphorylation Ser³⁵⁶, Ser³⁵⁸, and Thr⁴¹⁵ were found highly conserved among Primate species (Figure 1a, and 1b).

Noteworthy the apparent great variability in some basal branches of Primates, including *Carlito syrichta* (Haplorrhini) and Strepsirrhini, and the outgroup Dermoptera with *Galeopterus variegatus* as representative species. So, we seek out to elucidate the pattern of variation within lineages calculating the Ney & Gojobori mean distance within groups (Basal branch, Catarrhini, and Platyrrhini). Our results showed the basal branches (dS-dN= 0.256) as the most variable ones, contrasting with Simiiformes: Catarrhini (dS-dN= 0.041) and Platyrrhini (dS-dN= 0.040). However, PAML branch sites analysis does not find a significant branch with different ω values (dN/dS) in Primates (Table S4).

Regarding SLiMs diversity, the sites above cited were not found to change the protein profile of interactions with other proteins. The most frequent motifs for interactions were LIG_SH3_3 (SH3 ligand) and TRG_ENDOCYTIC_2 (Y-based sorting signal), with the emergence of LIG_SH3_5 (SH3 ligand) in the Simiiformes (Figure 2), and LIG_SH3_2 (SH3 ligand) only found in *Pongo abelii* species. While their composition among species is conservative (Figure 2).

Primates diversity

We evaluated the diversity within the PRLR – IC (E8) among some NWM species (*Callithrix geoffroyi*, *Callithrix penicillata*, and *Callicebus nigrifons*; Table S1). Additionally, sequences of *Aotus azarai*, *Aotus lemurinus*, *Aotus nancymae*, and *Aotus nigriceps* from Babb *et al.*, (2014) were also considered in the analysis. We stressed here that the putative selected sites (507, 532, and 572) with phenotypic associations, as above described, have intra-specific conservation.

We found in the *Callithrix geoffroyi* species four polymorphic sites, been two of them nonsynonymous substitutions, as follows: Ser⁴²²Thr, and Ala⁴⁴⁴Thr (GS= 58; Table S5). The Ser⁴²²Thr mutation was found only in heterozygosity in 3% of individuals, whereas Thr⁴⁴⁴Ala was present in heterozygosity (40%), and homozygosity (30%). On the other hand, we found the lowest PRLR - IC diversity in *Callithrix penicillata* species, with only one synonymous substitution at Pro³⁹⁴ (heterozygous; 2%; Table S5). One nonsynonymous substitution was also seen in *Callicebus nigrifons*, but position 569 (Val⁵⁶⁹Met; GS= 21) was only found in homozygosity (3%; Table S5). Our described variations here for the *Callithrix* and *Callicebus* genera are species-specific in comparison to Owl monkeys (*Aotus sp.*) intraspecific data

published (BABB *et al.*, 2014). Moreover, we found that only species of the *Callithrix* genus, *C. penicillata*, and *C. geoffroyi*, were sharing alleles (Figure 3).

According to our search for human variants using Ensembl, we found one unique nucleotide variant already reported in the literature, which was the synonymous Pro³⁹⁴, although, there is no phenotypic description.

We applied the classical population tests (HKA, McDonald–Kreitman, Tajima's D, and Fu & Li, or Fst-based population structure, data not shown) to detect deviations from neutrality in the PRLR - IC of Primates, neither of them was significant.

Discussion

We stand out through molecular evolutionary analyses a prominent role of PRLR – long form in the modulation of cell signaling in Primates, from which all the positively selected sites came from. Our LRT analysis between NsSites models indicated model 8 (admits positive selection) as the best molecular evolution model ($\omega = 1.9$; Table 1), in both, entire exon 8 and just within the long form. This protein region (IC) was previously suggested as the most variable in PRLR, considering Primates (interspecific) and *Aotus* (intraspecific) approaches (BABB *et al.*, 2014). Nevertheless, we proposed Leu⁵⁰⁷ to be evolving under positive selection (95% of ω probability = 1.9) among six sites from the PRLR - long form with a highest posterior probability (>70%). The previous study of PRLR in Primates have not tested sites-specific positive selection (BABB *et al.*, 2014). Changes found at Leu⁵⁰⁷ position and in other site suggested by BEB analysis (532, and 572; Figure 1a) were associated with evidence of the adaptive phenotypes. Adaptive traits included offspring size, sexual maturity ratio in female and twin births. These changes in positions 507, 532, and 572 (Figure 1a) are specifically found in Calithrichidae species (*Callithrix*, *Saguinus*, and *Leontopithecus*). Associations of PRLR genetic variations within species of economic interest also revealed alleles associated with phenotypic changes of prolificacy-related traits (e.g., litter size, number of teats, puppy survival, and milk production/composition (PUTNOVA, KNOLL *et al.*, 2002, RENS, GJ *et al.*, 2003, TOMÁS, CASELLAS *et al.*, 2006, LÜ, HU *et al.*, 2011). However, as in the case of NWM, twin births appear concomitantly with monogamous behavior, and it is very difficult to distinguish its real effect, whether through acting in specific behaviors of the group, or through its physiological action to maintain the offspring due to PRL pleiotropic effect.

We further explored the PRLR variability at the intraspecific level, which indicates no intraspecific variation regarding the amino acid sites above cited and suggests our results representative of interspecific variability. In turn, marmosets (*C. geoffroyi* and *C. penicillata*) were the only species sharing alleles, which were already reported to hybridize in nature (MALUKIEWICZ *et al.*, 2014). Unique variations found for each species appears to be a pattern for PRLR (IC) intraspecific variability. As highlighted by Babb *et al.*, (2014), the variability found in full PRLR within the platyrrhines or catarrhines have a substantial number of lineage-specific nucleotide changes that distinguished them from Strepsirrhini and Dermoptera.

Regarding SLiMs diversity, the sites above cited were not found to change protein profile of interactions with other proteins. However, an additional SH3 ligand (LIG_SH3_5) was observed emerging in the Simiiformes branch (Figure 2). SH3 are motifs for Src kinase bind proteins, members of the Src family that are kinases and modulate the signaling of other proteins acting in cellular processes, including cell growth, differentiation, cell shape, migration and survival, and specialized cell signals (PARSONS & PARSONS, 2004). Are highly expressed in the adult brain regulating neuronal plasticity and behavior through tyrosine phosphorylation of key substrates such as neurotransmitter receptors (OHNISHI *et al.*, 2011). Since these proteins have SH3 and SH2 domains (BOGGON & ECK, 2004), they can be function as additional modulators of PRLR signaling in the Simiiformes branch by an intramolecular interaction between the SH3 domain of Src with PRLR (IC), resulting in suppressed kinase activity (OHNISHI *et al.*, 2011). It denotes an adaptive novelty in the Haplorrhini branch regarding PRLR signaling upon a complex network of interactions performed by Src kinase.

Although there is an above-cited indication of adaptive novelties in primate branches, it was not found a signal of positive selection in Simiiformes, and in any other branches, agreeing with the finding of the Babb *et al.*, (2014), and contrasting with the reported PRL evidence of episodic evolution in the basal branch of Primates (WALLIS *et al.*, 2005). These data show distinct evolutionary strategies followed by PRL-PRLR systems during evolution.

We could identify here some specific phosphorylation target sites not tightly conserved among Primates (Figure 1). For instance, at Tyr⁴¹⁴ conserved residue within the Box 2 motif in Simiiformes, there is a Phe⁴¹⁴ only in Tarsiiformes (*Carlito syrichta*), Strepsirrhini, and in *Galeopterus variegatus*, a functional domain within PRLR (IC). This may imply in distinct patterns of phosphorylation shaped throughout PRLR-long form molecular evolution. Despite these variations, other putative sites for phosphorylation Ser³⁵⁶, Ser³⁵⁸, and Thr⁴¹⁵, were found

conserved among Primates, suggesting that their important role through kinase signaling has been maintained along the evolutionary history. The maintenance of such invariant positions (purifying selection) suggest that they also play an important role in cell signaling in other species, and their maintenance is further highlighted by their association with carcinogenic variant finds in humans, according to COSMIC (<http://cancer.sanger.ac.uk/cosmic>): Ser³⁵⁶Arg (Malignant melanoma; COSM3615593); Ser³⁵⁶Asn (Carcinoma; COSM5463437); Ser³⁵⁸ (Malignant melanoma; COSM3615592); and Ser³⁵⁸Phe (Malignant melanoma; COSM5717241). Either through Src protein family signaling or other, a PRLR proper intracellular signal is dependent on Tyr and Ser residues for cellular proliferation through stimulation of transcriptional factors (BERLANGA 1995).

Our previous studies found signals of positive selection and association with reproductive and life history traits in NWMs in the OXT system (VARGAS-PINILLA *et al.*, 2015; PARREIRAS-&-SILVA *et al.*, 2017), and in the 5-HT system (personal communication). It is widely known that these two systems interact with the PRL network, and such interaction is required for the regulation of social behavior in mammals and other animals. We suggest that the variants found in PRLR – long form, together with our findings in the other neuromodulation systems, are interacting and evolving together, leading to the observed behaviors, and taking an important part of the genetic background that allows Primates gain the novel adaptive traits that characterize them. Although these results should be considered very carefully, they constitute a very interesting genetic background for future biological tests.

Proteins with multiple functions evolve by acquiring new phenotypes and preserving their ancestral functions. Therefore, several gene portions, and consequently the protein domains, evolve unevenly through time, whereas some them are highly conserved due to purifying selection others tolerate great variation. In addition, a significant number of sites present signals of relaxed selective constraint, can be related with a condition compatible with the pre-adaptive model (RAMAN *et al.*, 2016, and references therein).

Conclusions

In summary, our results of intra and interspecies analyses highlight a prominent role of PRLR – long form in the modulation of cell signaling in Primates. The positively selected sites from PRLR – long form suggest amino acid changes playing an adaptive role in the reproductive traits of Primates, such as the number of offspring and sexual maturity in females. Our analysis indicates innovative features regarding Src kinase signaling in PRLR, specifically in Siimiformes branch. On the other hand, residues at Ser³⁵⁶, Ser³⁵⁸, and Thr⁴¹⁵ have been maintained conserved along Primates evolution, depicting their essential role in the intracellular signaling through kinase interactions. We highlighted here important PRLR (IC) features which could form the basis for future biological tests *in vitro*, further demonstrating their biological implication.

Acknowledgements

This research was financially supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), FAPEMIG and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). We thank Alcides Pissinatti from Centro de Primatologia do Rio de Janeiro by donation of Primate Samples and to the Taxonomic Collections Center of Universidade Federal de Minas Gerais for allowing to access. We also thank Tetsu Sakamoto for design a specific script to facilitate our use of TaxOnTree software.

References

- AKEN, B. L. et al. The Ensembl gene annotation system. **Database (Oxford)**, v. 2016, 2016. ISSN 1758-0463. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/27337980> >.
- BABB, P. L. et al. **Prolactin Receptor Gene Diversity in Azara's Owl Monkeys (*Aotus azarai*) and Humans (*Homo sapiens*) Suggests a Non-Neutral Evolutionary History among Primates**. *International Journal of Primatology*, p.129–155. 2014
- BANDEL, H. J.; FORSTER, P.; RÖHL, A. Median-joining networks for inferring intraspecific phylogenies. **Mol Biol Evol**, v. 16, n. 1, p. 37-48, Jan 1999. ISSN 0737-4038. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/10331250> >.
- BEN-JONATHAN, N.; HUGO, E. Prolactin (PRL) in adipose tissue: regulation and functions. **Adv Exp Med Biol**, v. 846, p. 1-35, 2015. ISSN 0065-2598. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/25472532> >.
- BENSON, D. A. et al. GenBank. **Nucleic Acids Res**, v. 42, n. Database issue, p. D32-7, Jan 2014. ISSN 1362-4962. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/24217914> >.
- BERLANGA, J. J. et al. Prolactin receptor is associated with c-src kinase in rat liver. *Mol Endocrinol*, v. 9, n. 11, p. 1461-7, Nov 1995. ISSN 0888-8809. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/8584023> >.
- BOGGON, T. J.; ECK, M. J. Structure and regulation of Src family kinases. **Oncogene**, v. 23, n. 48, p. 7918-27, Oct 2004. ISSN 0950-9232. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/15489910> >.
- BOLE-FEYSOT, C. et al. Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. **Endocr Rev**, v. 19, n. 3, p. 225-68, Jun 1998. ISSN 0163-769X. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/9626554> >.
- BRIDGES, R. S. Neuroendocrine regulation of maternal behavior. **Front Neuroendocrinol**, v. 36, p. 178-96, Jan 2015. ISSN 1095-6808. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/25500107> >.
- CASTELLANO, S. et al. Patterns of coding variation in the complete exomes of three Neandertals. *Proc Natl Acad Sci U S A*, v. 111, n. 18, p. 6666-71, May 20
- CLEVELAND, J.; SNOWDON, C. T. Social development during the first twenty weeks in the cotton-top tamarin (*Saguinus o. oedipus*). 32: 12 p. 1984.
- CLEVINGER, C. V. et al. The role of prolactin in mammary carcinoma. **Endocr Rev**, v. 24, n. 1, p. 1-27, Feb 2003. ISSN 0163-769X. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/12588805> >.

CLEVENGER, C. V.; MEDAGLIA, M. V. The protein tyrosine kinase P59fyn is associated with prolactin (PRL) receptor and is activated by PRL stimulation of T-lymphocytes. **Mol Endocrinol**, v. 8, n. 6, p. 674-81, Jun 1994. ISSN 0888-8809. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/7935483> >.

DINKEL, H. et al. The eukaryotic linear motif resource ELM: 10 years and counting. **Nucleic Acids Res**, v. 42, n. Database issue, p. D259-66, Jan 2014. ISSN 1362-4962. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24214962> >.

DUSANTER-FOURT, I. et al. Identification of JAK protein tyrosine kinases as signaling molecules for prolactin. Functional analysis of prolactin receptor and prolactin-erythropoietin receptor chimera expressed in lymphoid cells. **EMBO J**, v. 13, n. 11, p. 2583-91, Jun 1994. ISSN 0261-4189. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/8013458> >.

EXCOFFIER, L.; LISCHER, H. E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. **Mol Ecol Resour**, v. 10, n. 3, p. 564-7, May 2010. ISSN 1755-0998. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/21565059> >.

FRENCH, J. A. et al. Neuropeptide diversity and the regulation of social behavior in New World primates. **Front Neuroendocrinol**, Mar 2016. ISSN 1095-6808. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/27020799> >.

GRANTHAM, R. Amino acid difference formula to help explain protein evolution. *Science*, v. 185, n. 4154, p. 862-4, Sep 1974. ISSN 0036-8075. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/4843792> >.

GOFFIN, V.; KELLY, P. A. The prolactin/growth hormone receptor family: structure/function relationships. **J Mammary Gland Biol Neoplasia**, v. 2, n. 1, p. 7-17, Jan 1997. ISSN 1083-3021. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/10887515> >.

HORSEMAN, N. D. et al. Defective mammapoiesis, but normal hematopoiesis, in mice with a targeted disruption of the prolactin gene. **EMBO J**, v. 16, n. 23, p. 6926-35, Dec 1997. ISSN 0261-4189. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/9384572> >.

KAMILAR, J. M.; COOPER, N. Phylogenetic signal in primate behaviour, ecology and life history. **Philos Trans R Soc Lond B Biol Sci**, v. 368, n. 1618, p. 20120341, May 2013. ISSN 1471-2970. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/23569289> >.

KENT, W. J. et al. The human genome browser at UCSC. **Genome Res**, v. 12, n. 6, p. 996-1006, Jun 2002. ISSN 1088-9051. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/12045153> >.

KHONG, H. K. et al. Prolactin receptor mRNA is upregulated in discus fish (*Symphysodon aequifasciata*) skin during parental phase. **Comp Biochem Physiol B Biochem Mol Biol**, v. 153, n. 1, p. 18-28, May 2009. ISSN 1879-1107. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/19272315> >.

KUCKA, M. et al. The role of cyclic nucleotides in pituitary lactotroph functions. **Front Endocrinol (Lausanne)**, v. 4, p. 122, 2013. ISSN 1664-2392. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24062725> >.

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Mol Biol Evol**, v. 33, n. 7, p. 1870-4, Jul 2016. ISSN 1537-1719. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/27004904> >.

LESUEUR, L. et al. Comparison of long and short forms of the prolactin receptor on prolactin-induced milk protein gene transcription. **Proc Natl Acad Sci U S A**, v. 88, n. 3, p. 824-8, Feb 1991. ISSN 0027-8424. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/1992473> >.

LI, W. H.; WU, C. I.; LUO, C. C. Nonrandomness of point mutation as reflected in nucleotide substitutions in pseudogenes and its evolutionary implications. **J Mol Evol**, v. 21, n. 1, p. 58-71, 1984. ISSN 0022-2844. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/6442359> >.

LUCAS, B. K. et al. Null mutation of the prolactin receptor gene produces a defect in maternal behavior. **Endocrinology**, v. 139, n. 10, p. 4102-7, Oct 1998. ISSN 0013-7227. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/9751488> >.

LÜ, A. et al. Novel SNPs of the bovine PRLR gene associated with milk production traits. **Biochem Genet**, v. 49, n. 3-4, p. 177-89, Apr 2011. ISSN 1573-4927. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/21165768> >.

MALUKIEWICZ, J. et al. Hybridization effects and genetic diversity of the common and black-tufted marmoset (*Callithrix jacchus* and *Callithrix penicillata*) mitochondrial control region. **Am J Phys Anthropol**, v. 155, n. 4, p. 522-36, Dec 2014. ISSN 1096-8644. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/25186076> >.

MCNEILLY, A. S. et al. Plasma prolactin concentrations during the ovarian cycle and lactation and their relationship to return of fertility post partum in the common marmoset (*Callithrix jacchus*). **J Reprod Fertil**, v. 62, n. 2, p. 353-60, Jul 1981. ISSN 0022-4251. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/6788944> >.

MITCHELL, T. J.; JOHN, S. Signal transducer and activator of transcription (STAT) signalling and T-cell lymphomas. **Immunology**, v. 114, n. 3, p. 301-12, Mar 2005. ISSN 0019-2805. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/15720432> >.

NEDUVA, V.; RUSSELL, R. B. Linear motifs: evolutionary interaction switches. **FEBS Lett**, v. 579, n. 15, p. 3342-5, Jun 2005. ISSN 0014-5793. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15943979> >.

NIELSEN, R.; YANG, Z. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. **Genetics**, v. 148, n. 3, p. 929-36, Mar 1998. ISSN 0016-6731. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/9539414> >.

OHNISHI, H. et al. Src family kinases: modulators of neurotransmitter receptor function and behavior. **Trends Neurosci**, v. 34, n. 12, p. 629-37, Dec 2011. ISSN 1878-108X. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/22051158> >.

PARREIRAS-E-SILVA, L. T. et al. Functional New World monkey oxytocin forms elicit an altered signaling profile and promotes parental care in rats. **Proc Natl Acad Sci U S A**, Aug 2017. ISSN 1091-6490. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/28784762> >.

PARSONS, S. J.; PARSONS, J. T. Src family kinases, key regulators of signal transduction. **Oncogene**, v. 23, n. 48, p. 7906-9, Oct 2004. ISSN 0950-9232. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/15489908> >.

PERELMAN, P. et al. A molecular phylogeny of living primates. **PLoS Genet**, v. 7, n. 3, p. e1001342, Mar 2011. ISSN 1553-7404. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21436896> >.

POND, S. L.; FROST, S. D.; MUSE, S. V. HyPhy: hypothesis testing using phylogenies. **Bioinformatics**, v. 21, n. 5, p. 676-9, Mar 2005. ISSN 1367-4803. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/15509596> >.

PUTNOVA, L. et al. **A new Hpa II PCR-RFLP within the porcine prolactin receptor (PRLR) gene and study of its effect on litter size and number of teats.** *Animal breeding and Genetics*. 119: 7 p. 2002.

RAMAN, A. S.; WHITE, K. I.; RANGANATHAN, R. Origins of Allostery and Evolvability in Proteins: A Case Study. **Cell**, v. 166, n. 2, p. 468-480, Jul 2016. ISSN 1097-4172. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/27321669> >.

REN, D. et al. Genetic diversity in oxytocin ligands and receptors in new world monkeys. **PLoS One**, v. 10, n. 5, p. e0125775, 2015. ISSN 1932-6203. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25938568> >.

RENS, B. V.; GJ, E.; T, V. D. L. **Components of litter size in gilts with different prolactin receptor genotypes.** *Theriogenology*. 59: 26 p. 2003.

ROZAS, J. et al. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. **Mol Biol Evol**, v. 34, n. 12, p. 3299-3302, Dec 2017. ISSN 1537-1719. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/29029172> >.

SAMBROOK, J.; RUSSELL, D. W. Purification of nucleic acids by extraction with phenol:chloroform. **CSH Protoc**, v. 2006, n. 1, Jun 2006. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/22485786> >.

STEPHENS, M., SMITH, N. and DONNELLY, P. (2001). A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, 68, 978-989.

SCHNEIDER, H.; SAMPAIO, I. The systematics and evolution of New World primates - A review. **Mol Phylogenet Evol**, v. 82 Pt B, p. 348-57, Jan 2015. ISSN 1095-9513. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/24201058> >.

SCHRADIN, C. et al. Prolactin and paternal care: comparison of three species of monogamous new world monkeys (*Callicebus cupreus*, *Callithrix jacchus*, and *Callimico goeldii*). **J Comp Psychol**, v. 117, n. 2, p. 166-75, Jun 2003. ISSN 0735-7036. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/12856787> >.

STOREY, A. E.; ZIEGLER, T. E. Primate paternal care: Interactions between biology and social experience. **Horm Behav**, v. 77, p. 260-71, Jan 2016. ISSN 1095-6867. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/26253726> >.

R Core TEAM. **R: A language and environment for statistical computing**. Foundation for Statistical Computing. Vienna, Austria. 2017.

TOMÁS, A. et al. High amino acid variation in the intracellular domain of the pig prolactin receptor (PRLR) and its relation to ovulation rate and piglet survival traits. **J Anim Sci**, v. 84, n. 8, p. 1991-8, Aug 2006. ISSN 1525-3163. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/16864857> >.

VARGAS-PINILLA, P. et al. Evolutionary pattern in the OXT-OXTR system in primates: Coevolution and positive selection footprints. **Proc Natl Acad Sci U S A**, v. 112, n. 1, p. 88-93, Jan 2015. ISSN 1091-6490. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25535371> >.

WALLIS, O. C. et al. Molecular evolution of prolactin in primates. **J Mol Evol**, v. 60, n. 5, p. 606-14, May 2005. ISSN 0022-2844. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/15983870> >.

WEADICK, C. J.; CHANG, B. S. An improved likelihood ratio test for detecting site-specific functional divergence among clades of protein-coding genes. **Mol Biol Evol**, v. 29, n. 5, p. 1297-300, May 2012. ISSN 1537-1719. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/22319160> >.

YANG, Z. PAML 4: phylogenetic analysis by maximum likelihood. **Mol Biol Evol**, v. 24, n. 8, p. 1586-91, Aug 2007. ISSN 0737-4038. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17483113> >.

YANG, Z.; WONG, W. S.; NIELSEN, R. Bayes empirical bayes inference of amino acid sites under positive selection. **Mol Biol Evol**, v. 22, n. 4, p. 1107-18, Apr 2005. ISSN 0737-4038. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/15689528> >.

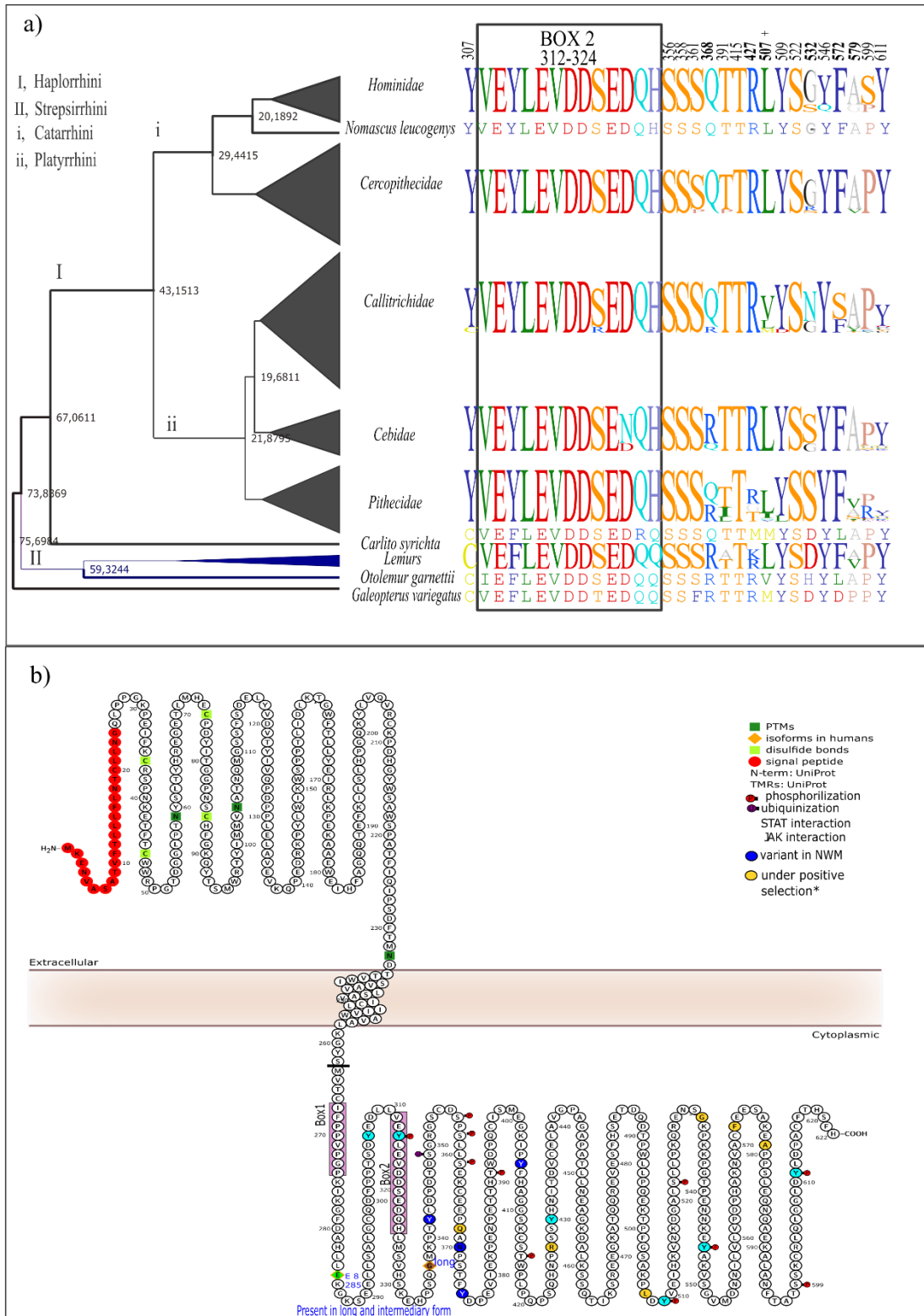


Figure 1 – PRLR (IC) variability within Primates: a) PRLR (IC) amino acid changes in putative adaptive positions indicated through PAML analysis ($> 70\%$ of ω probability to be under positive selection are shown with numbers in bold), and Tyr and Ser residues conservation important for protein phosphorylation. b) PRLR plot depicting protein regions and sites specificities, snake plot obtained through the open-source tool for visualization of proteoforms, Protter (Omasits, Ahrens et al. 2014). Functional information, as important Tyr residues for intracellular signaling through kinases and the assignment of intermediate and long PRLR form were obtained in Kline *et al.*,

(1999). * Residues with probability higher than 70% to be under $\omega = 1.9$. The tree topology with nodes and divergence time was obtained in the Time Tree of Life web server (<http://www.timetree.org>).

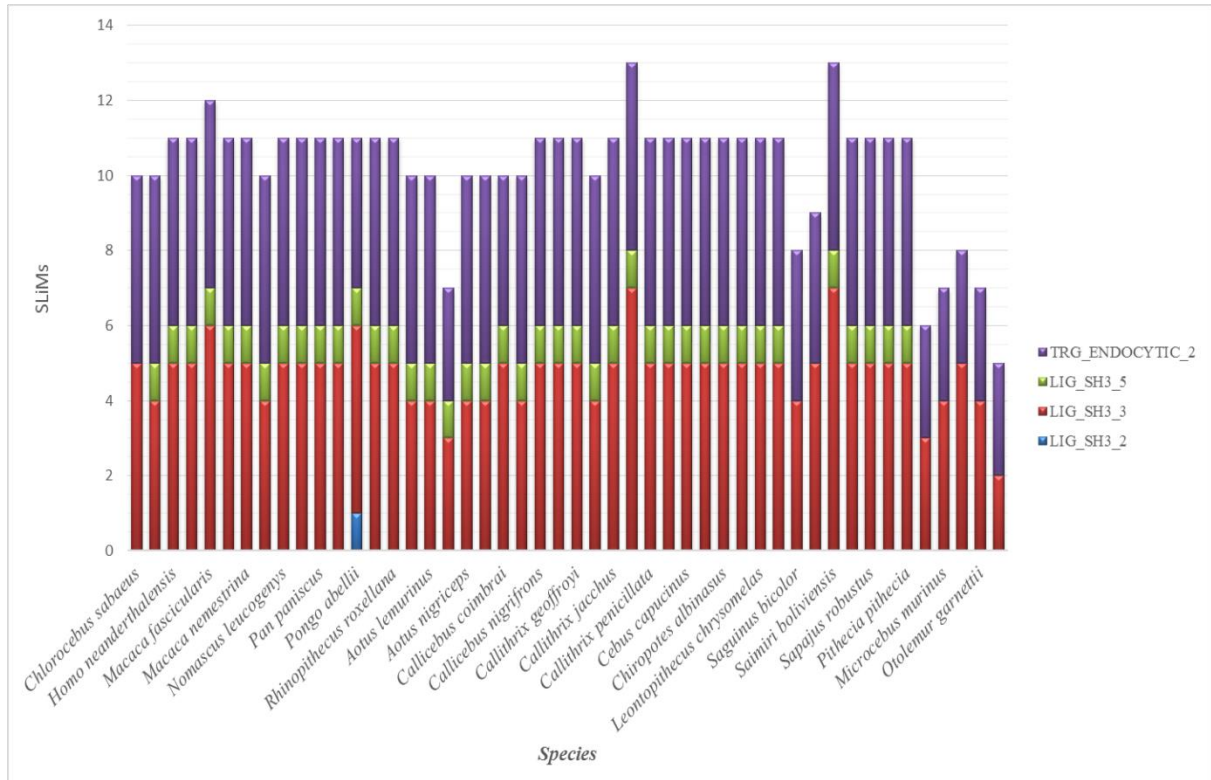


Figure 2 - Predicted SLiMs for PRLR (IC) in Primates. LIG SH3 in accordance to ELM description these domains (Src Homolgy 3) are small protein modules of about 50-60 residues. It is found in proteins involved in several biological processes as diverse as signal transduction pathways, cytoskeleton organization, membrane traffic or organelle assembly. SH3 domains bind to proline-rich peptides that form a left-handed polyproline type II helix (PPII) in one of two opposite orientations. Ligand orientation depends on the position of the positive residue in the target peptide. Peptides that bind in a type I orientation contain the consensus R/KXXPPXP, while the type II orientation conform to the PXXPXR/K consensus. TRG_Endocyt_2 are Tyrosine-based sorting signal. Intracellular traffic of many membrane-bound proteins requires sequences facing the cytoplasm. In many instances the sorting information is encoded in short peptide motifs, typically 4-6 amino acids, which are referred to as sorting signals. These motifs determine which vesicular traffic pathway is used to transport a particular molecule and hence determine its final destination. The Yxxphi motif is a sorting signal found in the cytosolic tail of various membrane proteins as an endocytic motif and also directs traffic within the endosomal and the secretory pathways. The Yxxphi motif interacts with mu subunit of AP (Adaptor Protein) complex (DINKEL *et al.*, 2016).

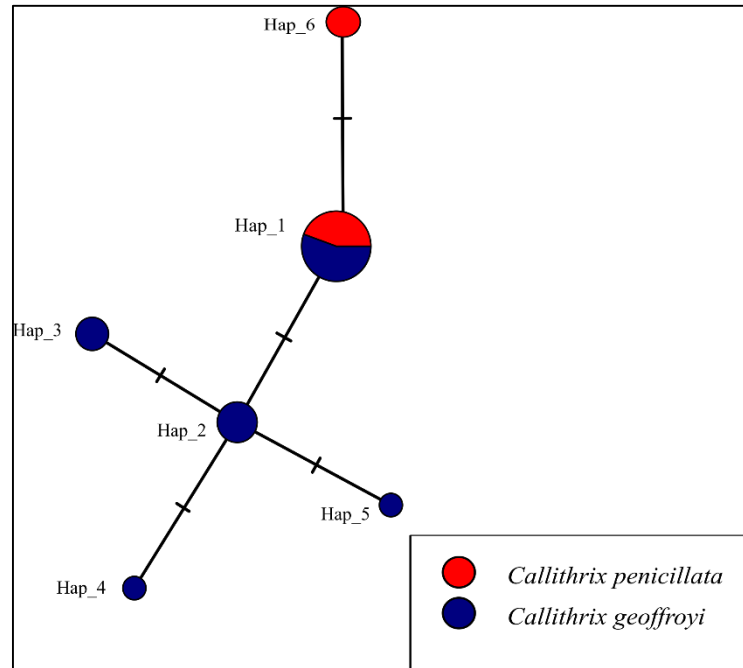


Figure 3 – Median - joining network of PRLR (IC) coding region from 12 individuals of *Callithrix* genus.

Table 1 - PRLR estimated parameters under different codon substitution models for Primate species in IC (E8).

Gene	Model	d_N/d_S	Estimated parameters	ℓ	P value
PRLR - E8	M1a: neutral	0.5481	$p_0 = 0.53567$, ($p_1 = 0.49457$) ($\omega_0 = 0.26627$), ($\omega_1 = 1.00000$)	-5258.76	M1a vs M2a $P > 0.999$
	M2a: selection	0.5859	$p_0 = 0.67127$, $p_1 = 0.26311$, ($p_2 = 0.06562$) ($\omega_0 = 0.30594$), ($\omega_1 = 1.00000$), $\omega_2 = 1.78957$	-5257.35	
	M7: β	0.5200	[$p = 0.53567$, $q = 0.49457$]	-5950.82	M7 vs M8 $P = 0.02$
	M8: β and ω	0.5572	$p_0 = 0.92900$, ($p_2 = 0.07100$), $p = 0.83516$ $q = 0.99992$, $\omega_2 = 1.89672$	-5946.98	
PRLR - Long (aa 338-621)	M1a: neutral	0.5438	$p_0 = 0.62335$, ($p_1 = 0.37665$) ($\omega_0 = 0.26815$), ($\omega_1 = 1.00000$)	-5309.56	M1a vs M2a $P > 0.999$
	M2a: selection	0.5827	$p_0 = 0.69087$, $p_1 = 0.23239$, ($p_2 = 0.07673$) ($\omega_0 = 0.31365$), ($\omega_1 = 1.00000$), $\omega_2 = 1.74185$	-5307.98	
	M7: β	0.5443	[$p = 0.69474$, $q = 0.58197$]	-5310.49	M7 vs M8 $P = 0.02$
	M8: β and ω	0.5830	$p_0 = 0.90125$, ($p_2 = 0.09875$), $p = 1.31981$ $q = 1.57117$, $\omega_2 = 1.74169$	-5306.67	

p = proportion of sites within each omega (ω) class; $\omega_0 < 1$ (negative selection), $\omega_1 = 1$ (neutral selection), and $\omega_2 > 1$ (positive selection). ℓ = Log Likelihood Values. Likelihood ratio tests were performed between neutral models (M1a- nearly neutral, and M7 - beta) and models that admits positive selection (M2a - selection, and M8, β & ω - beta +selection). The comparisons between neutral and positive selection models had 2 degrees of freedom. Within parentheses: fixed parameters; Within brackets: Beta parameters p and q . dN/dS = non-synonymous/synonymous rate ratio.

Table S1 – Primate species originally sampled for PRLR, and recovered from data mining.

SPECIES	REFSEQ*	E-value
<i>Galeopterus variegatus</i>	XP_008584330.1	1.00E-122
<i>Aotus azarai</i>	(Babb <i>et al.</i> , 2014)	NA
<i>Aotus lemurinus</i>	(Babb <i>et al.</i> , 2014)	NA
<i>Aotus nancymae</i>	XP_012315391.1	1.00E-148
<i>Aotus nigriceps</i>	(Babb <i>et al.</i> , 2014)	NA
<i>Callithrix jacchus</i>	NP_001254666.1	1.00E-143
<i>Callithrix humeralifera</i>	M1856	NA
<i>Callithrix melanura</i>	M1942	NA
<i>Callithrix penicillata</i>	MM1700138; MM1700143; MM1700046; MM1700154; MM1700226; MM1700050	NA
<i>Callithrix saterei</i>	M2260	NA
<i>Callithrix geoffroyi</i>	M2216; MM1700048; MM1700074; MM1700077; MM1700058; MM1700040	NA
<i>Callimico goeldii</i>	M2033	NA
<i>Leontopithecus chrysomelas</i>	M1859	NA
<i>Leontopithecus rosalia</i>	M1620	NA
<i>Saguinus bicolor</i>	M1896	NA
<i>Saguinus niger</i>	M1977	NA
<i>Saimiri boliviensis</i>	XP_003925979.1	5.00E-146
<i>Sapajus paella</i>	AY227708.1	NA
<i>Sapajus robustos</i>	M2209	NA
<i>Sapajus xanthosternos</i>	M0474	NA
<i>Cebus capucinus</i>	XM_017548276.1	NA
<i>Cacajao melanocephalus</i>	M2072	NA
<i>Chiropotes utahickae</i>	M2095	NA
<i>Chiropotes albinasus</i>	M2094	NA
<i>Callicebus donacophilus</i>	(Babb <i>et al.</i> , 2014)	NA
<i>Callicebus nigrifrons</i>	MM1700132; MM1700133; MM1700134	NA
<i>Callicebus coimbrai</i>	M2275	NA
<i>Pithecia pithecia</i>	(Babb <i>et al.</i> , 2014)	NA
<i>Cercocebus atys</i>	XP_011910550.1	5.00E-163
<i>Chlorocebus sabaesus</i>	XP_007959535.1	7.00E-163
<i>Colobus angolensis</i>	XP_011799730.1	NA
<i>Macaca fascicularis</i>	XP_005556762.1	4.00E-163
<i>Macaca mulatta</i>	XP_001092241.1	7.00E-164
<i>Macaca nemestrina</i>	XP_011724970.1	4.00E-163
<i>Mandrillus leucophaeus</i>	XP_011824502.1	5.00E-163
<i>Papio Anubis</i>	XP_009206518.1	9.00E-162
<i>Rhinopithecus roxellana</i>	XP_010382384.1	3.00E-163
<i>Rhinopithecus bieti</i>	XM_017894520.1	NA
<i>Microcebus murinus</i>	XP_012613586.1	3.00E-121
<i>Otolemur garnettii</i>	XP_003793026.1	2.00E-116
<i>Gorilla gorilla</i>	XP_004059023.1	1.00E-163
<i>Homo sapiens</i>	NP_000940.1	7.00E-167
<i>Homo neanderthalensis</i>	ENST00000382002 ^U	NA

<i>Pan paniscus</i>	XP_003806712.1	6.00E-164
<i>Pan troglodytes</i>	XP_001150064.2	6.00E-164
<i>Pongo abelii</i>	NP_001153263.1	3.00E-165
<i>Nomascus leucogenys</i>	XP_003274991.1	3.00E-164
<i>Propithecus coquereli</i>	XP_012495965.1	2.00E-123
<i>Carlito syrichta</i>	XP_008060865.1	2.00E-113

* Samples provided by the Taxonomic Collections Center of the Universidade Federal de Minas Gerais (UFMG) present the prefix "MM"; All the others were provided by Rio de Janeiro's Primatology Center (CPRJ). The sequences were downloaded from NCBI; U UCSC genome browse

Table S2 – Behavioural and reproductive traits recovered by Primate species.

Species	Gestation period/year	Interbirth interval/year	Interbirth gestation ratio	Offspring	Sexual maturity F/year	Sexual maturity M/year	Sexual maturity ratio FM	Gestation lifespan ratio	Social structure	Dominant Sexual	Dispersive Sexual	Mating system	Gestation	Weaning Age/days
<i>Aotus azarae</i>				1	2	2			MFPB			MGY	129.5	255
<i>Aotus lemurinus</i>				1	2.5	2.5			MFPB			MGY	133	
<i>Aotus nigriceps</i>		0.801		1	3.5	2.00E+00			MFPB			MGY	120	133
<i>Aotus nancymae</i>	0.364383 6 0.493150	0.5 1.972602	1.3721805	1	2.1216667	2.12E+00	1	0.0182192	MFPB	N	F	MGY		
<i>Cacajao melanocephalus</i>	7	7	4	1	3.5833333	3.58E+00	1	0.0273973	MMMMF	N	B	NonMGY		
<i>Callicebus donacophilus</i>				1	2	4			MFPB	N		MGY	126	
<i>Callicebus nigrifrons</i>	0.350684													
<i>Callicebus coimbrai</i>	9 0.424657	1	2.8515625	1	1.6666667	2	0.833	0.0140274	MFPB	N	F	MGY		
<i>Callimico goeldii</i>	5	0.5	1.1774194	1	1.1666667	7	1	0.0237239	MFPB	N	B	NonMGY	153	66.53
<i>Callithrix penicillata</i>				2	1.47945	1.48E+00			MFPB			MGY	150	56
<i>Callithrix saterei</i>	0.397260 3	0.5	1.2586207	2	1.2083333	1.67E+00	0.725	0.033105	OFMM	N	B	MGY		
<i>Callithrix geoffroyi</i>	0.438356 2	0.575342 5	1.3125	2	1.375	1.375	1	0.0438356	MFPB	N	B	MGY		
<i>Callithrix humeralifer</i>	0.383561 6	0.5	1.3035714	2	1.2083333	7	0.725	0.0306849	OFMM	N	B	MGY	144	99.01
<i>Callithrix jacchus</i>	0.405479 5	0.437808 2	1.0797297	2	1.3333333	3	1.231	0.0253425	MFPB	N	B	MGY	144	60.24
<i>Callithrix melanura</i>	0.397260 3	0.517808 2	1.3034483	2	1.2083333	7	0.725	0.0248288	OFMM	N	B	MGY		
<i>Carlito syrichta</i>	0.493150 7	0.5	1.0138889	1	2	2	1	0.035225	MFPB	N	B	NonMGY	179	82.49
<i>Cebus capucinus</i>				1	5	8			MMMMF	M		NonMGY	160	90
<i>Cercocebus atys</i>	0.457534 2	1	2.1856287	1	4.5833333	4	1.146	0.0170722	MMMMF	M	M	NonMGY		
<i>Chiropotes albinasus</i>	0.390411	2	5.122807	1	4	4	1	0.0260274	MMMMF	N	F	NonMGY		
<i>Chiropotes utahicki</i>	0.369863 0.449315	2	5.4074074	1	4	4	1	0.0246575	MMMMF	N	F	NonMGY		
<i>Chlorocebus sabaeus</i>	1	0.969863	2.1585366	1	2	5.00E+00	0.4	0.0149772	MMMMF	M	M	NonMGY		

<i>Colobus angolensis</i>	0.445205 5	1.232876 7	2.7692308	1	5	7	0.714	0.0148402	OMMF	M	M	NonMGY		
<i>Gorilla gorilla</i>	0.465753 4	3.92	8.4164706	1	10	15	0.667	0.0093151	OMMF	M	F	NonMGY	256	920.35
<i>Homo sapiens</i>	0.767123 3	3.698	4.8206071	1	13	14	0.929	0.009589	MFPB	M	F	NonMGY		
<i>Leontopithecus chrysomelas</i>	0.352054 8	0.675616 4	1.9190661	2	1.5	2	0.75	0.0165284	MFPB	N	B	MGY		
<i>Leontopithecus rosalia</i>	0.363013 7	0.5	1.3773585	2	1.5	2	0.75	0.0165006	MFPB	N	B	MGY	128	75.69
<i>Macaca fascicularis</i>	0.443835 6	1.091643 8	2.4595679	1	4	6.00E+00	0.667	0.0116799	MMMMF	N	M	NonMGY	165	283.53
<i>Macaca mulatta</i>	0.452054 8	1.056164 4	2.3363636	1	3.2	5.7	0.561	0.0125571	MMMMF	N	M	NonMGY	165	304.16
<i>Macaca nemestrina</i>	0.468493 2	1.109780 8	2.3688304	1	3.2	4.20E+00	0.762	0.0173516	MMMMF	M	M	NonMGY	172	292.6
<i>Mandrillus leucophaeus</i>	0.479452 1	1.406657 5	2.9338857	1	3.55	3.5	1.014	0.0167641	MMMMF	M	M	NonMGY	179	486.66
<i>Microcebus murinus</i>	0.167123 3	1	5.9836066	2	1.75	1.083333 3	1.615	0.0111416	OMMF	F	M	NonMGY	61	40.45
<i>Nomascus leucogenys</i>	0.575342 5	2.5	4.3452381	1	6.5	7	0.929	0.0205479	MFPB	F	B	MGY		
<i>Otolemur garnettii</i>	0.356164 4	0.591643 8	1.6611538	1	1	1.67E+00	0.6	0.0237443	MMMMF	F	B	NonMGY	132	139.2
<i>Pan paniscus</i>	0.657534 2	4.7	7.1479167	1	14	1.40E+01	1 8.52E	0.0160374	FF	F	F	NonMGY	232	1081.3
<i>Pan troglodytes</i>	0.630137 0.493150	5 2.062493	7.9347826	1	11.5	1.35E+01	-01	0.0126027	FF	M	F	NonMGY	229	1260.8
<i>Papio anubis</i>	0.493150 7	2.062493 2	4.1822778	1	7.5	8.5	0.882	0.0195695	MMMMF	F	M	NonMGY	178.9	596.6
<i>Pithecia pithecia</i>				1		2.63						NonMGY	146	
<i>Pongo abelii</i>	0.695890 4	8.087671 2	11.622047	1	12.3	19	0.647	0.0126526	SOL	M	M	NonMGY		
<i>Propithecus coquereli</i>	0.443835 6	1	2.2530864	1	3.25	3.25	1	0.0161395	MMMMF	F	M	NonMGY	140.9	180.96
<i>Rhinopithecus roxellana</i>	0.575342 5	1.808219 2	3.1428571	1	4.5	7	0.643	0.0221286	OMMF	M	B	NonMGY	199.3	4
<i>Saguinus bicolor</i>	0.397260 3	0.534246 6	1.3448276	2	1.5	2	0.75	0.0209084	OFMM	N	B	MGY	160	
<i>Saguinus niger</i>	0.383561 6	0.690411	1.8	2	1.75	1.75	1	0.0511416	OFMM	N	B	MGY		
<i>Saimiri boliviensis</i>	0.441095 9	2	4.5341615	1	2.5	5	0.5	0.0245053	MMMMF	M	M	NonMGY	157.7	9

<i>Sapajus apella</i>	0.452054			1	4	7				MMMMF	M	M	NonMGY	155
<i>Sapajus robustus</i>	8	1	2.2121212	1	4	6.00E+00	0.667	0.010274		MMMMF	M	M	NonMGY	
<i>Sapajus xanthosternos</i>	0.452054	1.833333												
	8	3	4.0555556	1	4	6.00E+00	0.667	0.0150685		MMMMF	N	M	NonMGY	

Dipersive Sexual: M = Male, F = Female, N = No sex-biased dispersal. Social Structure: MFPB = Male-Female pair bonds, FF = Fission-fussion, SOL = Solitary, OMMF = One male multiple females, OFMM = One female multiple males, MMMF = Multiple males and multiple females. Dominant sex: M = Male, F = Female, N = No dominant sex. Mating system: MGY = Monogamous, Non-MGY = Non-monogamous. *For the sake of simplicity, all polyandrous/polygynous systems were classified as non-monogamous.

Table S4 – Branch site tests within primate branches.

Primates PRLR (IC)	Parameters estimates Foreground vs. Background	ℓ	<i>P value</i>
<i>M2a_rel</i>	$p_0= 0.62619, p_1= 0.34254, p_2= 0.03127; \omega_0= 0.24295, \omega_1= 1.00000, \omega_2= 2.20596$	-5949.60	
<i>Catarrhini</i>	$p_0= 0.51683; p_1= 0.32245; p_2= 0.16072$ <i>Brackground:</i> $\omega_0= 0.36402; \omega_1= 1.00000; \omega_2= 0.00000$ <i>Foreground:</i> $\omega_0= 0.36402; \omega_1= 1.00000; \omega_2= 0.00000$	-5948.65	>0.999
<i>Platyrrhini</i>	$p_0= 0.51683; p_1= 0.32245; p_2= 0.16072$ <i>Brackground:</i> $\omega_0= 0.36402; \omega_1= 1.00000; \omega_2= 0.00000$ <i>Foreground:</i> $\omega_0= 0.36402; \omega_1= 1.00000; \omega_2= 0.00000$	-5948.65	>0.999
<i>Strepsirrhini</i> *	$p_0= 0.51681; p_1= 0.32254; p_2= 0.16065$ <i>Brackground:</i> $\omega_0= 0.36388; \omega_1= 1.00000; \omega_2= 0.00000$ <i>Foreground:</i> $\omega_0= 0.36388; \omega_1= 1.00000; \omega_2= 0.26841$	-5948.65	>0.999

p = proportion of sites within each omega (ω) class; $\omega < 1$ (negative selection), $\omega = 1$ (neutral selection), and $\omega > 1$ (positive selection).

ℓ = Log Likelihood Values; Likelihood ratio tests were performed between the neutral model (M2a – Neutral model) and the model that admits positive selection (Clade C; $df = 1$; performed for each group: Simiiformes (Catarrhini and Platyrrhini), and Strepsirrhini; * The test performed in Strepsirrhini branch included *Galeopterus variegatus* species).

Table S5 – Primates intraspecific variations within *PRLR* (IC).

Species	Genotypes														Alleles											
	294Ser 882 C>T		394Pro 1182 A>C			Ser422Thr 1265 G>C		Thr444Ala 1330 G>A			490Thr 1470 G>A		Val569Met 1705 A>G		294Ser 882 C>T		394Pro 1182 A>C		Ser422Thr 1265 G>C		Thr444Ala 1330 G>A		490Thr1470 G>A		Val569Met 1705 A>G	
	CC	CT	AA	CC	AC	GG	GC	GG	GA	AA	GG	GA	AA	GG	C	T	A	C	G	C	G	A	G	A	A	G
<i>Callithrix penicillata</i> (n=5)			0.6	0.2	0.2	-	-	-	-	-	-	-	-	-			0.3	0.7	-	-	-	-	-	-	-	-
<i>Callithrix geoffroyi</i> (n=7)	0.9	0.1	-	-	-	0.7	3	0.3	0.4	0.3	0.8	0.2	-	-	0.9	0.1	-	-	0.9	0.1	0.5	0.5	0.9	0.1	-	-
<i>Callicebus nigrifrons</i> (n=3)	-	-	-	-	-	-	-	-	-	-	-	-	0.7	0.3	-	-	-	-	-	-	-	-	-	-	0.7	0.3

**CAPÍTULO V *DADOS ADICIONAIS: GENE DO RECEPTOR DA
PROLACTINA (PRLR) EM MARSUPIAIS***

Abaixo podem ser encontrados dados originais para o gene do receptor de prolactina (*PRLR*) em amostras de marsupiais que foram gerados para a presente Tese e cujas análises estão em andamento. Nosso interesse nesse clado de mamíferos se deu por suas características únicas, como um sofisticado sistema de amamentação, incluindo mudanças na composição do leite para combinar com cada estágio do desenvolvimento do filhote durante todo seu crescimento no marsúpio (NICHOLAS *et al.*, 1997). Por exemplo, o canguru walabi (*Macropus eugenii*), a espécie modelo para investigar estudos sobre a reprodução marsupial, tem uma gravidez de 26,5 dias seguida de um período de lactância de 300-350 dias, enquanto o rato (*Rattus sp.*) tem uma gravidez de 20 dias seguida de um período de aleitamento de 20 a 24 dias (GUERNSEY *et al.*, 2017 e referências lá contidas). Apesar dessas diferenças, Guernsey e colegas (2017) demonstram que a placenta relativamente simples do walabi apresenta um padrão de expressão gênica que é uma reminiscência da observada em mamíferos placentários. Os autores descobriram que a expressão gênica sofre mudanças num dado momento da gestação do walabi, crítico para o crescimento e proliferação da placenta. Esse mesmo tipo de mudança dinâmica na expressão de genes placentários também foi documentada em ratos e humanos, sugerindo que os padrões de expressão de genes que mudam rapidamente estão conservados nas subclasses (Metatheria e Eutheria) de mamíferos Theria (GUERNSEY *et al.*, 2017). Além disso, Guernsey e colegas (2017) sugeriram que os genes reprodutivos foram cooptados durante a evolução, tanto para funções na placenta quanto na glândula mamária, uma estratégia evolutiva bem-sucedida de nutrição do filhote, através da placenta e da mama. Dentre os genes que mostraram sinais de expressão no contexto dos estudos de Guernsey e colegas, estão alguns genes investigados na presente Tese, incluindo *PRL*, *PRLR* e *PRLH*.

Os dados foram obtidos utilizando-se os *primers* PRLR_FOp (GAAGTGGACGACAGCGAAG) e PRLR_ROp (CAAGTCGGATCCAGGTAAT) para gerar um amplicon de 911 pb (TM= 58°C), desenhados a partir de sequência da espécie *Monodelphis domestica* como referência. A conservação das sequências alvo entre marsupiais foi verificada através dos dados disponíveis no UCSC genome browser (<http://genome.ucsc.edu/>). Dados das tabelas 2 e 3 foram obtidos de acordo com metodologias descritas em (NEI & GOJOBORI, 1986) e nos capítulos III e IV da presente Tese, respectivamente. A discussão a respeito dos achados, incluindo comparações com mamíferos placentários, ainda não foram realizadas.

Tabela 1 - Amostras amplificadas para PRLR – IC (E8) em Didelphimorphia e espécies recuperadas nos bancos de dados.

Ordem	Família	Espécies	REFSEQ*	E-value
Diprotodontia	Phalangeridae	<i>Trichosurus vulpecula</i>	AF098296.1	-
Diprotodontia	Phascolarctidae	<i>Phascolarctos cinereus</i>	XM_020995639.1	-
Didelphimorphia	Didelphidae	<i>Monodelphis domestica</i>	XP_007487526.1	4.00E-109
Didelphimorphia	Didelphidae	<i>Monodelphis brevicaudata</i>	1397	-
Didelphimorphia	Didelphidae	<i>Monodelphis glirina</i>	1486	-
Didelphimorphia	Didelphidae	<i>Monodelphis americana</i>	1470	-
Didelphimorphia	Didelphidae	<i>Monodelphis kunsii</i>	1531	-
Didelphimorphia	Didelphidae	<i>Didelphis albiventris</i>	1779	-
Didelphimorphia	Didelphidae	<i>Philander frenatus</i>	951	-
Didelphimorphia	Didelphidae	<i>Metachirus nudicaudatus</i>	1191	-
Didelphimorphia	Didelphidae	<i>Gracilinanus agilis</i>	1363	-

*Referência da sequência recuperada no GenBank (<https://www.ncbi.nlm.nih.gov/nucleotide/>) e/ou da amostra amplificada em laboratório.

Tabela 2 – Distância de Ney & Gojobori calculada para as espécies de marsupiais amostradas.

<u>Espécies</u>	<i>Sarcophilus harrisi</i>	<i>Trichosurus vulpecula</i>	<i>Phascolarctos cinereus</i>	<i>Didelphis albiventris</i>	<i>Philander frenatus</i>	<i>Metachirus nudicaudatus</i>	<i>Gracilinanus agilis</i>	<i>Monodelphis domestica</i>	<i>Monodelphis brevicaudata</i>	<i>Monodelphis glirina</i>	<i>Monodelphis americana</i>	<i>Monodelphis kunyi</i>
<i>Sarcophilus harrisi</i>												
<i>Trichosurus vulpecula</i>	0.267											
<i>Phascolarctos cinereus</i>	0.313	0.408										
<i>Didelphis albiventris</i>	0.212	0.255	0.265									
<i>Philander frenatus</i>	0.244	0.289	0.27	0.211								
<i>Metachirus nudicaudatus</i>	0.259	0.369	0.347	0.407	0.501							
<i>Gracilinanus agilis</i>	0.19	0.275	0.28	0.312	0.365	0.719						
<i>Monodelphis domestica</i>	0.184	0.289	0.241	0.208	0.258	0.36	0.227					
<i>Monodelphis brevicaudata</i>	0.192	0.278	0.255	0.219	0.277	0.401	0.247	0.14				
<i>Monodelphis glirina</i>	0.192	0.265	0.255	0.197	0.245	0.34	0.206	0.069	-1			
<i>Monodelphis americana</i>	0.193	0.262	0.266	0.256	0.286	0.389	0.302	0.275	0.348	0.229		
<i>Monodelphis kunyi</i>	0.194	0.256	0.234	0.221	0.276	0.35	0.203	0.222	0.281	0.166	0.222	

Gradiente de cores do azul ao vermelho, mostrando valores mais baixos aos mais altos respectivamente.

Tabela 3 – Parâmetros estimados para *PRLR* através de diferentes modelos de substituição de códons.

	Modelo	d_N/d_S	Parâmetros estimados	ℓ	Valor <i>p</i>
<i>Marsupiais</i> <i>PRLR</i> (E8)	M1a: neutral	0.4430	$p_0=0.67497, (p_1=0.32503)$ $(\omega_0=0.17477), (\omega_1=1.00000)$	-1480.31	
	M2a: selection	0.4430		-1480.31	M1a vs M2a $P > 0.999$
	M7: β	0.3932	$p_0=0.67497, p_1=0.25773, (p_2=0.06731)$ $(\omega_0=0.17477), (\omega_1=1.00000), \omega_2=1.00000$	-1479.02	
	M8: β and ω	0.3932	$[p=0.59893, q=0.92229]$ $p_0=0.99999, (p_2=0.00001), p=0.59893$ $q=0.92231, \omega_2=1.00000$	-1479.02	M7 vs M8 $P > 0.999$

p = proporção de sítios dentro de cada classe omega (ω); $\omega < 1$ (seleção negativa), $\omega_1 = 1$ (neutralidade) e $\omega_2 > 1$ (seleção positiva). ℓ = Valores de Log Likelihood. Os testes de razão de verossimilhança foram realizados entre os modelos neutros (M1a- quase neutro e M7 - beta) e modelos que admitem seleção positiva (seleção M2a e seleção M8, β e ω - beta + seleção). As comparações entre modelos de seleção neutros e positivos apresentaram 2 graus de liberdade. Entre parênteses: parâmetros fixos; Entre parênteses: parâmetros beta p e q . d_N / d_S = taxa de mudanças não sinônimas/ sinônimas.

CAPÍTULO VI *DADOS ADICIONAIS: VARIABILIDADE DO*
HORMÔNIO LIBERADOR DE PROLACTINA (PRLH) EM PRIMATAS DO
NOVO MUNDO

Como visto anteriormente, existem diferenças marcantes entre mamíferos placentários e marsupiais, mas estudos recentes mostram também similaridades instigantes relacionadas à expressão de genes com funções na reprodução (GUERNSEY *et al.*, 2017). Para ampliar nosso conhecimento a respeito, resolvemos investigar também a variação do gene liberador de prolactina (*PRLH*) em amostras de macacos do Novo Mundo (NWM). Os dados foram obtidos utilizando-se os seguintes *primers*: (Exon 1 *forw*: PRLH1_F: CTGATGCCTGTGTGACGCTG/ PRLH2_F: CTGATGCCTGTGTAACCCTG; Rev: PRLH1_R GCTTTGGGAAGCAGGAGTTC; TM = 62°C/ Exon 2 *forw*: E2_PRLH: GAACTCCTGCTTCCCAAAGC; PRLH2_R AGGATTGAAGGGAAAGGAGC; TM = 62°C), desenhados utilizando a espécie *Callithrix jacchus* como referência. A conservação das sequências foi avaliada através dos dados de espécies disponíveis no UCSC genome browser (<http://genome.ucsc.edu/>). Dados da tabela 3 foram obtidos de acordo com metodologias descritas nos capítulos III e IV da presente tese, respectivamente. Discussão a respeito dos achados, incluindo comparações com outras categorias taxonômicas, ainda não foram realizadas.

Tabela 1 - Amostras amplificadas para o gene *PRLH* em Primatas.

Amostra	Nome	Exon 1	Exon 2
M2160	<i>Brachyteles arachnoides</i>	V	X
M2072	<i>Cacajao melanocephalus</i>	V	V
M2201	<i>Callibella humilis</i>	V	X
M2275	<i>Callicebus coimbrai</i>	V	V
M2074	<i>Cebuella pygmaea</i>	V	X
M2209	<i>Sapajus robustus</i>	X	V
M2094	<i>Chiropotes albinasus</i>	X	V
M2095	<i>Chiropotes uthaikaie</i>	X	V
M1859	<i>Leontopithecus chrysomelas</i>	V	V
M1620	<i>Leontopithecus rosalia</i>	V	V
M1856	<i>Callithrix humeralifer</i>	V	V
M1942	<i>Callithrix melanura</i>	V	V
M2260	<i>Callithrix saterei</i>	V	V
M1977	<i>Saguinus niger</i>	V	V

Tabela 2 – Sequências recuperadas para *PRLH* em Primatas.

Espécies	Referência*
<i>Aotus nancymae</i>	XM_012434024.1 ^N
<i>Callithrix jacchus</i>	XM_008999759.2 ^N
<i>Cebus capucinus</i>	XM_017537749.1 ^N
<i>Chlorocebus sabaesus</i>	A0A0D9R1E3 ^E
<i>Gorilla gorilla gorilla</i>	XM_019022703.1 ^N
<i>Homo sapiens</i>	NM_015893.1 ^N
<i>Macaca mulatta</i>	XM_001084860.2 ^N
<i>Nomascus leucogenys</i>	G1RJK6 ^E
<i>Otolemur garnettii</i>	H0Y186 ^E
<i>Pan paniscus</i>	XM_003813105.2 ^N
<i>Pan troglodytes</i>	XM_016950786.1 ^N
<i>Papio anubis</i>	XM_017947198.1 ^N
<i>Pongo abelii</i>	XM_002813037.1 ^N
<i>Rhinopithecus bieti</i>	XM_017890283.1 ^N
<i>Rhinopithecus roxellana</i>	XM_010357982.1 ^N
<i>Saimiri boliviensis boliviensis</i>	XM_010347318.1 ^N
<i>Carlito syrichta</i>	XM_008054089.1 ^N

*Banco de dado: ^N ; NCBI, ^E Ensembl

Tabela 3 – Padrões evolutivos para *PRLH* em Primatas.

	Modelo	d_N/d_S	Parâmetros estimados	ℓ	Valor p	
Peptídeo sinal	<i>M1a</i> : neutro	0.1410	$p_0=0.99999,(p_1=0.00001)$	-235.75	<i>M1a vs M2a P</i> > 0.999	
	<i>M2a</i> : seleção	0.1396	$(\omega_0=0.14100),(\omega_1=1.00000)$ $p_0=1.00000,p_1=0.00000,(p_2=0.00000)$ $(\omega_0=0.13964),(\omega_1=1.00000),\omega_2=1.00000$	-235.45		
	<i>M7</i> : β	0.1428	$[p=16.52424,q=99.00000]$	-235.77	<i>M7 vs M8 P</i> > 0.999	
	<i>M8</i> : β e ω	0.2232	$p_0=0.99999,(p_2=0.00001),p=16.52426$ $q=2.66488,\omega_2=1.00000$	-235.77		
	Peptídeo	<i>M1a</i> : neutro	0.4272	$p_0=0.64241,(p_1=0.35759)$	-429.89	<i>M1a vs M2a P</i> > 0.999
		<i>M2a</i> : seleção	0.6968	$(\omega_0=0.10832),(\omega_1=1.00000)$ $p_0=0.63504,p_1=0.27965,(p_2=0.08530)$ $(\omega_0=0.13162),(\omega_1=1.00000),\omega_2=3.90992$	-428.23	
<i>M7</i> : β		0.4026	$[p=0.27949,q=0.41469]$	-430.14	<i>M7 vs M8 P</i> > 0.999	
<i>M8</i> : β e ω		0.6052	$p_0=0.90081,(p_2=0.09919),p=0.58982$ $q=1.34743,\omega_2=3.35011$	-427.49		

p = proporção de sítios dentro de cada classe omega (ω); $\omega < 1$ (seleção negativa), $\omega = 1$ (neutralidade) e $\omega > 1$ (seleção positiva). ℓ = Valores de Log Likelihood. Os testes de razão de verossimilhança foram realizados entre os modelos neutros (*M1a*- quase neutro e *M7* - beta) e modelos que admitem seleção positiva (seleção *M2a* e seleção *M8*, β e ω - beta + seleção). As comparações entre modelos de seleção neutros e positivos apresentaram 2 graus de liberdade. Entre parênteses: parâmetros fixos; Entre parênteses: parâmetros beta p e q . d_N / d_S = taxa de mudanças não sinônimas/ sinônimas.

CAPÍTULO VII *CONCLUSÃO*

Os capítulos III e IV da presente Tese apresentam discussões específicas referentes aos resultados obtidos aqui. Já os achados brevemente apresentados nos capítulos V e VI, e que não serão discutidos aqui, foram introduzidos ao corpo da presente Tese com o objetivo de salientar o esforço empregado para que mais informações acerca dos genes da rede *OXT-AVP-PRL* sejam conhecidas em espécies na fauna brasileira, pouco ou nada representada nos bancos de dados públicos, incluindo também espécies ameaçadas de extinção. Além disso, devido à natureza dos genes em questão, buscou-se representar grupos de espécies de categorias taxonômicas diferentes quanto a traços reprodutivos (fisiológicos e comportamentais) como é o caso dos primatas e marsupiais. Desse modo, pode-se antever que mais artigos serão redigidos a partir do trabalho iniciado na presente Tese.

Visto o exposto acima, o presente Capítulo apresenta somente uma discussão e conclusão geral sobre os assuntos estudados e resultados obtidos. Para isso, quando oportuno, alguns achados apresentados nos anexos também serão comentados.

Um dos maiores desafios da atualidade é conseguir conectar variações em nível de genoma com aquelas em nível de fenótipos. Além disso, conectar esses achados com cenários evolutivos torna esse tipo de estudo e abordagem altamente desafiadores. Nosso grupo de pesquisa tem buscado contribuir com essa área de conhecimento em especial considerando o sistema *OXT-AVP* e seu papel na emergência de fenótipos adaptativos em primatas do Novo Mundo (anexos I, II e III; VARGAS-PINILLA et al., 2015, 2016; PARREIRAS-&-SILVA et al., 2017). Interessantemente, também verificamos que variações em genes desse mesmo sistema estavam implicadas na emergência de fenótipos ligados à domesticação em diferentes clados de mamíferos (FAM et al., 2018, no prelo; anexo V). No entanto, é esperado que fenótipos de natureza complexa, como por exemplo, traços reprodutivos, resultem da ação (e interação) de uma rede de genes, compreendendo sistemas interligados. Assim sendo, nossos estudos se iniciam com a rede funcional da *OXT-AVP* (*OXT*, *OXTR*, *AVP*, *AVPR1A*, *AVPR1B* e *AVPR2*) e aqui ampliamos para a rede da prolactina (*PRL*), inicialmente com dados originais do receptor *PRLR* em macacos do Novo Mundo (cujo manuscrito já se encontra em sua primeira versão; Capítulo IV) e marsupiais (Capítulo V).

Num primeiro momento, os receptores de *OXT* e *AVP* (*AVPR1a*, *AVPR1b*, *AVPR2* e *OXTR*) foram explorados mais exaustivamente através de diferentes abordagens *in silico*, conforme pode ser apreciado no Capítulo III da presente Tese (PARÉ et al., 2016). Esse estudo descreve os padrões evolutivos dessa família num conjunto de mamíferos placentários, incluindo a delimitação de um perfil de distribuição do padrão de desordem de proteínas ao longo da evolução, que se correlaciona positivamente com maiores valores de ω . Além disso,

investiga os padrões de interação através do SLiMs encontrados, sugerindo uma mudança adaptativa observada na espécie *Dipodomys ordii* (PARÉ *et al.*, 2016). Nessa espécie adaptada a ambientes áridos, foi observada a fixação de um motivo linear de interação para proteínas MAPKs (*Mitogen-Activated Protein Kinases*) em AVPR2. Essas proteínas são envolvidas em diversas respostas a estímulos celulares, incluindo estresse osmótico. Uma vez que AVPR2 promove a homeostase da água e *Dipodomys ordii* vive em ambientes limitados quanto à disponibilidade desse recurso, foi sugerido que essa mudança tenha prováveis implicações adaptativas. Esse achado salienta a importância de avaliar a alteração não sinônima num contexto maior, considerando também o impacto desta na interação com outros elementos genéticos.

Num segundo momento, genes relacionados à PRL foram explorados em diferentes análises. Inicialmente, a região codificadora do receptor PRLR foi sequenciada em dois grupos taxonômicos distintos: macacos do Novo Mundo e Didelphimorphia (Capítulos IV, “*Molecular evolution of the Prolactin Receptor and implications for adaptive traits in primates*”; e V, resultados preliminares). No capítulo IV, descrevemos a alta variabilidade de *PRLR* em primatas, no qual mudanças especificamente encontradas em NWM apresentaram uma alta probabilidade de estarem sob seleção positiva, o que pode estar ligado à emergência de fenótipos específicos, como a diminuição do tempo de maturidade sexual em fêmeas e o tamanho da prole. Além disso, avaliamos se as mudanças de aminoácidos alteravam a sinalização e interação entre moléculas através de predição de motivos lineares (SLiMs). Por exemplo, descrevemos que nos Simiiformes (macacos do Velho e Novo Mundo), há um motivo linear adicional de interação com domínios SH3 (Src homology 3). Os domínios SH3 e os seus locais de ligação foram descritos para centenas de proteínas; eles fornecem à célula um meio particularmente conveniente e adaptável de interação seletiva proteína-proteína, que pode ser de importância funcional (MAYER, 2015). Essas variações constituem-se num interessante repertório a ser avaliado em futuros ensaios experimentais, por exemplo, como aquele realizado por nosso grupo de pesquisa (PARREIRAS-&-SILVA *et al.*, 2017; anexo III). A alta variabilidade de PRLR (IC) em primatas havia sido relatada anteriormente (BABB *et al.*, 2014), sendo essa região a mais variável da proteína, por isso foi considerada no planejamento de nosso estudo. No entanto, Babb *et al.* (2014) não sugere sítios específicos sob seleção positiva, diferente do que foi evidenciado aqui, particularmente relacionado à forma longa do receptor em associação com características específicas.

Desse modo, podemos concluir dizendo: Um dos maiores desafios para as Ciências Biomédicas na atualidade é conseguir conectar variações em nível de genoma com aquelas observadas em nível de fenótipos. Normalmente, o foco das atenções são populações humanas e a identificação de variantes em genes que possam estar correlacionadas com a diversidade normal e patológica dentro de nossa espécie (BRADLEY & LAWLER, 2011). De menos interesse, mas não menos relevante e desafiador, são as investigações que buscam desvendar o perfil genético por trás de traços fenotípicos que caracterizam uma espécie como um todo, ou mesmo um grupo taxonômico maior (gênero, família, ordem, etc). Em outras palavras, como um gene (ou um grupo deles) e suas variantes se conectam com fenótipos adaptativos em populações naturais? Existem várias ferramentas metodológicas disponíveis para responder à pergunta feita acima, mas o ponto de partida necessariamente é aquele que gera o dado genético original na espécie sob investigação. Para isso, basicamente, duas abordagens iniciais são geralmente utilizadas: a) caracterização do genoma completo; e b) caracterização de gene (s) candidato (s). Os estudos do tipo (b), naturalmente mais econômicos, focam na investigação de um ou mais genes (ou regiões genômicas) de uma determinada rede cujas funções e implicações para um determinado fenótipo (ou um grupo de fenótipos correlacionados) já são previamente bem conhecidas. A estratégia de genes candidatos vem sendo utilizada com sucesso por nosso grupo de pesquisa como mostra a presente Tese, visto que temos conseguido investigar a complexa conexão entre um determinado repertório genético e fenótipos adaptativos correspondentes. Desdobramentos a partir de nossos achados envolvem necessariamente novos estudos funcionais e a investigação de mais genes da rede OXT- AVP- PRL, considerando sua porção codificadora e regulatória, o que não é nada trivial. Mas desafios assim nos movem!

REFERÊNCIAS

- ACHER, R.; CHAUVET, J.; CHAUVET, M. T. Man and the chimaera. Selective versus neutral oxytocin evolution. **Adv Exp Med Biol**, v. 395, p. 615-27, 1995. ISSN 0065-2598. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/8714026> >.
- ADÁN, N. et al. Arthritis and prolactin: a phylogenetic viewpoint. **Gen Comp Endocrinol**, v. 203, p. 132-6, Jul 2014. ISSN 1095-6840. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/24508497> >.
- AKEN, B. L. et al. The Ensembl gene annotation system. **Database (Oxford)**, v. 2016, 2016. ISSN 1758-0463. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/27337980> >.
- AMEMIYA, Y. et al. Somatolactin in the white sturgeon and African lungfish and its evolutionary significance. **Gen Comp Endocrinol**, v. 114, n. 2, p. 181-90, May 1999. ISSN 0016-6480. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/10208767> >.
- AMRINE-MADSEN, H. et al. Nuclear gene sequences provide evidence for the monophyly of australidelphian marsupials. **Mol Phylogenet Evol**, v. 28, n. 2, p. 186-96, Aug 2003. ISSN 1055-7903. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/12878458> >.
- ARISTIDE, L. et al. Modeling lineage and phenotypic diversification in the New World monkey (Platyrrhini, Primates) radiation. **Mol Phylogenet Evol**, v. 82 Pt B, p. 375-85, Jan 2015. ISSN 1095-9513. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/24287474> >.
- ASTÚA, D. et al. First evidence of gregarious denning in opossums (Didelphimorphia, Didelphidae), with notes on their social behaviour. **Biol Lett**, v. 11, n. 6, p. 20150307, Jun 2015. ISSN 1744-957X. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/26085500> >.
- BABB, P. L. et al. **Prolactin Receptor Gene Diversity in Azara's Owl Monkeys (*Aotus azarai*) and Humans (*Homo sapiens*) Suggests a Non-Neutral Evolutionary History among Primates**. International Journal of Primatology, p.129–155. 2014
- BANDEL, H. J.; FORSTER, P.; RÖHL, A. Median-joining networks for inferring intraspecific phylogenies. **Mol Biol Evol**, v. 16, n. 1, p. 37-48, Jan 1999. ISSN 0737-4038. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/10331250> >.
- BARBERIS, C.; MOUILLAC, B.; DURROUX, T. Structural bases of vasopressin/oxytocin receptor function. **J Endocrinol**, v. 156, n. 2, p. 223-9, Feb 1998. ISSN 0022-0795. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/9518866> >.
- BARGIOTA, S. I. et al. The Effects of Antipsychotics on Prolactin Levels and Women's Menstruation. **Schizophr Res Treatment**, v. 2013, p. 502697, 2013. ISSN 2090-2085. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/24490071> >.

BEN-JONATHAN, N.; HUGO, E. Prolactin (PRL) in adipose tissue: regulation and functions. **Adv Exp Med Biol**, v. 846, p. 1-35, 2015. ISSN 0065-2598. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/25472532> >.

BENSON, D. A. et al. GenBank. **Nucleic Acids Res**, v. 42, n. Database issue, p. D32-7, Jan 2014. ISSN 1362-4962. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/24217914> >.

BOCKAERT, J.; PIN, J. P. Molecular tinkering of G protein-coupled receptors: an evolutionary success. **EMBO J**, v. 18, n. 7, p. 1723-9, Apr 1999. ISSN 0261-4189. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/10202136> >.

BOGGON, T. J.; ECK, M. J. Structure and regulation of Src family kinases. **Oncogene**, v. 23, n. 48, p. 7918-27, Oct 2004. ISSN 0950-9232. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/15489910> >.

BOLE-FEYSOT, C. et al. Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. **Endocr Rev**, v. 19, n. 3, p. 225-68, Jun 1998. ISSN 0163-769X. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/9626554> >.

BRADLEY, B. J.; LAWLER, R. R. Linking genotypes, phenotypes, and fitness in wild primate populations. **Evol Anthropol**, v. 20, n. 3, p. 104-19, 2011 May-Jun 2011. ISSN 1520-6505. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/22034168> >.

BRIDGES, R. S. Neuroendocrine regulation of maternal behavior. **Front Neuroendocrinol**, v. 36, p. 178-96, Jan 2015. ISSN 1095-6808. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/25500107> >.

CHRISTOFFELS, A. et al. Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. **Mol Biol Evol**, v. 21, n. 6, p. 1146-51, Jun 2004. ISSN 0737-4038. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15014147> >.

CLEVINGER, C. V. et al. The role of prolactin in mammary carcinoma. **Endocr Rev**, v. 24, n. 1, p. 1-27, Feb 2003. ISSN 0163-769X. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/12588805> >.

CLEVINGER, C. V.; MEDAGLIA, M. V. The protein tyrosine kinase P59fyn is associated with prolactin (PRL) receptor and is activated by PRL stimulation of T-lymphocytes. **Mol Endocrinol**, v. 8, n. 6, p. 674-81, Jun 1994. ISSN 0888-8809. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/7935483> >.

UNIPROT CONSORTIUM, U. The Universal Protein Resource (UniProt) in 2010. **Nucleic Acids Res**, v. 38, n. Database issue, p. D142-8, Jan 2010. ISSN 1362-4962. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19843607> >.

COOKE, N. E. et al. Human prolactin. cDNA structural analysis and evolutionary comparisons. **J Biol Chem**, v. 256, n. 8, p. 4007-16, Apr 1981. ISSN 0021-9258. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/6260780> >.

COSTA-NETO, C. M.; PARREIRAS-E-SILVA, L. T.; BOUVIER, M. A Pluridimensional View of Biased Agonism. **Mol Pharmacol**, v. 90, n. 5, p. 587-595, Nov 2016. ISSN 1521-0111. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/27638872> >.

DALE, H. H. On some physiological actions of ergot. **J Physiol**, v. 34, n. 3, p. 163-206, May 1906. ISSN 0022-3751. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16992821> >.

DINKEL, H. et al. The eukaryotic linear motif resource ELM: 10 years and counting. **Nucleic Acids Res**, v. 42, n. Database issue, p. D259-66, Jan 2014. ISSN 1362-4962. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24214962> >.

WILSON, D. E., REEDER, D. M. Mammal Species of the World. A Taxonomic and Geographic Reference (3rd ed), Johns Hopkins University Press, 2,142 pp. (Available from Johns Hopkins University Press, 1-800-537-5487 or (410) 516-6900, 2005 or at <http://www.press.jhu.edu>).

DUNCAN, L. E.; KELLER, M. C. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. **Am J Psychiatry**, v. 168, n. 10, p. 1041-9, Oct 2011. ISSN 1535-7228. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/21890791> >.

DUSANTER-FOURT, I. et al. Identification of JAK protein tyrosine kinases as signaling molecules for prolactin. Functional analysis of prolactin receptor and prolactin-erythropoietin receptor chimera expressed in lymphoid cells. **EMBO J**, v. 13, n. 11, p. 2583-91, Jun 1994. ISSN 0261-4189. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/8013458> >.

EXCOFFIER, L.; LISCHER, H. E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. **Mol Ecol Resour**, v. 10, n. 3, p. 564-7, May 2010. ISSN 1755-0998. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/21565059> >.

FARON-GÓRECKA, A. et al. Involvement of prolactin and somatostatin in depression and the mechanism of action of antidepressant drugs. **Pharmacol Rep**, v. 65, n. 6, p. 1640-6, 2013. ISSN 1734-1140. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24553012> >.

FERGUSON, D. R. The genetic distribution of vasopressins in the peccary (*Tayassu angulatus*) and Warthog (*Phacochoerus aethiopicus*). **Gen Comp Endocrinol**, v. 12, n. 3, p. 609-13, Jun 1969. ISSN 0016-6480. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/5769950> >.

FLEMING, A. S. et al. Testosterone and prolactin are associated with emotional responses to infant cries in new fathers. **Horm Behav**, v. 42, n. 4, p. 399-413, Dec 2002. ISSN 0018-506X. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/12488107> >.

FRENCH, J. A. et al. Neuropeptide diversity and the regulation of social behavior in New World primates. **Front Neuroendocrinol**, Mar 2016. ISSN 1095-6808. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/27020799> >.

FUJII, R. et al. Tissue distribution of prolactin-releasing peptide (PrRP) and its receptor. **Regul Pept**, v. 83, n. 1, p. 1-10, Aug 1999. ISSN 0167-0115. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/10498338> >.

GABOR, C. S. et al. Interplay of oxytocin, vasopressin, and sex hormones in the regulation of social recognition. **Behav Neurosci**, v. 126, n. 1, p. 97-109, Feb 2012. ISSN 1939-0084. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22141469> >.

GIMPL, G.; FAHRENHOLZ, F. The oxytocin receptor system: structure, function, and regulation. **Physiol Rev**, v. 81, n. 2, p. 629-83, Apr 2001. ISSN 0031-9333. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/11274341> >.

GOFFIN, V.; KELLY, P. A. The prolactin/growth hormone receptor family: structure/function relationships. **J Mammary Gland Biol Neoplasia**, v. 2, n. 1, p. 7-17, Jan 1997. ISSN 1083-3021. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/10887515> >.

GREEN, C. D. HEREDITARY TALENT AND CHARACTER. <http://psychclassics.yorku.ca/Galton/talent.htm>, p. *Classics in the History of Psychology*, 2017. Acesso em: 14/06/2016.

GUERNSEY, M. W. et al. Molecular conservation of marsupial and eutherian placentation and lactation. **Elife**, v. 6, Sep 2017. ISSN 2050-084X. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/28895534> >.

GRUBER, C. W. Physiology of invertebrate oxytocin and vasopressin neuropeptides. *Exp Physiol*, v. 99, n. 1, p. 55-61, Jan 2014. ISSN 1469-445X. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23955310> >.

GWEE, P. C. et al. Characterization of the neurohypophysial hormone gene loci in elephant shark and the Japanese lamprey: origin of the vertebrate neurohypophysial hormone genes. **BMC Evol Biol**, v. 9, p. 47, 2009. ISSN 1471-2148. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19243634> >.

HIGASHIDA, H. et al. CD38 and its role in oxytocin secretion and social behavior. **Horm Behav**, v. 61, n. 3, p. 351-8, Mar 2012. ISSN 1095-6867. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22227279> >.

HORSEMAN, N. D.; GREGERSON, K. A. Prolactin actions. **J Mol Endocrinol**, v. 52, n. 1, p. R95-106, Feb 2014. ISSN 1479-6813. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24130130> >.

HORSEMAN, N. D. et al. Defective mammopoiesis, but normal hematopoiesis, in mice with a targeted disruption of the prolactin gene. **EMBO J**, v. 16, n. 23, p. 6926-35, Dec 1997. ISSN 0261-4189. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/9384572> >.

HOYLE, C. H. Neuropeptide families: evolutionary perspectives. **Regul Pept**, v. 73, n. 1, p. 1-33, Jan 1998. ISSN 0167-0115. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/9537670> >.

IBATA, Y. et al. Morphological survey of prolactin-releasing peptide and its receptor with special reference to their functional roles in the brain. **Neurosci Res**, v. 38, n. 3, p. 223-30, Nov 2000. ISSN 0168-0102. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/11070188> >.

KAMILAR, J. M.; COOPER, N. Phylogenetic signal in primate behaviour, ecology and life history. **Philos Trans R Soc Lond B Biol Sci**, v. 368, n. 1618, p. 20120341, May 2013. ISSN 1471-2970. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/23569289> >.

KENT, W. J. et al. The human genome browser at UCSC. **Genome Res**, v. 12, n. 6, p. 996-1006, Jun 2002. ISSN 1088-9051. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/12045153> >.

KHONG, H. K. et al. Prolactin receptor mRNA is upregulated in discus fish (*Symphysodon aequifasciata*) skin during parental phase. **Comp Biochem Physiol B Biochem Mol Biol**, v. 153, n. 1, p. 18-28, May 2009. ISSN 1879-1107. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/19272315> >.

KOEHBACH, J. et al. Insights into the molecular evolution of oxytocin receptor ligand binding. **Biochem Soc Trans**, v. 41, n. 1, p. 197-204, Feb 2013. ISSN 1470-8752. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23356283> >.

KOSHIMIZU, T. A. et al. Vasopressin V1a and V1b receptors: from molecules to physiological systems. **Physiol Rev**, v. 92, n. 4, p. 1813-64, Oct 2012. ISSN 1522-1210. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23073632> >.

KUCKA, M. et al. The role of cyclic nucleotides in pituitary lactotroph functions. **Front Endocrinol (Lausanne)**, v. 4, p. 122, 2013. ISSN 1664-2392. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24062725> >.

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Mol Biol Evol**, v. 33, n. 7, p. 1870-4, Jul 2016. ISSN 1537-1719. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/27004904> >.

LAGMAN, D. et al. The vertebrate ancestral repertoire of visual opsins, transducin alpha subunits and oxytocin/vasopressin receptors was established by duplication of their shared genomic region in the two rounds of early vertebrate genome duplications. **BMC Evol Biol**, v. 13, p. 238, 2013. ISSN 1471-2148. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24180662> >.

LANGMEAD, C. J. et al. Characterization of the binding of [(125)I]-human prolactin releasing peptide (PrRP) to GPR10, a novel G protein coupled receptor. **Br J Pharmacol**, v. 131, n. 4, p. 683-8, Oct 2000. ISSN 0007-1188. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/11030716> >.

LEE, A. G. et al. A novel form of oxytocin in New World monkeys. **Biol Lett**, v. 7, n. 4, p. 584-7, Aug 2011. ISSN 1744-957X. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21411453> >.

LEE, H. J. et al. Oxytocin: the great facilitator of life. **Prog Neurobiol**, v. 88, n. 2, p. 127-51, Jun 2009. ISSN 1873-5118. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19482229> >.

LESUEUR, L. et al. Comparison of long and short forms of the prolactin receptor on prolactin-induced milk protein gene transcription. **Proc Natl Acad Sci U S A**, v. 88, n. 3, p. 824-8, Feb 1991. ISSN 0027-8424. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/1992473> >.

LUCAS, B. K. et al. Null mutation of the prolactin receptor gene produces a defect in maternal behavior. **Endocrinology**, v. 139, n. 10, p. 4102-7, Oct 1998. ISSN 0013-7227. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/9751488> >.

LUO, Z. X. et al. A Jurassic eutherian mammal and divergence of marsupials and placentals. **Nature**, v. 476, n. 7361, p. 442-5, Aug 2011. ISSN 1476-4687. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/21866158> >.

LÜ, A. et al. Novel SNPs of the bovine PRLR gene associated with milk production traits. **Biochem Genet**, v. 49, n. 3-4, p. 177-89, Apr 2011. ISSN 1573-4927. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/21165768> >.

MALUKIEWICZ, J. et al. Hybridization effects and genetic diversity of the common and black-tufted marmoset (*Callithrix jacchus* and *Callithrix penicillata*) mitochondrial control region. **Am J Phys Anthropol**, v. 155, n. 4, p. 522-36, Dec 2014. ISSN 1096-8644. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/25186076> >.

MARANO, R. J.; BEN-JONATHAN, N. Minireview: Extrapituitary prolactin: an update on the distribution, regulation, and functions. **Mol Endocrinol**, v. 28, n. 5, p. 622-33, May 2014. ISSN 1944-9917. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24694306> >.

MAYER, B. J. The discovery of modular binding domains: building blocks of cell signalling. **Nat Rev Mol Cell Biol**, v. 16, n. 11, p. 691-8, 11 2015. ISSN 1471-0080. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/26420231> >.

MCCALL, C.; SINGER, T. The animal and human neuroendocrinology of social cognition, motivation and behavior. **Nat Neurosci**, v. 15, n. 5, p. 681-8, May 2012. ISSN 1546-1726. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22504348> >.

MCNEILLY, A. S. et al. Plasma prolactin concentrations during the ovarian cycle and lactation and their relationship to return of fertility post partum in the common marmoset (*Callithrix jacchus*). **J Reprod Fertil**, v. 62, n. 2, p. 353-60, Jul 1981. ISSN 0022-4251. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/6788944> >.

MITCHELL, T. J.; JOHN, S. Signal transducer and activator of transcription (STAT) signalling and T-cell lymphomas. **Immunology**, v. 114, n. 3, p. 301-12, Mar 2005. ISSN 0019-2805. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/15720432> >.

MONASTERIO, N.; VERGARA, E.; MORALES, T. Hormonal influences on neuroimmune responses in the CNS of females. **Front Integr Neurosci**, v. 7, p. 110, 2013. ISSN 1662-5145. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/24478642> >.

NADEAU, J. H. Modifier genes in mice and humans. **Nat Rev Genet**, v. 2, n. 3, p. 165-74, Mar 2001. ISSN 1471-0056. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/11256068> >.

NEDUVA, V.; RUSSELL, R. B. Linear motifs: evolutionary interaction switches. **FEBS Lett**, v. 579, n. 15, p. 3342-5, Jun 2005. ISSN 0014-5793. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15943979> >.

NEI, M.; GOJOBORI, T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. **Mol Biol Evol**, v. 3, n. 5, p. 418-26, Sep 1986. ISSN 0737-4038. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/3444411> >.

NICHOLAS, K. et al. The tammar wallaby: a model to study putative autocrine-induced changes in milk composition. **J Mammary Gland Biol Neoplasia**, v. 2, n. 3, p. 299-310, Jul 1997. ISSN 1083-3021. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/10882313> >.

NIELSEN, R.; YANG, Z. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. **Genetics**, v. 148, n. 3, p. 929-36, Mar 1998. ISSN 0016-6731. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/9539414> >.

NORTON, T. W. Ecology of greater gliders, *petauroides volans kerr* 1792, in relation to variations in habitat quality in eucalypt forests in south-east new south wales. 1988. 222 (Doctor). Department of Forestry, The Australian National University

OHNISHI, H. et al. Src family kinases: modulators of neurotransmitter receptor function and behavior. **Trends Neurosci**, v. 34, n. 12, p. 629-37, Dec 2011. ISSN 1878-108X. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/22051158> >.

ORMANDY, C. J. et al. Mouse prolactin receptor gene: genomic organization reveals alternative promoter usage and generation of isoforms via alternative 3'-exon splicing. **DNA Cell Biol**, v. 17, n. 9, p. 761-70, Sep 1998. ISSN 1044-5498. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/9778035> >.

PARREIRAS-E-SILVA, L. T. et al. Functional New World monkey oxytocin forms elicit an altered signaling profile and promotes parental care in rats. **Proc Natl Acad Sci U S A**, Aug 2017. ISSN 1091-6490. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/28784762> >.

PARSONS, S. J.; PARSONS, J. T. Src family kinases, key regulators of signal transduction. **Oncogene**, v. 23, n. 48, p. 7906-9, Oct 2004. ISSN 0950-9232. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/15489908> >.

PARÉ, P. et al. Oxytocin and arginine vasopressin receptor evolution: implications for adaptive novelties in placental mammals. **Genet Mol Biol**, v. 0, p. 0, Aug 2016. ISSN 1415-4757. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/27505307> >.

PATIL, M. J.; HENRY, M. A.; AKOPIAN, A. N. Prolactin receptor in regulation of neuronal excitability and channels. **Channels (Austin)**, v. 8, n. 3, p. 193-202, 2014. ISSN 1933-6969. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/24758841> >.

PERELMAN, P. et al. A molecular phylogeny of living primates. **PLoS Genet**, v. 7, n. 3, p. e1001342, Mar 2011. ISSN 1553-7404. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21436896> >.

PLOMIN, R. et al. **Behavioral genetics : a primer**. Sixth edition. 2000. ISBN 9781429242158 (hardback) : 151.99
1429242159 (hbk.) : 151.99.

POND, S. L.; FROST, S. D.; MUSE, S. V. HyPhy: hypothesis testing using phylogenies. **Bioinformatics**, v. 21, n. 5, p. 676-9, Mar 2005. ISSN 1367-4803. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/15509596> >.

POZZI, L. et al. Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes. **Mol Phylogenet Evol**, v. 75, p. 165-83, Jun 2014. ISSN 1095-9513. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/24583291> >.

PUTNOVA, L. et al. **A new Hpa II PCR-RFLP within the porcine prolactin receptor (PRLR) gene and study of its effect on litter size and number of teats.** Animal breeding and Genetics. 119: 7 p. 2002.

R CORE TEAM. **R: A language and environment for statistical computing**. Foundation for Statistical Computing. Vienna, Austria. 2017.

RAMAN, A. S.; WHITE, K. I.; RANGANATHAN, R. Origins of Allosterity and Evolvability in Proteins: A Case Study. **Cell**, v. 166, n. 2, p. 468-480, Jul 2016. ISSN 1097-4172. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/27321669> >.

REN, D. et al. Genetic diversity in oxytocin ligands and receptors in new world monkeys. **PLoS One**, v. 10, n. 5, p. e0125775, 2015. ISSN 1932-6203. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25938568> >.

RENS, B. V.; GJ, E.; T, V. D. L. **Components of litter size in gilts with different prolactin receptor genotypes.** Theriogenology. 59: 26 p. 2003.

ROZAS, J. et al. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. **Mol Biol Evol**, v. 34, n. 12, p. 3299-3302, Dec 2017. ISSN 1537-1719. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/29029172> >.

SAMBROOK, J.; RUSSELL, D. W. Purification of nucleic acids by extraction with phenol:chloroform. **CSH Protoc**, v. 2006, n. 1, Jun 2006. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/22485786> >.

SCHNEIDER, H.; SAMPAIO, I. The systematics and evolution of New World primates - A review. **Mol Phylogenet Evol**, v. 82 Pt B, p. 348-57, Jan 2015. ISSN 1095-9513. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/24201058> >.

SCHRADIN, C. et al. Prolactin and paternal care: comparison of three species of monogamous new world monkeys (*Callicebus cupreus*, *Callithrix jacchus*, and *Callimico goeldii*). **J Comp Psychol**, v. 117, n. 2, p. 166-75, Jun 2003. ISSN 0735-7036. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/12856787> >.

SIEHLER, S.; MILLIGAN, G. **G Protein-Coupled Receptors**. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi, Dubai, Tokyo, Mexico City: Cambridge University Press, 2011. 430 ISBN 9780521112086.

SKOGLUND, P.; GÖTHERSTRÖM, A.; JAKOBSSON, M. Estimation of population divergence times from non-overlapping genomic sequences: examples from dogs and wolves. **Mol Biol Evol**, v. 28, n. 4, p. 1505-17, Apr 2011. ISSN 1537-1719. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/21177316> >.

STOOP, R. Neuromodulation by Oxytocin and Vasopressin. **Neuron**, v. 76, n. 1, p. 142-159, 2012.

STOREY, A. E.; ZIEGLER, T. E. Primate paternal care: Interactions between biology and social experience. **Horm Behav**, v. 77, p. 260-71, Jan 2016. ISSN 1095-6867. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/26253726> >.

TABOR, H. K.; RISCH, N. J.; MYERS, R. M. Candidate-gene approaches for studying complex genetic traits: practical considerations. **Nat Rev Genet**, v. 3, n. 5, p. 391-7, 05 2002. ISSN 1471-0056. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/11988764> >.

TAYLOR, M. M.; SAMSON, W. K. The prolactin releasing peptides: RF-amide peptides. **Cell Mol Life Sci**, v. 58, n. 9, p. 1206-15, Aug 2001. ISSN 1420-682X. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/11577979> >.

TOMÁS, A. et al. High amino acid variation in the intracellular domain of the pig prolactin receptor (PRLR) and its relation to ovulation rate and piglet survival traits. **J Anim Sci**, v. 84, n. 8, p. 1991-8, Aug 2006. ISSN 1525-3163. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/16864857> >.

VAN ANDERS, S. M.; GOLDEY, K. L.; KUO, P. X. The Steroid/Peptide Theory of Social Bonds: integrating testosterone and peptide responses for classifying social behavioral contexts. **Psychoneuroendocrinology**, v. 36, n. 9, p. 1265-75, Oct 2011. ISSN 1873-3360. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21724336> >.

VARGAS-PINILLA, P. et al. Evolutionary pattern in the OXT-OXTR system in primates: Coevolution and positive selection footprints. **Proc Natl Acad Sci U S A**, v. 112, n. 1, p. 88-93, Jan 2015. ISSN 1091-6490. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25535371> >.

VIEIRA, C. M. D. A. G. **A evolução molecular da rede gênica da oxitocina em primatas e outros vertebrados**. 2012. Dissertação de Mestrado. Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul

WALLIS, M. Molecular evolution of the neurohypophysial hormone precursors in mammals: Comparative genomics reveals novel mammalian oxytocin and vasopressin analogues. **Gen Comp Endocrinol**, v. 179, n. 2, p. 313-8, Nov 2012. ISSN 1095-6840. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22995712> >.

WALLIS, O. C. et al. Molecular evolution of prolactin in primates. **J Mol Evol**, v. 60, n. 5, p. 606-14, May 2005. ISSN 0022-2844. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/15983870> >.

WEADICK, C. J.; CHANG, B. S. An improved likelihood ratio test for detecting site-specific functional divergence among clades of protein-coding genes. **Mol Biol Evol**, v. 29, n. 5, p. 1297-300, May 2012. ISSN 1537-1719. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/22319160> >.

WILLIAMSON, T. E.; BRUSATTE, S. L.; WILSON, G. P. The origin and early evolution of metatherian mammals: the Cretaceous record. **Zookeys**, n. 465, p. 1-76, 2014. ISSN 1313-2989. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/25589872> >.

YAMASHITA, K.; KITANO, T. Molecular evolution of the oxytocin-oxytocin receptor system in eutherians. **Mol Phylogenet Evol**, v. 67, n. 2, p. 520-8, May 2013. ISSN 1095-9513. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23485918> >.

YANG, Z. PAML 4: phylogenetic analysis by maximum likelihood. **Mol Biol Evol**, v. 24, n. 8, p. 1586-91, Aug 2007. ISSN 0737-4038. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17483113> >.

YANG, Z.; WONG, W. S.; NIELSEN, R. Bayes empirical bayes inference of amino acid sites under positive selection. **Mol Biol Evol**, v. 22, n. 4, p. 1107-18, Apr 2005. ISSN 0737-4038. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/15689528> >.

YOUNG, L. J.; FLANAGAN-CATO, L. M. Editorial comment: oxytocin, vasopressin and social behavior. **Horm Behav**, v. 61, n. 3, p. 227-9, Mar 2012. ISSN 1095-6867. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22443808> >.

ZERBINO, D. R. et al. Ensembl 2018. **Nucleic Acids Res**, Nov 2017. ISSN 1362-4962. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/29155950> >.

ZHU, M.; ZHAO, S. Candidate gene identification approach: progress and challenges. **Int J Biol Sci**, v. 3, n. 7, p. 420-7, Oct 2007. ISSN 1449-2288. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/17998950> >.

Gene Cards. Disponível em: <<http://www.genecards.org/>>. Acesso em: 1 dez. 2013.

Online Mendelian Inheritance in Man. Disponível em: <<http://omim.org/>>. Acesso em: 1 dez. 2013.

Ensemble genome browser. Disponível em: <<http://www.ensembl.org/>>. Acesso em: 1 dez. 2013.

Prolactin Signaling. Disponível em: <<https://www.qiagen.com>>. Acesso em: 1 dez. 2017.

TIMETREE: The Time scale of Life. Disponível em: <<http://www.timetree.org/>>. Acesso em: 1 dez. 2013.

UCSC Genome Browser database . Disponível em: <<http://genome.ucsc.edu/>>. Acesso em: 1 dez. 2013.

ORIGIN AND EVOLUTION OF MARSUPIALS: The University of Edinburgh. Disponível em: <<https://www.ed.ac.uk/>>. Acesso em: 1 fev 2017.

Charles Darwin – Darwin Online. Disponível em <<http://darwin-online.org.uk>>. Acesso em: 1 fev de 2017.

Neill, J.T. (2004). Nature vs nurture in intelligence. Disponível em <wilderdom.com>. Acesso em: 2 fev de 2017.

ANEXO I

A versão completa do manuscrito pode ser acessada em: <https://www.ncbi.nlm.nih.gov/pubmed/25535371>.

Evolutionary pattern in the OXT-OXTR system in primates: Coevolution and positive selection footprints

Pedro Vargas-Pinilla^{a,1}, Vanessa Rodrigues Paixão-Côrtes^{a,1}, Pamela Paré^a, Luciana Tovo-Rodrigues^a, Carlos Meton de Alencar Gadelha Vieira^a, Agatha Xavier^a, David Comas^b, Alcides Pissinatti^c, Marialva Sinigaglia^a, Maurício Menegatti Rigo^a, Gustavo Fioravanti Vieira^a, Aldo B. Lucion^d, Francisco Mauro Salzano^{a,2}, and Maria Cátira Bortolini^{a,2}

^aDepartamento de Genética, Instituto de Biotecnologia, Universidade Federal do Rio Grande do Sul, 91501-970 Porto Alegre, RS, Brazil; ^bInstitut de Biologia Evolutiva, Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, 08003 Barcelona, Spain; ^cCentro de Primatologia do Rio de Janeiro, 20940-200 Rio de Janeiro, RJ, Brazil; and ^dDepartamento de Fisiologia, Instituto de Ciências Básicas da Saúde, Universidade Federal do Rio Grande do Sul, 90050-170 Porto Alegre, RS, Brazil

Contributed by Francisco Mauro Salzano, November 26, 2014 (sent for review July 11, 2014; reviewed by Guido Barbujani and Rafal Slusarz)

Oxytocin is a nonapeptide involved in a wide range of physiologic and behavioral functions. Until recently, it was believed that an unmodified oxytocin sequence was present in all placental mammals. This study analyzed oxytocin (OXT) in 29 primate species and the oxytocin receptor (OXTR) in 21 of these species. We report here three novel OXT forms in the New World monkeys, as well as a more extensive distribution of a previously described variant (Leu8Pro). In structural terms, these OXTs share the same three low-energy conformations in solution during molecular dynamic simulations, with subtle differences in their side chains. A consistent signal of positive selection was detected in the Cebidae family, and OXT position 8 showed a statistically significant ($P = 0.013$) correlation with litter size. Several OXTR changes were identified, some of them promoting gain or loss of putative phosphorylation sites, with possible consequences for receptor internalization and desensitization. OXTR amino acid sites are under positive selection, and intramolecular and intermolecular coevolutionary processes with OXT were also detected. We suggest that some New World monkey OXT-OXTR forms can be correlated to male parental care through the increase of cross-reactivity with its correlated vasopressin system.

OXT | OXTR | primates | coevolution | behavior

Oxytocin has crucial functions related to physiological processes and social behaviors in primates and other placental mammals. A nonapeptide (Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly) (1), oxytocin (OXT-8Leu) is both a neurotransmitter released by neuronal cells in synapses and a hormone, activating receptors distant from the site of its synthesis through the circulatory system (2). In mammals, OXT acts as a hormone in uterine contraction during parturition and in milk ejection while lactating. It is also a key central nervous system neurotransmitter, regulating/modulating complex social and reproductive behaviors (i.e., pair bonding and parental care) (3–7).

Until recently, it was believed that the OXT amino acid chain was the same in all placental mammals. However, Lee and colleagues (8) reported a T > C change in four New World monkeys (NWms), *Saimiri sciureus*, *Cebus apella*, *Callithrix jacchus*, and *Aotus nancimae*, substituting leucine to proline at position 8 (OXT-8Pro). This form was also found in *Tupaia belangeri*, a tree shrew species of Southeast Asia (8). OXT differs from its paralog vasopressin (AVP; Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly) at positions 3 and 8. Variation at position 8 also identifies nonplacental OXT/AVP-like nonapeptides, such as mesotocin, present in some marsupials (7, 9). These findings dispel the notion of a universal OXT amino acid sequence for placental mammals. They also suggest that residue variability at position 8, in some cases associated with variations at positions 2–5, may be connected with the recognition, binding, and activation of receptors, potentially leading to species-specific functional changes (7, 10).

OXT activity depends on adequate interaction with its unique receptor, OXTR, although it can also bind to the vasopressin receptors (AVPR1a, AVPR1b, and AVPR2) with lower affinity (11–13). Similar to other receptors that use G proteins as transducer signals across the cell membranes, OXTR is composed of seven transmembrane (TM1–TM7), four extracellular (N-terminal tail-ECL3), and four intracellular (ICL1-C-terminal tail) domains. ECL and ICL are important for the interaction with OXT and G proteins, respectively, whereas TMs are connected with both functions (7, 11).

In contrast to what is observed for placental mammal OXT, OXTR presents hundreds of variants in regulatory and coding regions, including at the intraspecific level. In humans, OXTR single-nucleotide polymorphisms have been associated with several social behavioral phenotypes (14).

The presence of OXT-OXTR-related systems throughout the animal kingdom indicates that their typical roles in placental mammals are likely exaptations of ancient functions, such as regulation of fluid balance and egg-laying (15, 16). Studies have attempted to investigate both the interaction of OXT-OXTR-like systems and their coevolution (11, 17). However, our knowledge about this nonapeptide-receptor system, including the extent of its variability in the primate order, remains limited.

Significance

It was previously believed that placental mammals present no variability in oxytocin (OXT). The present study reports novel data on the diversity of OXT and its receptor (OXTR) in primate species, including New World monkeys. Contrary to prior expectations, we found three novel OXT forms and several OXTR nonsynonymous changes not previously described. In the Cebidae family, signals of positive selection were found for an OXT variant at position 8, which is associated with larger litter sizes. We detected positive selection for OXTR forms and report a coevolutionary process between changes in OXT and OXTR.

Author contributions: P.V.-P., V.R.P.-C., C.M.d.A.G.V., F.M.S., and M.C.B. designed research; P.V.-P., V.R.P.-C., C.M.d.A.G.V., and A.X. performed research; A.P. contributed new reagents/analytic tools; P.V.-P., V.R.P.-C., P.P., L.T.-R., M.S., M.M.R., G.F.V., and M.C.B. analyzed data; A.P. provided information about the animals investigated; and P.V.-P., V.R.P.-C., P.P., D.C., A.B.L., F.M.S., and M.C.B. wrote the paper.

Reviewers included: G.B., Università di Ferrara; and R.S., University of Gdansk.

The authors declare no conflict of interest.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession nos. [KM186262](#) to [KM186289](#)).

¹P.V.-P. and V.R.P.-C. contributed equally to this work.

²To whom correspondence may be addressed. Email: francisco.salzano@ufrgs.br or maria.bortolini@ufrgs.br.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1419399112/-DCSupplemental.

ANEXO II

A versão completa do manuscrito pode ser acessada em: <https://www.ncbi.nlm.nih.gov/pubmed/27562397>.

Progesterone Response Element Variation in the *OXTR* Promoter Region and Paternal Care in New World Monkeys

Pedro Vargas-Pinilla¹ · Paul Babb² · Leandro Nunes¹ · Pâmela Paré¹ ·
Gabrielle Rosa¹ · Aline Felkl¹ · Dânae Longo¹ · Francisco M. Salzano¹ ·
Vanessa R. Paixão-Côrtes³ · Gislene Lopes Gonçalves¹ · Maria Cátira Bortolini¹

Received: 30 December 2015 / Accepted: 10 August 2016
© Springer Science+Business Media New York 2016

Abstract Paternal care is a complex social behavior common in primate species with socially monogamous mating systems and twin births. Evolutionary causes and consequences of such behavior are not well understood, nor are their neuroendocrine and genetic bases. However, the neuropeptide oxytocin (OXT) and its receptor (OXTR) are associated with parental care in mammalian lineages. Here we investigated the interspecific variation in the number of progesterone response elements (PREs) in the *OXTR* promoter region of 32 primate species, correlating genetic data with behavior, social systems, and ecological/life-history parameters, while controlling for phylogeny. We verified that PREs are only present in New World monkeys and that PRE number is significantly correlated with the presence of paternal care in this branch. We suggest that PRE number could be an essential part of the genetic repertoire that allowed the emergence of taxon-specific complex social behaviors, such as paternal care in marmosets and tamarins.

Edited by Stephen Maxson.

Electronic supplementary material The online version of this article (doi:10.1007/s10519-016-9806-2) contains supplementary material, which is available to authorized users.

✉ Maria Cátira Bortolini
maria.bortolini@ufrgs.br

¹ Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, Porto Alegre, RS 91501-970, Brazil

² Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine at the University of Pennsylvania, SCTR Room 10-130 Bldg. 421, 3400 Civic Center Blvd., Philadelphia, PA 19104, USA

³ Departamento de Biologia, Instituto de Biologia, Universidade Federal da Bahia, Campus de Ondina, Salvador, BA 40170-290, Brazil

Keywords Behavioral genetics · Callitrichinae · Molecular evolution · Oxytocin system

Introduction

Parental care is a complex social behavior associated with animal breeding systems, life history, and ecology (Houston et al. 2005; Kokko and Jennions 2008; Alonzo 2010). Parental care is fundamental for offspring fitness in many animal taxa, and thus can also be characterized as an effective evolutionary strategy.

In mammals, behavioral differences among *taxa* rely not only on the type of care associated with different stages of offspring development and survival but also on the relative contribution of males and females to paternal and maternal care, respectively. Lactation, for instance, is an exclusive maternal attribute linked to the first infancy, but other activities such as carrying, grooming, sharing food, feeding, defending, and teaching can be provided by the father and/or other group members.

Fathers that regularly contribute to offspring care occur in only ~5 % of mammals but evolved independently at least sixty times within these lineages, including four times within the primate order (Lukas and Clutton-Brock 2013). This kind of male behavior is thus equally interesting from an evolutionary perspective when compared with the much more usual maternal care observed in mammals.

Environmental, physiological, epigenetic, and genetic factors promote the neuroendocrine changes associated with parental care in different taxa (Adkins-Regan 2005; Székely et al. 2010; Lucion and Bortolini 2014; French et al. 2016). For this reason, it is not a trivial task to identify the genetic factors involved in both maternal and paternal care, their specific (e.g., breast-feeding) and

ANEXO III

A versão completa do manuscrito pode ser acessada em: <https://www.ncbi.nlm.nih.gov/pubmed/28784762>.

Functional New World monkey oxytocin forms elicit an altered signaling profile and promotes parental care in rats

Lucas T. Parreiras-e-Silva^{a,1}, Pedro Vargas-Pinilla^{b,1}, Diego A. Duarte^{a,1}, Dânae Longo^{b,c}, Grace Violeta Espinoza Pardo^d, Andrea Dolor Finkler^d, Vanessa Rodrigues Paixão-Côrtes^e, Pâmela Paré^b, Diego L. Rovaris^b, Eduardo B. Oliveira^a, Rafael Andrade Caceres^f, Gislene L. Gonçalves^b, Michel Bouvier^{g,h}, Francisco M. Salzano^{b,2}, Aldo B. Lucion^{d,2}, Claudio M. Costa-Neto^{a,2}, and Maria Cátira Bortolini^{b,2}

^aDepartamento de Bioquímica e Imunologia, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, 14040-900 Ribeirão Preto, São Paulo, Brazil; ^bDepartamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, 91501-970 Porto Alegre, Rio Grande do Sul, Brazil; ^cInstituto Federal Farroupilha, 97555-000 Alegrete, Rio Grande do Sul, Brazil; ^dDepartamento de Fisiologia, Instituto de Ciências Básicas da Saúde, Universidade Federal do Rio Grande do Sul, 90050-170 Porto Alegre, Rio Grande do Sul, Brazil; ^eDepartamento de Biologia Geral, Instituto de Biologia, Universidade Federal da Bahia, 40170-290 Salvador, Bahia, Brazil; ^fUniversidade Federal de Ciências da Saúde de Porto Alegre, 90050-170 Porto Alegre, Rio Grande do Sul, Brazil; ^gDepartment of Biochemistry and Molecular Medicine, Université de Montréal, Montréal, QC H3T 1J4, Canada; and ^hInstitute for Research in Immunology and Cancer, Université de Montréal, Montréal, QC H3T 1J4, Canada

Contributed by Francisco M. Salzano, July 12, 2017 (sent for review December 1, 2016; reviewed by Ali Salahpour and Froylán Vargas-Martínez)

The neurohormone oxytocin is a key player in the modulation of reproductive and social behavioral traits, such as parental care. Recently, a correlation between different forms of oxytocin and behavioral phenotypes has been described in the New World Monkeys (NWMs). Here, we demonstrate that, compared with the Leu⁸OXT found in most placental mammals, the Cebidae Pro⁸OXT and *Saguinus* Val³Pro⁸OXT taxon-specific variants act as equiefficacious agonists for the G_q-dependent pathway but are weaker agonists for the β-arrestin engagement and subsequent endocytosis toward the oxytocin receptor (OXTR). Upon interaction with the AVPR1a, Pro⁸OXT and the common Leu⁸OXT yielded similar signaling profiles, being equally efficacious on G_q and β-arrestin, while Val³Pro⁸OXT showed reduced relative efficacy toward β-arrestin. Intranasal treatment with either of the variants increased maternal behavior and also promoted unusual paternal care in rats, as measured by pup-retrieval tests. We therefore suggest that Val³Pro⁸OXT and Pro⁸OXT are functional variants, which might have been evolutionarily co-opted as an essential part of the adaptive genetic repertoire that allowed the emergence of taxon-specific complex social behaviors, such as intense parental care in the Cebidae and the genus *Saguinus*.

paternal care | biased agonism | primates | oxytocin variants | social behaviors

In placental mammals, the nonapeptides oxytocin (Leu⁸OXT) and arginine vasopressin (AVP) differ by two amino acid residues at positions 3 and 8 (OXT: Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly; AVP: Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly). Both act as hormones and neurotransmitters and play a key role in the modulation of reproductive and social behavioral traits, such as parental care and bonding between sexual partners. It is thought that the oxytocin system plays a greater role in females, while the AVP system primarily influences bonding and parental care in males (1–4).

Until a few years ago, it was thought that the nine amino acid sequence of oxytocin, known since the 1950s (5), was strictly conserved across all placental mammals. Recently, this paradigm was challenged when several taxon-specific forms of oxytocin were identified in New World monkeys (NWMs). In a first instance, Lee et al. (6) reported a Leu-to-Pro substitution at position 8 (Pro⁸OXT) in four species. Later, four more NWM forms of oxytocin, Ala⁸OXT, Thr⁸OXT, Phe²OXT, and Val³Pro⁸OXT were identified by our group (7) and by Ren et al. (8). On the other hand, the AVP amino acid sequence was shown to be conserved in species expressing the novel oxytocin forms (9, 10).

The Pro⁸OXT variant emerged at least ~20 Mya (10) and is expressed by all genera of the Cebidae family. In the genus *Saguinus* (family Cebidae, subfamily Callitrichinae) (11), an additional mutation led to the substitution of Ile to Val at position 3 (Val³Pro⁸OXT). These two variants might form an essential part of the genetic repertoire which allowed certain NWM to evolve behavioral traits that are otherwise rare among mammals, such as social monogamy (8) and intense caregiving behavior by males (7). These behaviors are typical of the Cebidae family, which is also characterized by small body size and, in some subfamilies such as Callitrichinae, twin births. Furthermore, a clear positive selection signal was detected for the Pro⁸OXT variant in the Cebidae (7). Based on these findings, it seems plausible that these recently described taxon-specific oxytocin variants might have relevant functional and evolutionary implications. In this context, it is noteworthy that an independent

Significance

Several forms of the oxytocin neurohormone have been found in New World monkeys (NWMs). Previous research has suggested an association between these forms and behaviors typical of this primate branch, including paternal care and monogamy. Our study provides genetic, pharmacological, behavioral, and in silico evidence supporting this connection. Rats treated intranasally with two NWM oxytocin variants showed an increase in some parental care behaviors. The same two variants were found to elicit different cell-signaling profiles in cell-based assays compared with ancestral oxytocin. Our findings highlight how mutations in the OXT DNA sequence coding for a nonapeptide result in distinct signaling profiles that may be linked to the emergence of novel adaptive traits, in this case, paternal care and monogamy.

Author contributions: L.T.P.-e-S., P.V.-P., D.A.D., D.L., A.B.L., C.M.C.-N., and M.C.B. designed research; L.T.P.-e-S., P.V.-P., D.A.D., D.L., G.V.E.P., A.D.F., P.P., and G.L.G. performed research; E.B.O., R.A.C., M.B., A.B.L., C.M.C.-N., and M.C.B. contributed new reagents/analytic tools; L.T.P.-e-S., P.V.-P., D.A.D., D.L., V.R.P.-C., D.L.R., M.B., A.B.L., C.M.C.-N., and M.C.B. analyzed data; and L.T.P.-e-S., P.V.-P., D.A.D., D.L., R.A.C., M.B., F.M.S., A.B.L., C.M.C.-N., and M.C.B. wrote the paper.

Reviewers: A.S., University of Toronto; and F.V.-M., Cinvestav I.P.N.

The authors declare no conflict of interest.

¹L.T.P.-e-S., P.V.-P., and D.A.D. contributed equally to this work.

²To whom correspondence may be addressed. Email: francisco.salzano@ufrgs.br, alucion@ufrgs.br, claudio@fmrp.usp.br, or maria.bortolini@ufrgs.br.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1711687114/-DCSupplemental.

ANEXO IV

O artigo encontra-se em revisão na revista American Journal of Primatology (Número especial sobre sistema OXT-AVP e comportamento).

AVPR1b variation and the emergence of adaptive phenotypes in platyrrhine monkeys

Journal:	<i>American Journal of Primatology</i>
Manuscript ID	AJP-17-0182.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Fam, Bibiana; Universidade Federal do Rio Grande do Sul Instituto de Biociencias, Departamento de genética e biologia molecular Vargas-Pinilla, Pedro; Universidade Federal do Rio Grande do Sul Instituto de Biociencias, Departamento de Genética e Biologia Molecular Pare, Pamela; Universidade Federal do Rio Grande do Sul Instituto de Biociencias, Departamento de Genética e Biologia Molecular Viscardi, Lucas; Universidade Federal do Rio Grande do Sul Instituto de Biociencias, Departamento de Genética e Biologia Molecular Reales, Guillermo; Universidade Federal do Rio Grande do Sul Instituto de Biociencias, Departamento de Genética e Biologia Molecular Felkl, Aline; Universidade Federal do Rio Grande do Sul Instituto de Biociencias, Departamento de Genética e Biologia Molecular Franco, Alvaro; Universidade Federal do Rio Grande do Sul Instituto de Biociencias, Departamento de Genética e Biologia Molecular Lucion, Aldo; Universidade Federal do Rio Grande do Sul Instituto de Ciencias Basicas da Saude, Departamento de Fisiologia Costa-Neto, Claudio ; Universidade de Sao Paulo Faculdade de Medicina de Ribeirao Preto, Departamento de Bioquimica Pissinatti, Alcides; CPRJ/FEEMA Salzano, Francisco; Universidade Federal do Rio Grande do Sul Instituto de Biociencias, Departamento de genética e biologia molecular Paixão-Côrtes, Vanessa; Universidade Federal da Bahia Bortolini, Maria; Universidade Federal do Rio Grande do Sul Instituto de Biociencias, Departamento de Genética e Biologia Molecular
Keywords:	New World monkeys, molecular evolution, AVPR1b, social behavior, reproductive traits

ANEXO V

A versão completa do manuscrito pode ser acessada em:

http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-47572018005005105.



Oxytocin and arginine vasopressin systems in the domestication process

Bibiana S. O. Fam¹, Pamela Paré¹, Aline B. Felkl¹, Pedro Vargas-Pinilla¹, Vanessa R. Paixão-Côrtes¹, Lucas Henriques Viscardi¹ and Maria Cátira Bortolini¹

¹*Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.*

Abstract

Domestication is of unquestionable importance to the technological revolution that has given rise to modern human societies. In this study, we analyzed the DNA and protein sequences of six genes of the oxytocin and arginine vasopressin systems (*OXT-OXTR*; *AVP-AVPR1a*, *AVPR1b* and *AVPR2*) in 40 placental mammals. These systems play an important role in the control of physiology and behavior. According to our analyses, neutrality does not explain the pattern of molecular evolution found in some of these genes. We observed specific sites under positive selection in *AVPR1b* ($\omega = 1.429$, $p = 0.001$) and *AVPR2* ($\omega = 1.49$, $p = 0.001$), suggesting that they could be involved in behavior and physiological changes, including those related to the domestication process. Furthermore, *AVPR1a*, which plays a role in social behavior, is under relaxed selective constraint in domesticated species. These results provide new insights into the nature of the domestication process and its impact on the OXT-AVP system.

Keywords: Oxytocin and receptors, vasopressin and receptors, animal domestication, molecular evolution, positive selection.

Received: May 4, 2017; Accepted: October 1, 2017.

Introduction

The phenomenon of domestication did not go unnoticed by Charles Darwin. With *The Variation of Animals and Plants under Domestication*, published in January 1868, he devoted a whole book to the mechanisms underlying this intriguing process (Darwin, 1868). It is now well known that the complex process of animal domestication involves systematic selective pressures imposed by humans according to their needs and wishes. In a classic experiment modeling domestication, Russian researchers subjected silver foxes to rigorous artificial selection for tameness and correlated traits over more than 50 years. Although part of the lineages retained their ancestral traits, foxes selected for tamability showed a loss of wild-type behavior within relatively few generations, acquiring several classical morphological attributes of domesticated animals such as white spotting, floppy ears, and curly tails (Trut *et al.*, 2009). The selection of traits, such as tameness and the reduction of innate stress, aggression, fear and anxiety, allowed domesticated animals to coexist and coevolve with humans within their constructed niches. This complex process arguably contributed to the rapid spread of our species across the globe, as well as facilitating the emergence of our civilization (Künzler and Sachser, 1999; Wiener and Wilkinson, 2011; Wright, 2015).

In recent years, a picture of the genetic basis for domestication has started to emerge (Schubert *et al.*, 2014; Park *et al.*, 2015; Marsden *et al.*, 2016). Genes of the immune system, neuronal development and behavior were shown to have been co-opted as part of the domestication process (Albert *et al.*, 2012; Wright, 2015). For instance, gray wolves (*Canis lupus lupus*) and dogs (*Canis lupus familiaris*) are highly differentiated in their expression of serotonin receptor genes, consistent with behavioral changes as part of domestication (Li *et al.*, 2013). Importantly, domestication can lead to both the rapid fixation of alleles associated with phenotypes of interest and a relaxation of selective constraints previously imposed by natural selection, as has been described for dogs, horses (*Equus caballus*) and cows (*Bos taurus*) (Zeder, 2012).

In placental mammals, the paralogous nonapeptides oxytocin (OXT: Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly) and arginine vasopressin (AVP: Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly) play an important role in physiological and behavioral functions, such as water homeostasis, vasoconstriction, lactation, uterine contractions, parental care, control of aggression, anxiety and stress (Gimpl and Fahrenholz, 2001; Cagliani *et al.*, 2009; Young and Flanagan-Cato, 2012; McCall and Singer, 2012). In addition to the respective nonapeptides, the genes *OXT* and *AVP* encode the carrier protein neurophysin (NP) and a signal peptide (SP). While the nonapeptides OXT and AVP are relatively well conserved in placental mammals, neurophysin and the signal peptide are more variable. The patterns of variation found in the portion of the *OXT* and *AVP* genes encoding

Send correspondence to: Maria Cátira Bortolini. Departamento de Genética, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil. E-mail: maria.bortolini@ufrgs.br.