

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

GABRIELLE ROSA SILVA

**ASPECTOS EVOLUTIVOS DO GENE *LNPEP*:
DE MAMMALIA A PRIMATES**

PORTO ALEGRE, MAIO DE 2019.

GABRIELLE ROSA SILVA

**ASPECTOS EVOLUTIVOS DO GENE *LNPEP*:
DE MAMMALIA A PRIMATES**

Dissertação submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do grau de Mestre em Genética e Biologia Molecular.

ORIENTADORA: DR^a MARIA CÁTIRA BORTOLINI

PORTO ALEGRE, MAIO DE 2019.

ASPECTOS EVOLUTIVOS DO GENE *LNPEP*: DE MAMMALIA A PRIMATES

GABRIELLE ROSA SILVA

Esta dissertação foi apresentada em seminário público ocorrido no Anfiteatro do Departamento de Genética da UFRGS, no dia 02 de maio de 2019, e aprovada por todos os membros da banca examinadora, composta pelas professoras doutoras abaixo relacionadas:

DRA. GISLENE LOPES GONÇALVES
(PPGBM/UFRGS)

DRA. MÁRCIA PINHEIRO MARGIS
(PPGBM/UFRGS)

DRA. VANESSA RODRIGUES PAIXÃO-CÔRTEZ
(UFBA)

Porto Alegre, Rio Grande do Sul, Brasil.

Maio de 2019.

*Aos meus amados pais, Odila e Celso,
e aos meus sobrinhos preciosos, Ana Valentina e João Pedro.*

AGRADECIMENTOS

À Universidade Federal do Rio Grande do Sul, que me acolhe desde a graduação, pela perseverança em todas as etapas do desenvolvimento científico – mesmo quando o conhecimento geral peregrina timidamente em meio a trevas.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo auxílio em todo o período deste mestrado e pela função que exerce.

Ao Elmo Cardoso, pela subsistência de virtudes humanas ainda que lhe seja atribuído o papel mais impiedoso e indiferente aos apelos filantrópicos – o gerenciamento do cerne burocrático de um departamento de pós-graduação –, o que é totalmente espantoso.

À professora Eliane Kaltchuk, pela bondade, disponibilidade e pelos valiosos conselhos que agora carrego comigo.

Ao Pedro Vargas-Pinilla, pelas revisões e aprimorada visão científica.

Ao Fernando Gazalle e à Muriel Kubiak, pelas mais sortidas propostas de resolução de problemas. Agradeço imensamente a confiança e o suporte químico.

À Ana Mastella, minha singular amiga, pelo estímulo criativo, amparo incessante e pela compreensão das minhas mais variadas formas de humor.

Aos meus irmãos Daniele e Eduardo e à Ingrid Bandeira, minhas sinceras graças pela acalentadora ajuda nos momentos mais frágeis.

Ao Boni e à Brenda, pela amizade inestimável e companhia a qualquer hora do dia, pela alegria que me proporcionam e pelo amor com que sempre me recebem.

Ao Francisco, por ter me ensinado grandes lições sobre a vida – e mostrado que frequentemente é possível restabelecer o que já foi considerado perdido.

Aos meus pais, pela presença carinhosa nos dias mais difíceis. Pelas caminhadas nas tardes de sol que resgatavam o ânimo. Por cada gesto sutilmente afetuoso e gentil.

À Ana Valentina e ao João Pedro, por serem exatamente quem são.

SUMÁRIO

| | |
|---|----|
| ÍNDICE DE TABELAS | ix |
| ÍNDICE DE FIGURAS | x |
| RESUMO | 11 |
| ABSTRACT | 12 |
| 1. INTRODUÇÃO | 13 |
| 1.1. PEPTIDASES – DOS FUNDAMENTOS À SISTEMÁTICA | 13 |
| 1.2. LEUCIL-CISTINIL AMINOPEPTIDASE – DO GENE À ENZIMA | 17 |
| 1.3. BREVE EPÍLOGO EVOLUTIVO | 23 |
| 1.4. JUSTIFICATIVA DO ESTUDO | 31 |
| 2. OBJETIVOS | 32 |
| 2.1. OBJETIVO GERAL | 32 |
| 2.2. OBJETIVOS ESPECÍFICOS | 32 |
| 3. MÉTODOS | 33 |
| 3.1. CONJUNTO DE DADOS INICIAIS | 33 |
| 3.2. ANÁLISES EVOLUTIVAS | 33 |
| 3.3. EFEITOS DE MUTAÇÕES SOBRE A ESTABILIDADE DA PROTEÍNA | 35 |
| 4. RESULTADOS E DISCUSSÃO | 36 |
| 4.1. CONJUNTO DE DADOS E CONSIDERAÇÕES INICIAIS | 36 |
| 4.2. ANÁLISE DAS TAXAS EVOLUTIVAS | 42 |
| 5.3. FILOGENIA EM MAMMALIA E ANCESTRAIS EM PRIMATES | 46 |
| 6. CONSIDERAÇÕES FINAIS E CONCLUSÃO | 51 |
| 7. REFERÊNCIAS | 52 |

ÍNDICE DE TABELAS

| | |
|--|----|
| TABELA 1. IDENTIFICADORES DOS MECANISMOS CATALÍTICOS SEGUNDO A IUBBMB | 14 |
| TABELA 2. FAMÍLIA M1 DE METALOPEPTIDASES (HGNC ID: 1437)..... | 16 |
| TABELA 3. VARIANTES DO PEPTÍDEO OXT ENTRE PRIMATAS..... | 25 |
| TABELA 4. CLASSIFICAÇÃO EM PLATYRRHINI..... | 30 |
| TABELA 5. TÁXONS, REFERÊNCIAS E CLASSIFICAÇÃO | 37 |
| TABELA 6. RESULTADOS DOS MODELOS M8, M7 E M8A EM <i>NSSITES</i> | 42 |
| TABELA 7. SÍTIOS SOB SELEÇÃO POSITIVA SEGUNDO O MODELO M8..... | 42 |
| TABELA 8. RESULTADOS ESTATÍSTICOS PARA O MODELO CLADE C | 45 |
| TABELA 9. PROPORÇÃO DE SÍTIOS SOB DIFERENTES TIPOS DE SELEÇÃO | 45 |
| TABELA 10. SUBSTITUIÇÕES NÃO SINÔNIMAS ENTRE AS LINHAGENS DE PRIMATAS | 47 |
| TABELA 11. VALORES ATRIBUÍDOS A CADA CATEGORIA DE EFEITO SOBRE A ESTABILIDADE.... | 48 |
| TABELA 12. RESULTADOS DO PROGRAMA FOLDX..... | 49 |

ÍNDICE DE FIGURAS

| | |
|--|----|
| FIGURA 1. INFORMAÇÕES DO UNIPROT..... | 17 |
| FIGURA 2. DOMÍNIOS FUNCIONAIS DE LNPEP..... | 18 |
| FIGURA 3. VISÃO ESQUEMÁTICA DAS REGIÕES DE LNPEP E DO SÍTIO DE CLIVAGEM..... | 18 |
| FIGURA 4. LNPEP SOLÚVEL..... | 18 |
| FIGURA 5. O MOTIVO DE CLIVAGEM PARA A FORMA SOLÚVEL DE LNPEP..... | 19 |
| FIGURA 6. SUPEREXPRESSION DAS PROTEASES ADAMS SOBRE A LIBERAÇÃO DE LNPEP..... | 19 |
| FIGURA 7. ESTRUTURA PRESUMIDA DAS PROTEASES..... | 20 |
| FIGURA 8. DISPOSIÇÃO DOS GENES DA SUBFAMÍLIA OXITOCINASE..... | 21 |
| FIGURA 9. SUBFAMÍLIA OXITOCINASE..... | 21 |
| FIGURA 10. REGIÃO DE INTERAÇÃO CATALÍTICA ENTRE OXT E LNPEP..... | 24 |
| FIGURA 11. A ESTRUTURA DO NONAPEPTÍDEO ENTRE VERTEBRADOS..... | 25 |
| FIGURA 12. CLASSIFICAÇÃO DOS SIMIIFORMES..... | 29 |
| FIGURA 13. TOPOLOGIA CONSENSO..... | 37 |
| FIGURA 14. ALINHAMENTO DOS AMINOÁCIDOS DE LNPEP..... | 38 |
| FIGURA 15. PHE154 NÃO É RESTRITO À FAMÍLIA HOMINIDAE..... | 39 |
| FIGURA 16. VARIANTES NO MOTIVO GAMEN..... | 40 |
| FIGURA 17. GAMEN EM ERAP1 E ERAP2..... | 41 |
| FIGURA 18. OS VALORES DE Ω EM LNPEP..... | 43 |
| FIGURA 19. LOCALIZAÇÃO DOS SÍTIOS SOB SELEÇÃO POSITIVA..... | 43 |
| FIGURA 20. VISÃO ESTRUTURAL DE LNPEP..... | 44 |
| FIGURA 21. FILOGENIA DE LNPEP..... | 46 |

RESUMO

O gene *LNPEP* codifica a enzima oxitocinase, uma proteína de membrana zinco-dependente com a propriedade de inativar os nonapeptídeos oxitocina e vasopressina, além de outros peptídeos cíclicos. Apesar de suas potenciais atribuições fisiológicas, a diversidade desse gene e de seu produto funcional permanece desconhecida na maioria dos vertebrados. Nesta pesquisa, investigamos a evolução molecular do gene *LNPEP* e da oxitocinase na classe Mammalia e entre a ordem Primates. Foram incluídas 45 espécies de mamíferos, dos quais 26 são primatas. Identificamos variantes no motivo de ligação ao substrato GAMEN em *Callithrix jacchus*, *Saguinus midas* (G428V) e em *Saimiri boliviensis* (G428A). A variante G428A ocorre também em *Phascolarctos cinereus*, um marsupial. O motivo de clivagem Phe154↓Ala155, considerado até o momento exclusivo de Hominidae, está presente nas famílias Indriidae e Cheirogaleidae (ordem Primates), em duas espécies em Arctiodarctyla e em uma espécie da ordem Afrosoricida. Verificamos também seleção positiva nas posições 338, 793, 831, 838, 882, 884 e 1023 ($P < 0.001$). A proporção de sítios com seleção adaptativa mostrou-se significativamente maior em Platyrrhini em relação às outras espécies consideradas ($P < 0.001$). A diferença de energia livre de Gibbs ($\Delta\Delta G$) resultante das variantes do motivo GAMEN permite classificar as mutações como altamente estabilizadoras ($P < 0.01$). As mutações no motivo de ligação ao substrato são exclusivas de LNPEP, não sendo observadas nas duas outras proteínas da Subfamília Oxitocinase, ERAP1 e ERAP2. Os resultados obtidos neste trabalho evidenciam a relevância evolutiva de LNPEP e indicam potencial coevolução com seus peptídeos substratos. Futuros estudos podem elucidar a funcionalidade dessas variações sobre características reprodutivas, especialmente na subfamília Callitrichinae (Cebidae, Primates).

Palavras-chave: Aminopeptidase placentária, Família M1, Metalopeptidases, Primatas do Novo Mundo.

ABSTRACT

The *LNPEP* gene encodes the enzyme oxytocinase, a zinc-dependent membrane protein with the property of inactivating the nonapeptides oxytocin and vasopressin, in addition to other cyclic peptides. Despite the potential physiological attributions, the diversity of this gene and its functional product remains unresolved for other vertebrates. In this work, we investigated the molecular evolution of the *LNPEP* gene and the oxytocinase in the Class Mammalia and among the Order Primates. Were included 45 species of mammals, of which 26 are primates. We identified variants in the substrate binding region, GAMEN, in *Callithrix jacchus*, *Saguinus midas* (G428V) and *Saimiri boliviensis* (G428A). The G428A variant also occurs in *Phascolarctos cinereus*, a marsupial. The cleavage motif Phe154↓Ala155, considered exclusive to Hominidae, is also present in the families Indriidae and Cheirogaleidae (order Primates), in two species in Arctiodarctyla and in one taxon of the order Afrosoricida. Additionally, we found positive selection at positions 338, 793, 831, 838, 882, 884 and 1023 ($P < 0.001$). The proportion of sites with adaptive selection was significantly higher in Platyrrhini than in other species ($P < 0.001$). The Gibbs free energy difference ($\Delta\Delta G$) resulting from the variants of the GAMEN motif allows to classify the mutations as highly stabilizing ($P < 0.01$). Finally, mutations in the substrate binding motif described here are exclusive of LNPEP and are not observed in the two other proteins of the Oxytocinase Subfamily, ERAP1 and ERAP2. The results obtained in this work show the evolutionary relevance of LNPEP and indicate potential coevolution with its substrate peptides. Future studies may elucidate the functionality of these variations on reproductive traits, particularly in the subfamily Callitrichinae (Cebidae, Primates).

Keywords: Placental aminopeptidase, Family M1, Metallopeptidases, New World Monkeys.

1. INTRODUÇÃO

O propósito desta Introdução é apresentar a enzima oxitocinase, codificada pelo gene *LNPEP*, mas não ser restrita simplesmente a uma descrição de suas características. Optou-se, em vez disso, por contextualizá-la fisiológica e evolutivamente a fim de que sua importância possa, ao menos em parte, ser interpretada com maior precisão, o que teria potencial para disseminar um justo interesse pelo assunto. Da mesma forma, busca-se esclarecer, ao longo da leitura, a escolha de *LNPEP* como tema desta pesquisa. Para tanto, e empregando o maior didatismo facultado a esta autora, esta seção parte de uma exposição das propriedades gerais de enzimas peptidases, incluindo os critérios sob os quais são classificadas, para então dedicar maior detalhamento à própria oxitocinase. A isso, segue-se uma revisão de características gerais e reprodutivas da classe Mammalia. Como último tópico introdutório, há um panorama da variabilidade dos substratos mais conhecidos da enzima, os nonapeptídeos OXT e AVP. Esses peptídeos foram objetos de intensos estudos nos últimos dez anos, período em que variações funcionalmente relevantes em suas estruturas foram descobertas na ordem Primates, especificamente na parvordem Platyrrhini.

1.1. PEPTIDASES – DOS FUNDAMENTOS À SISTEMÁTICA

Os processos bioquímicos que transcorrem nas células são, em geral, sucessivas etapas de reações mediadas por catálise enzimática. Embora exceções compostas por RNA existam, as chamadas ribozimas, a ampla maioria dos catalisadores biológicos é representada por enzimas, cuja natureza é essencialmente proteica. As enzimas aumentam exponencialmente as taxas de reação ao reduzir a energia livre entre substrato e produto (Watson, 2014). Essa propriedade é decorrente da configuração do sítio ativo, região onde a catálise ocorre. O sítio ativo é geralmente uma cavidade ou sulco na superfície da enzima com propriedades complementares em forma e interação ao estado de transição da reação. Os contatos favoráveis que se formam quando os reagentes se associam a esse sítio compensam, em parte, a distorção a que são submetidos (Voet, Voet and Pratt, 2016). A especificidade enzimática está relacionada à precisão com que o sítio ativo está configurado em relação ao substrato.

Até meados de 2018, eram reconhecidas seis classes de enzimas. A sétima e nova classe, composta por translocases, foi incluída em agosto desse ano (Tipton, 2018). Segundo as regras do Comitê de Nomenclatura da União Internacional de Bioquímica e Biologia Molecular (NC-IUBBMB), responsável por essa sistemática, cada enzima é descrita por uma sequência de números precedidas pelas iniciais “EC” (*Enzyme Commission*). O primeiro número, que indica o tipo de reação catalítica, corresponde a uma das sete categorias propostas, resumidas na Tabela 1. As classes são estruturadas, assim, conforme o tipo de reação que seus membros catalisam.

Tabela 1. Identificadores dos mecanismos catalíticos segundo a IUBBMB

| Identificador | Mecanismo Catalítico |
|---------------------------|---|
| EC1: Oxirredutases | Enzimas que catalisam reações de oxidação-redução. O substrato que sofre oxidação é o doador de hidrogênio. |
| EC 2: Transferases | Enzimas que transferem um grupamento funcional de um composto (o doador) para outro (aceptor). |
| EC 3: Hidrolases | Agrupa enzimas que catalisam a hidrólise de diversos tipos de ligações covalentes. Inclui as peptidases, que hidrolisam ligações peptídicas. |
| EC 4: Liases | Catalisam a clivagem de ligações C-C, C-O, C-N, entre outras, por hidrólise ou oxidação. Possuem dois substratos envolvidos em uma direção e apenas um na direção oposta. |
| EC 5: Isomerases | Catalisam alterações geométricas ou estruturais dentro de uma molécula. |
| EC 6: Ligases | Catalisam a síntese de uma molécula a partir da ligação entre duas outras moléculas, com a simultânea hidrólise de ATP ou de outro composto trifosfatado. |
| EC 7: Translocases | Catalisam o movimento de íons ou moléculas entre as faces (lados 1 e 2) das membranas ou dentro delas. |

Tabela 1. Categorias de enzimas. Conforme o mecanismo catalítico, são reconhecidas sete classes pela NC-IUBBMB. As categorias foram atualizadas em agosto de 2018 com a inclusão de translocases (EC 7).

Pertencentes às hidrolases (EC 3), as enzimas que operam o catabolismo de proteínas pela hidrólise de ligações peptídicas são identificadas como peptidases (EC 3.4). Várias linhagens em Vertebrata apontam evolução convergente em EC 3.4 (Rawlings and Salvesen, 2013). Desse modo, diferentes peptidases podem realizar a mesma reação a partir de mecanismos catalíticos distintos. O banco de dados MEROPS (Rawlings *et al.*, 2018) mantém uma classificação evolutiva das peptidases, onde estas são avaliadas e agrupadas de acordo com similaridades entre suas sequências de aminoácidos. Cada conjunto formado por peptidases homólogas compõe uma família, identificada por uma letra correspondente ao tipo catalítico das enzimas nela contidas e por um número único. As famílias consideradas homólogas – aquelas com relações evolutivas evidenciadas por suas estruturas terciárias, pela ordem dos resíduos de seus sítios catalíticos na cadeia polipeptídica e/ou pelos motivos comuns ao redor desses resíduos – são, por fim, reunidas em um clã (o equivalente a uma superfamília), nomeado com duas letras referentes ao modo catalítico das famílias incluídas. Há, até o momento, mais de cinquenta clãs relacionados no banco de dados.¹

O clã MA, por exemplo, refere-se às metalopeptidases, ou seja, peptidases cujo mecanismo catalítico envolva um metal. Dito de maneira mais específica, as metalopeptidases são particularizadas pelo modo que operam o ataque nucleofílico: nelas, um cátion metal divalente – em geral, zinco; mas possivelmente, embora com menor frequência, cobalto, manganês, níquel ou cobre – é o responsável pela ativação da molécula de água. O íon metal é mantido no lugar por três aminoácidos ligantes. Todas possuem o motivo **HEXXH**, no qual os dois resíduos de Histidina, His (H), são zinco-ligantes; ao passo que o resíduo Glutamato, Glu (E), tem função catalítica. Esse motivo não é exclusivo de metalopeptidases, mas em MA há pormenores identificáveis: o motivo HEXXH pode ser descrito com maior exatidão como Xaa-Xbb-Xcc-**His-Glu**-Xbb-Xbb-**His**-Xbb-Xdd, sendo Xaa hidrofóbico ou uma Treonina (Thr); Xbb, neutro; Xcc é qualquer aminoácido, exceto Prolina (Pro); e Xdd é hidrofóbico. Essa disposição também é chamada de consenso de Jongeneel (Jongeneel, Bouvier and Bairoch, 1989), pesquisador que legou, triunfante em sua luminosa simplicidade, este notavelmente prático e mnemônico padrão ao mundo científico.

¹ Os clãs e as respectivas famílias podem ser visualizados em https://www.ebi.ac.uk/merops/cgi-bin/statistics_index?type=P.

Dentre o conjunto MA, a família M1 é composta por exopeptidases que agem sobre o N-terminal de polipeptídeos; seus membros, dispostos na Tabela 2, são, em geral, identificados como aminopeptidases (EC 3.4.11), enzimas constitutivas de funções celulares essenciais (Hernández *et al.*, 2015). Além de aminopeptidases citoplasmáticas, M1 inclui enzimas de membrana. Nas peptidases ligadas à membrana, a região N-terminal citoplasmática é seguida por um domínio transmembrana, e o restante da molécula é extracelular e altamente glicosado. Assume-se que a família M1 estava presente no ancestral de todos os eucariotos. De fato, seus membros são amplamente distribuídos, presentes em bactérias Gram-negativas e Gram-positivas, Archaea, Protozoa, Fungi e Animalia (Rawlings and Salvesen, 2013).

Tabela 2. Família M1 de Metalopeptidases (HGNC ID: 1437)

| Símbolo | Nome | Sinônimos | Cromossomo | RefSeq |
|----------------|--|---|-------------------|------------------|
| ANPEP | alanyl aminopeptidase | LAP1, gp150, p150 | 15q26.1 | -- |
| C9orf3 | Chr. 9 open reading frame 3 | C9ORF3, FLJ14675, APO, AOPEP, AP-O | 9q22.32 | NM_032823 |
| ENPEP | glutamyl aminopeptidase | gp160, CD249 | 4q25 | -- |
| ERAP1 | endoplasmic reticulum aminopeptidase 1 | A-LAP, ARTS-1, PILS-AP, KIAA0525, ERAP1 | 5q15 | NM_016442 |
| ERAP2 | endoplasmic reticulum aminopeptidase 2 | L-RAP, LRAP | 5q15 | NM_022350 |
| LNPEP | Leucyl/cystinyl aminopeptidase | CAP, PLAP, P-LAP, IRAP | 5q15 | NM_005575 |
| LTA4H | leukotriene A4 hydrolase | -- | 12q23.1 | NM_000895 |
| LVRN | laeverin | APQ, FLJ90650, AQPEP, TAQPEP | 5q23.1 | NM_173800 |
| NPEPPS | aminopeptidase puromycin sensitive | PSA, MP100 | 17q21.32 | NM_006310 |
| RNPEP | arginyl aminopeptidase | APB | 1q32.1 | NM_020216 |
| RNPEPL1 | arginyl aminopeptidase like 1 | -- | 2q37.3 | NM_018226 |
| TRHDE | thyrotropin releasing hormone degrading enzyme | PGPEP2, TRHDE, PAP-II | 12q21.1 | NM_013381 |
| TAF2 | TATA-box binding protein associated factor 2 | TAFII150, CIF150 | 8q24.12 | NM_003184 |

Tabela 2. Família M1 de metalopeptidases. M1 é formada por metalopeptidases zinco-dependentes nas quais dois dos ligantes são histidinas, formando o motivo 'HEXXH'. Em destaque, LNPEP.

1.2. LEUCIL-CISTINIL AMINOPEPTIDASE – DO GENE À ENZIMA

O gene *LNPEP* codifica um produto com termos muitas vezes intercambiáveis na literatura científica: Leucil-Cistinil Aminopeptidase (LNPEP), Aminopeptidase Placentária (P-LAP) e Cistinil Aminopeptidase (CAP)² são exemplos. Em essência, todas essas denominações referem-se à mesma enzima – o produto proteico do gene *LNPEP*, purificado primeiramente como uma proteína solúvel (Tsujimoto *et al.*, 1992) e homóloga à IRAP (*insulin-regulated aminopeptidase*) murina (Keller *et al.*, 1995). Em favor da precisão conceitual, a enzima será referida, neste trabalho, como **oxitocinase**; nas vezes em que for abreviada será denominada **LNPEP**. O termo “oxitocinase” faz menção à propriedade de inativação de oxitocina, resultado da clivagem da ligação peptídica entre a cisteína N-terminal e o resíduo adjacente de tirosina presente no peptídeo. Além de oxitocina (OXT), entretanto, arginina-vasopressina (AVP), somatostatina, angiotensina III, Dinorfina A, Lys-Bradicinina e Met-Encefalina (Met-Enk) são substratos passíveis de catálise (Rogi *et al.*, 1996).

O transcrito originado do gene *LNPEP* possui 3078 nucleotídeos. Após a tradução, a proteína contém 1025 aminoácidos. Por modificação pós-traducional, é possível obter a forma sérica, gestacional, composta por 871 aminoácidos (FIGURA 1). A FIGURA 2 mostra os sítios que limitam os domínios da proteína. A enzima LNPEP possui 110 resíduos de aminoácidos em sua porção N-terminal citoplasmática, região seguida por uma fração transmembrana com 21 resíduos. O segmento C-terminal extracelular é composto por 894 resíduos de aminoácidos (UniProt Consortium, 2018). Estas informações estão apresentadas na FIGURA 3. A forma gestacional, por fim, está representada na FIGURA 4.

| Position(s) | Description | Graphical view | Length |
|-------------|--|--|--------|
| 1 – 1025 | Leucyl-cystinyl aminopeptidase |  | 1025 |
| 155 – 1025 | Leucyl-cystinyl aminopeptidase, pregnancy serum form |  | 871 |

Figura 1. Informações do UniProt. A forma integral, com 1025 aminoácidos, pode sofrer modificação pós-traducional e originar uma proteína solúvel com 871 aminoácidos.

² Nome adotado por MEROPS: *cystinyl aminopeptidase*.

Abaixo, algumas ilustrações baseadas na estrutura de LNPEP.

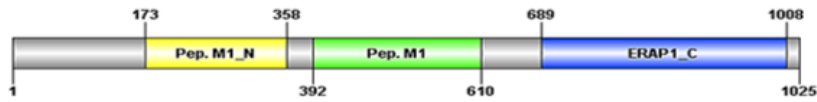


Figura 2. Domínios funcionais de LNPEP. As posições são baseadas na referência Q9UIQ6 (UniProt). Imagem construída com o programa DOG 2.0.

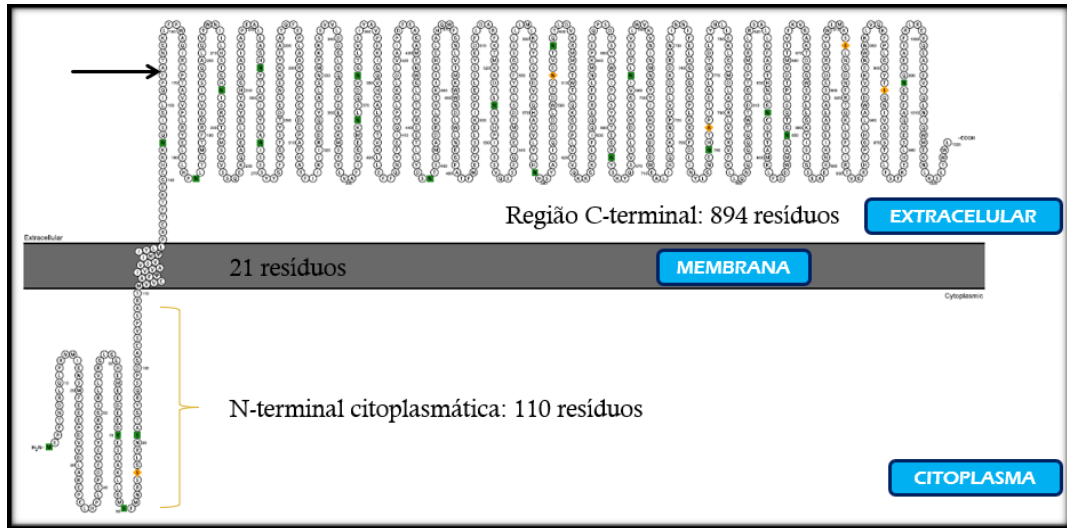


Figura 3. Visão esquemática das regiões de LNPEP e do sítio de clivagem. A seta indica a localização dos sítios 154 e 155. Imagem construída com o programa Protter.

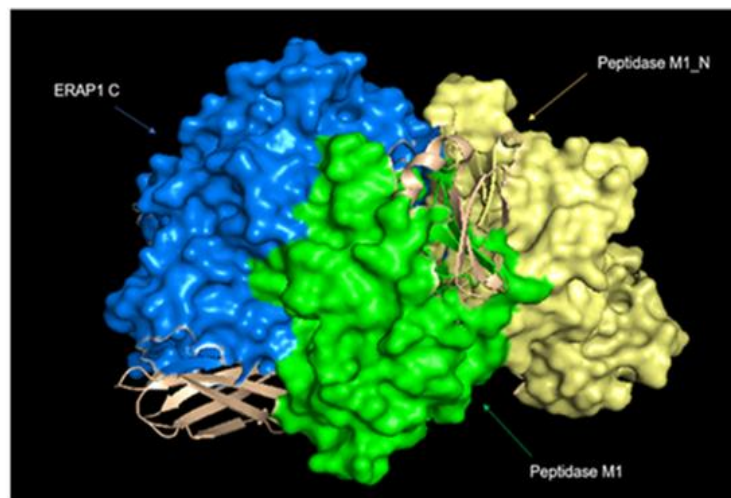


Figura 4. LNPEP solúvel. A enzima contém 871 aminoácidos após modificação pós-tradução. Imagem construída no programa Pymol.

A liberação de LNPEP no plasma seria, segundo Pham et al. (2009), uma característica exclusiva da família Hominidae. Esta conclusão é baseada na observação de que o sítio de clivagem Phe154↓Ala155 não estava conservado nas outras espécies por eles analisadas, conforme mostra a FIGURA 5.

| | | | |
|---------------|-----|--------------------------------|-----|
| HUMAN | 141 | CHKKNQSIGLIQPFATNGKLFPPWAQIRLP | 169 |
| CHIMPANZEE | 141 | CHKKNQSIGLIQPFATNGKLFPPWAQIRLP | 169 |
| RHESUS MONKEY | 292 | CHKKNQSIGLIQPIATNGKLFPPWAQIRLP | 320 |
| HORSE | 141 | CHKKNQSMVLIQPLATNGKLFPPWAQIRLP | 169 |
| PIG | 141 | CHKKNQSLGLIQPLATNGKLFPPWAQVRLP | 169 |
| OPOSSUM | 190 | CHKKNQSMELIQPIATNGKLFPPWAQIRLP | 218 |
| DOG | 141 | CHKRNQSMGVIQPLATNGKLFPPWAQVRLP | 169 |
| RAT | 141 | CHKTNQSAELIQPIATNGKVFPPWAQIRLP | 169 |
| MOUSE | 141 | CHKTNQSAELIQPVATNGKVFPPWAQIRLP | 169 |
| COW | 163 | CHKNNRPTPIQPIATNGKLFPPWAQIRLP | 191 |
| PLATYPUS | 255 | CHKTNQSMEWTLPIATNGKVFPPWAQLRPL | 253 |

Figura 5. O motivo de clivagem para a forma solúvel de LNPEP. Em negrito, o resíduo considerado exclusivo de Hominidae por Pham et al. (2009).

Atribui-se às metaloproteases ADAM9 e ADAM12 a atividade de reconhecimento da sequência de aminoácidos no sítio de clivagem de LNPEP e, por consequência, o papel de converter a peptidase à forma solúvel a partir da conformação da proteína ligada à membrana (Ito *et al.*, 2004). Essa relação funcional é resultante da verificação de que o aumento da forma solúvel de LNPEP é precedido por superexpressão dos genes *ADAM9* e *ADAM12*, embora apenas *ADAM12* seja expresso na placenta (FIGURA 6).

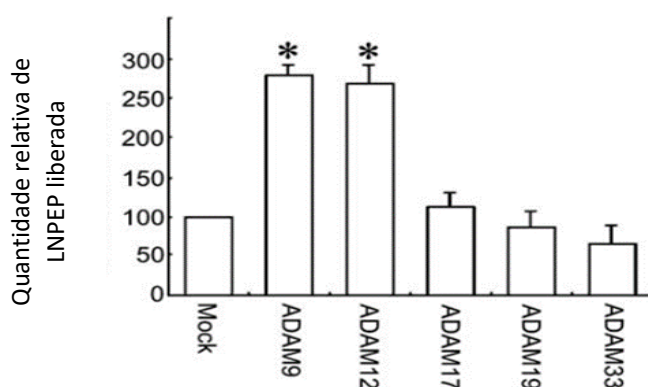


Figura 6. Superexpressão das proteases ADAMs sobre a liberação de LNPEP. ADAM9 e ADAM12 promovem maior liberação do domínio solúvel extracelular de LNPEP a partir de sua estrutura integral ligada à membrana. Análise densitométrica. Valores mostrados como média \pm desvio padrão. Adaptado de Ito et al (2004).

Uma vez que proteases da família ADAM agem na superfície celular e liberam proteínas integrais da membrana, possivelmente contribuiriam também para a liberação preferencial de LNPEP placentária. A FIGURA 7, abaixo, ilustra a estrutura genérica de uma protease como, por exemplo, ADAM, bem como o local onde essas enzimas atuam para liberar a forma solúvel de uma proteína de membrana.

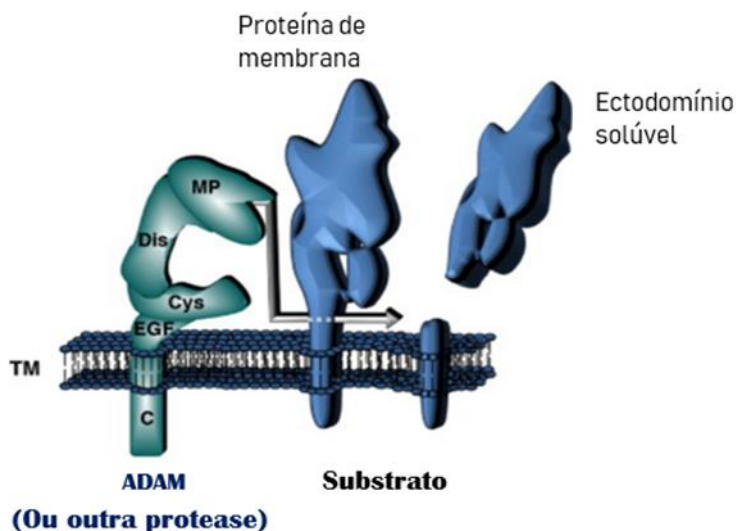


Figura 7. Estrutura presumida das proteases. ADAM possui uma região extracelular onde está localizado o domínio metaloprotease (MP). Ainda estão presentes os domínios desintegrina (Dis), EGF e o domínio rico em cisteína (Cys). O domínio citosólico é indicado por (C). ADAM promove a liberação das partes solúveis extracelulares de proteínas de membrana (ectodomínios solúveis). TM: região transmembrana. Adaptado de Reiss and Saftig (2009).

No gene *LNPEP*, os motivos de ligação ao zinco e de exopeptidase são codificados pelos éxons 6 e 7, e a conservação dentro deles é fundamental para a atividade enzimática. A funcionalidade dos resíduos do sítio ativo foi estimada em um estudo de análise mutacional sítio-dirigida, onde foi demonstrado que mutações nos resíduos G428, A429 e N432 alteraram a atividade catalítica. As mutações G428D, G428E e G428Q aboliram a catálise de Leu-Enk; as mutações A429G e N432A resultaram em uma enzima incapaz de catalisar Leu-Enk e AVP. Os resultados, em suma, mostraram que Met, Glu e Asn são essenciais para a atividade catalítica, enquanto Gly e Ala são fundamentais para a ligação ao substrato e também a inibidores (Ye *et al.*, 2007). Em mais um estudo baseado em mutagênese de *LNPEP*, substituições de Phe544 por alanina, isoleucina e valina resultaram em alteração na velocidade de hidrólise ou mesmo em perda completa da atividade catalítica (Albiston *et al.*, 2010).

Análises estruturais e buscas por homologia basearam o estudo de Tsujimoto e Hattori (2005), no qual propuseram que as aminopeptidases derivadas de adipócitos (ERAP1) e as aminopeptidases derivadas de leucócitos (ERAP2) fossem reunidas à LNPEP em um grupo distinto dentro a família M1, nomeado Subfamília Oxitocinase (FIGURA 8). A identidade de LNPEP em relação à ERAP1 e ERAP2 é de 43% e 49%, respectivamente. Os genes humanos que codificam essas enzimas estão localizados de modo contíguo na região cromossômica 5q15, o que poderia indicar uma diversificação a partir de um único gene ancestral (FIGURA 9). Os motivos de ligação ao zinco, de estrutura HEXXH(X)₁₈E, e o motivo de exopeptidase GAMEN, estão presentes nas três enzimas e são, segundo os autores, idênticos entre elas.

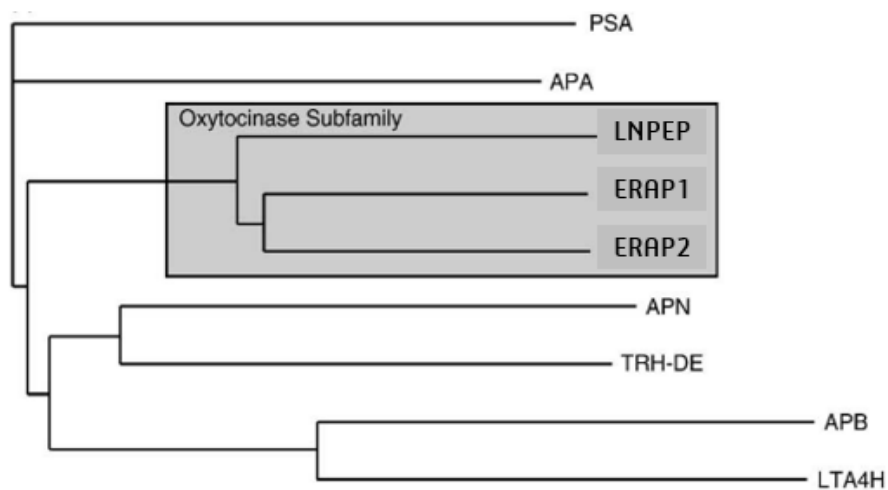


Figura 8. Subfamília Oxitocinase. As relações filogenéticas dos integrantes das peptidases M1, segundo Tsujimoto e Hattori. Em destaque, a Subfamília Oxitocinase. Adaptado do artigo original.

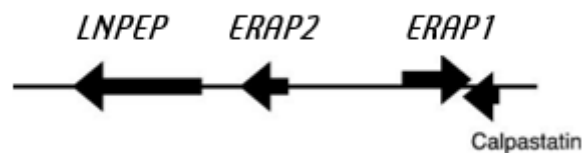


Figura 9. Disposição dos genes da Subfamília Oxitocinase. Os genes *LNPEP*, *ERAP1* e *ERAP2* são adjacentes na região cromossômica 5q15. Adaptado do artigo original de Tsujimoto e Hattori (2005).

A atividade da oxitocinase é acentuada no período gestacional; embora permaneça baixa nos três primeiros meses, a atividade sérica aumenta progressivamente durante o segundo e terceiro trimestres, atingindo um valor máximo com a proximidade da parturição, quando então declina rapidamente (Nomura *et al.*, 2005). A enzima estaria presente na placenta e no soro materno como um mecanismo adjuvante de degradação e sustentação de níveis apropriados de oxitocina e vasopressina nos tecidos fetais, placentários e maternos. Os peptídeos OXT e AVP são agentes uterotônico e vasoativo, respectivamente, e são produzidos pelo feto no decorrer da gestação ou durante acidemia. Assim, os níveis crescentes de oxitocinase na circulação materna previnem contrações uterinas prematuras e controlam a pressão sanguínea materno-fetal (Tian, Huang and Wen, 2016). A propriedade de clivar ligações de cisteína em substratos sintéticos e a capacidade de catalisar hormônios peptídeos são, por isso, os atributos mais evidentes e mais bem estudados de LNPEP.

Enquanto a expressão de *LNPEP* tem uma função crítica na manutenção da dinâmica de sinalização apropriada durante a gestação, a proteína LNPEP e seus mRNAs são também expressos em uma variedade de tecidos e órgãos, incluindo cérebro, rins e coração. Essa ampla distribuição da expressão de *LNPEP* atribui à proteína codificada funções além daquelas envolvidas na gestação, e estudos nos últimos anos forneceram evidências para os papéis de LNPEP na lactação, cognição, função diurética renal e resposta a estressores (Nakada *et al.*, 2011; Elkins *et al.*, 2017). Em humanos, a oxitocinase é detectável em diversas regiões cerebrais, incluindo o córtex (Matsumoto *et al.*, 2001), sendo responsável pela degradação de OXT quando o peptídeo é liberado no sistema nervoso central. Inibidores da atividade catalítica de IRAP intensificaram a retenção da memória espacial e acentuaram a memória aversiva em ratos (Albiston *et al.*, 2003). Com esse mesmo modelo animal, estudos demonstraram que IRAP está presente na amígdala, uma parte central do sistema límbico, e em áreas cerebrais relacionadas mais diretamente à cognição, incluindo hipocampo e córtex (Fernando *et al.*, 2005).

1.3. BREVE EPÍLOGO EVOLUTIVO

De Oxitocinase ao nonapeptídeo OXT

A enzima LNPEP libera um aminoácido N-terminal, Cys↓Xaa, no qual a metade da cistina é envolvida em uma ponte dissulfeto (FIGURA 10). Sabe-se que oxitocina, vasopressina, somatostatina, angiotensina III e Met-enkefalina são substratos da aminopeptidase e, assim, sujeitos à clivagem; por outro lado, calcitonina e endotelinas não o são. Os peptídeos OXT, AVP e Met-Enk estão envolvidos em uma grande variedade de funções fisiológicas e comportamentais. A modulação da memória é um dos aspectos já estudados (Yeatman *et al.*, 2008). Por exemplo, o desenvolvimento de inibidores de LNPEP foi apontado como uma alternativa terapêutica a doenças associadas à perda de memória. O mecanismo que sustentaria a aplicação de LNPEP nessa dinâmica envolve os agentes neuroendócrinos já citados. No caso da oxitocina, para tomar um modelo apenas, há uma meia-vida de aproximadamente 20 minutos no fluido cérebro-espinhal (Ludwig and Leng, 2006) e de três minutos no plasma (Gazis, 1978). A inibição da oxitocinase ampliaria a meia-vida desse hormônio, bem como a de outros neuropeptídeos que influenciam o aprendizado e a memória.

O gene *OXT* possui três éxons e codifica uma pré-proteína formada por 125 aminoácidos, a chamada pró-proteína oxitocina-neurofisina I. Da clivagem desse produto, apenas nove resíduos de aminoácidos originam o peptídeo denominado oxitocina, enquanto todo o restante constitui a neurofisina I, uma proteína carreadora (Gimpl and Fahrenholz, 2010). OXT é liberado na circulação a partir da neurohipófise, onde é armazenado após ser produzido no hipotálamo: é, portanto, produzido no cérebro, mas sua atuação é sistêmica. Por conta disso, é reconhecido tanto como hormônio quanto por neurotransmissor (Wallis, 2012; Haas, Anderson and Smith, 2013; Olazábal, 2014). A oxitocina tem importância destacada em vários aspectos que particularizam a classe Mammalia; entre eles, a lactação e a viviparidade, estando também associada a comportamentos sociais complexos (Pond, 1977; Lee *et al.*, 2010; Yamashita and Kitano, 2013).

A figura a seguir exhibe simplificadamente a ligação de OXT que é clivada pela enzima LNPEP.

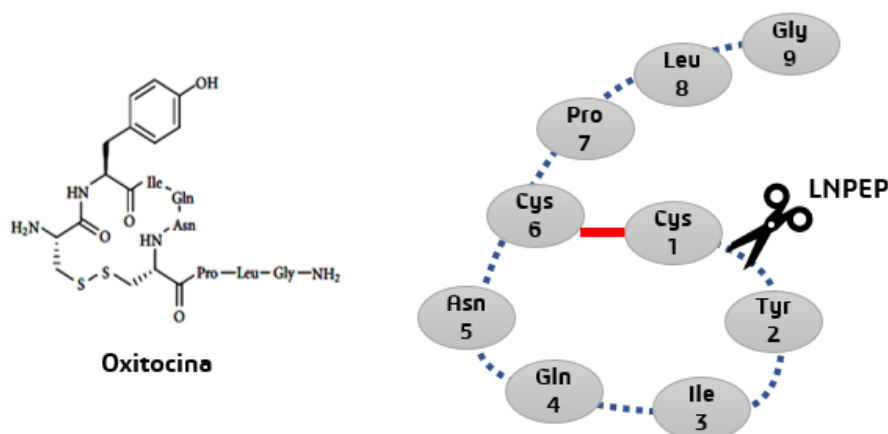


Figura 10. Região de interação catalítica entre OXT e LNPEP. Nessa ilustração, baseada na literatura, o peptídeo oxitocina é mostrado à direita. Já à esquerda, é assinalada a região clivada pela enzima LNPEP. A linha vermelha indica a ligação dissulfeto.

A sequência consenso de nove aminoácidos da oxitocina (Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly) era considerada idêntica entre os eutérios até 2011. Esta noção, entretanto, foi desfeita com a descoberta de uma substituição na posição 8 (Pro⁸OXT) (Lee *et al.*, 2011) em quatro espécies de Primatas do Novo Mundo – *Saimiri sciureus*, *Cebus apella*, *Callithrix jacchus* e *Aotus nancimae*. Em 2015, novas variantes foram identificadas (Vargas-Pinilla *et al.*, 2015), além da observação de que a presença de Pro⁸OXT na família Cebidae é generalizada, o que indica que a distribuição desta variante é mais abrangente do que aquela previamente observada por Lee *et al.* (2011). Posteriormente essa variabilidade foi ampliada com o registro de Phe²OXT em *Alouatta caraya* (família Atelidae). A Tabela 3 resume as variantes descobertas entre primatas e a FIGURA 11 apresenta as estruturas do nonapeptídeo encontradas em diversas espécies.

Tabela 3. Variantes do peptídeo OXT entre primatas

| Variantes | Posições no nonapeptídeo OXT | | | | | | | | | Espécies | Ref. |
|--|------------------------------|------------|------------|-----|-----|-----|-----|------------|-----|----------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | | |
| Consenso | Cys | Tyr | Ile | Gln | Asn | Cys | Pro | Leu | Gly | <i>a</i> | i |
| Pro⁸OXT | Cys | Tyr | Ile | Gln | Asn | Cys | Pro | Pro | Gly | <i>b</i> | ii |
| Val³Pro⁸OXT | Cys | Tyr | Val | Gln | Asn | Cys | Pro | Pro | Gly | <i>c</i> | |
| Ala⁸OXT | Cys | Tyr | Ile | Gln | Asn | Cys | Pro | Ala | Gly | <i>d</i> | iii |
| Thr⁸OXT | Cys | Tyr | Ile | Gln | Asn | Cys | Pro | Thr | Gly | <i>e</i> | |
| Phe²OXT | Cys | Phe | Ile | Gln | Asn | Cys | Pro | Leu | Gly | <i>f</i> | iv |

Tabela 3. Panorama das variantes de OXT. Espécies: a) *Homo sapiens*; b) *Saimiri sciureus*, *Cebus apella*, *Callithrix jacchus* e *Aotus nancimae*; c) *Saguinus bicolor*, *Saguinus martinsi* e *Saguinus niger*; d) *Cacajao melanocephalus*, *Chiropotes utahikae*; e) *Chiropotes albinasus*; f) *Alouatta caraya*. Referências: i) Tuppy et al., 1953) ii) Lee et al., 2011; iii) Vargas-Pinilla et al., 2015; iv) Ren et al., 2015.

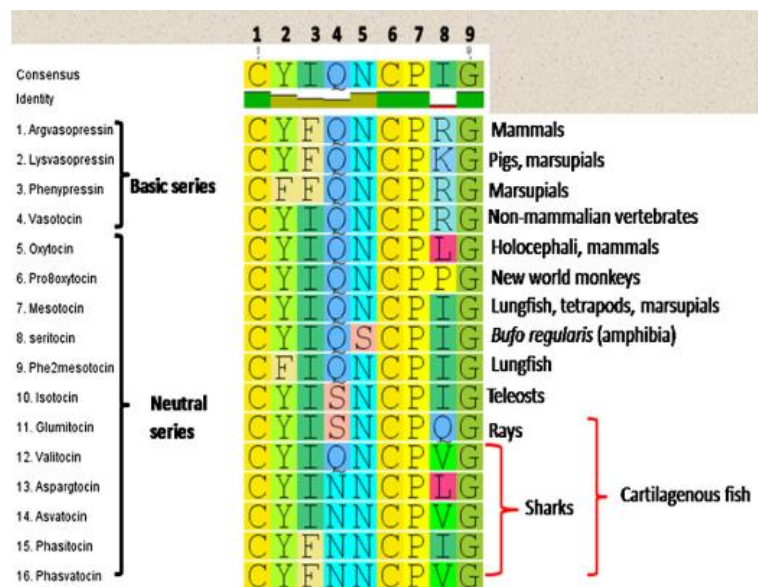


Figura 11. A estrutura do nonapeptídeo entre vertebrados. Imagem apresentada em Banerjee et al. (2017).

De Mammalia a Primates

Antes de restringirmos nossa visão aos entes mamalianos, é interessante assinalar que uma pesquisa relativamente recente abordou a evolução de LNPEP em uma espécie de peixe, *Cyprinodon salinus* (Elkins *et al.*, 2017). O tema – praticamente inédito – resultou em alguns achados relevantes. Os transcritos codificantes da enzima LNPEP, concluíram os autores, foram identificados com relativa abundância nas gônadas, hipotálamo, prosencéfalo, colículo superior, brânquias e músculo esquelético de *C. salinus* adultos. Uma avaliação posterior revelou que os mRNAs de *LNPEP* estavam em concentrações elevadas no hipotálamo de animais subordinados socialmente. Em relação aos sexos desses espécimes com fenótipo comportamental caracterizado por dominância, os transcritos permaneceram em menor quantidade no cérebro dos machos. Essas informações sugerem uma associação entre comportamento e transcrição de *LNPEP* no hipotálamo (Ebstein *et al.*, 2012; Elkins *et al.*, 2017). De modo inteiramente fortuito, ainda, durante as buscas por demais espécies em bancos de dados, os autores notaram que havia uma alteração do motivo GAMEN em *Python bivittatus*: nessa espécie, há uma ligeira modificação na ordem dos resíduos – AGMEN. Essa disposição dos resíduos do motivo de ligação parece ser a norma na subordem Serpentes (ordem Squamata).

Em relação aos mamíferos, precisamente primatas, a análise do genoma de *Callithrix jacchus*, alvo contínuo de estudos originados e derivados de concepções evolutivas acerca de minúcias genéticas subjacentes ao comportamento (ver, por exemplo, Worley *et al.* 2014; Harris *et al.* 2014; Saito 2015; Garber *et al.* 2015), indicou seleção adaptativa na região cromossômica que inclui o gene *LNPEP* (Worley *et al.*, 2014). A motivação desses estudos envolvendo esse modelo primata recai sobre as características incomuns dos calitriquíneos e as decorrentes adaptações fisiológicas e comportamentais. Existe uma alta fecundidade, frequente gemelaridade, e a organização social inclui cooperação entre machos e fêmeas. Em cada grupo familiar há, em geral, apenas uma fêmea ativa reprodutivamente, considerada dominante. A ovulação, e consequente reprodução, é suprimida nas demais fêmeas (Aristide *et al.*, 2015).

A reprodução, de fato, está entre os parâmetros distintivos entre os três principais grupos que formam a classe Mammalia. Em uma classificação geral, pode-se distinguir duas subclasses: Prototheria e Theria. Prototheria, representada apenas pela ordem Monotremata, retém o caráter ancestral de postura de ovos. A subclasse Theria é ainda subdividida em duas infraclasses: Metatheria e Eutheria (Murphy *et al.*, 2001). Em Metatheria estão os chamados marsupiais; os mamíferos placentários, por outro lado, compõem a infraclasse Eutheria. As diferenças entre as infraclasses de mamíferos térios recaem sobre o desenvolvimento e a duração do período gestacional.

Embora nem todas as espécies em Metatheria apresentem marsúpio, todos os marsupiais nascem em um estágio de desenvolvimento precoce após um rápido período de crescimento embrionário (Murphy *et al.*, 2001; Gundling and Wildman, 2015). O desenvolvimento intrauterino permite a viviparidade dos mamíferos placentários. Como todos os amniotas, mamíferos térios têm fertilização interna. Adicionalmente, há um sistema reprodutivo especializado. A clivagem é única porque desenvolve uma população celular especializada no embrião que permite a integração no tecido uterino materno. Com essa integração, um novo órgão, a placenta, é formado e permite um desenvolvimento embrionário prolongado no útero (Wildman *et al.*, 2006). Marsupiais são também vivíparos; contudo, muito da organogênese do neonato ocorre durante a lactação – a gestação, na qual a nutrição materna é suprida pela placenta, é muito curta (cerca de duas semanas).

A forma de placentação é muito variável entre os mamíferos. Primatas antropoides geralmente exibem implantação superficial, com aderência à parede uterina pelo trofoblasto, mas sem penetração endometrial completa ou invasão das camadas mais profundas da parede uterina, como acontece em *Homo sapiens*. Humanos têm uma implantação mais invasiva, com remodelação completa dos vasos uterinos. Em todas as placentas de primatas antropoides, entretanto, a camada de trofoblasto fetal (ou córion) está em contato direto com o suprimento sanguíneo materno, o que é chamado de placentação hemocorial (Gundling and Wildman, 2015) Uma associação entre córion e membrana vitelina parece ser a condição ancestral em amniotas. Essa condição está presente em alguns gêneros vivíparos da ordem Squamata.

Os integrantes atuais da ordem Primates são classificados dentre as subordens Strepsirrhini e Haplorrhini, e a divergência entre esses clados ocorreu há aproximadamente 75 milhões de anos (Pozzi *et al.*, 2014). Strepsirrhini, ou prossímios, compreendem as infraordens Lemuriformes, Lorisiformes e Chiromyiformes. Integram esses grupos, respectivamente às infraordens citadas, os lêmures, galagos e aye-ayes (Mittermeier, R.A., Rylanads, A.B. & Wilson, 2013). Os prossímios conservam caracteres ancestrais dos mamíferos: hábito de vida solitário, comunicação especialmente dependente de sinais olfativos e vocais, presença de *tapetum lucidum* (membrana ocular que reflete a luz e melhora a visão em condições de baixa luminosidade) e baixa taxa metabólica (Mittermeier, R.A., Rylanads, A.B. & Wilson, 2013).

A segunda subordem primata, Haplorrhini, é composta pelas infraordens Tarsiiformes e Simiiformes. Os membros da infraordem Tarsiiformes, como *Carlito syrichta*, habitam o Sudeste Asiático. Apresentam características comuns a prossímios e a antropoides (Mittermeier, R.A., Rylanads, A.B. & Wilson, 2013). Entre as similaridades com Strepsirrhini, nota-se a presença de garras em alguns dedos dos pés (em Tarsiiformes, as garras estão no segundo e no terceiro dedos podais), múltiplos mamilos e útero bicornado. A interface placentária, no entanto, é hemocorial – análoga a de vários antropoides e distinta daquela encontrada em lêmures, por exemplo. As fêmeas passam por ciclos reprodutivos mensais (Martin, 2008). O período gestacional tem duração de seis meses e dá origem a apenas um filhote, que nasce com 30% do peso materno e é amamentado por dois meses. Durante esse tempo, no qual também desenvolve habilidades de captura de presas, o cuidado provido é exclusivamente materno – não há, portanto, registro de cuidado paternal ou aloparental (Mittermeier, R.A., Rylanads, A.B. & Wilson, 2013).

Apesar das similaridades com prossímios e antropoides, os Tarsiiformes são, em si, no mínimo insólitos: o tamanho extremamente grande dos olhos (que, a propósito, são maiores que o próprio cérebro) é absolutamente singular. Contudo, além dos olhos, existem ainda outras peculiaridades menos evidentes: embora sejam de hábito noturno, Tarsiiformes não possuem *tapetum lucidum*, mas sim uma fóvea retinal, característica de animais com atividade diurna. Dentre os Simiiformes, ou antropoides, o hábito diurno é generalizado: ainda que existam exceções, é o padrão da ampla maioria das espécies. O olfato é comparativamente reduzido em relação às outras infraordens, mas a visão é relativamente aprimorada (Mittermeier, R.A., Rylanads, A.B. & Wilson, 2013).

Entre os antropoides, a infância é prolongada, há geralmente um forte vínculo materno. Quanto à classificação, Simiiformes são divididos em duas parvordens: Catarrhini e Platyrrhini (veja a FIGURA 12). A parvordem Catarrhini inclui as superfamílias Cercopithecoidea (Macacos do Velho Mundo) e Hominoidea (hominídeos): compreende, portanto, todos os símios dos continentes asiático, africano e europeu, o que, é claro, inclui a espécie humana. Os Macacos do Velho Mundo constituem um grupo monofilético. Apesar de contar com apenas uma família, Cercopithecidae, são o grupo mais diverso entre os primatas. A superfamília Hominoidea, composta pelos “grandes primatas”, é dividida em duas famílias: Hylobatidae e Hominidae. Duas subfamílias compõem Hominidae: Ponginae e Homininae (Fleagle, 2013).

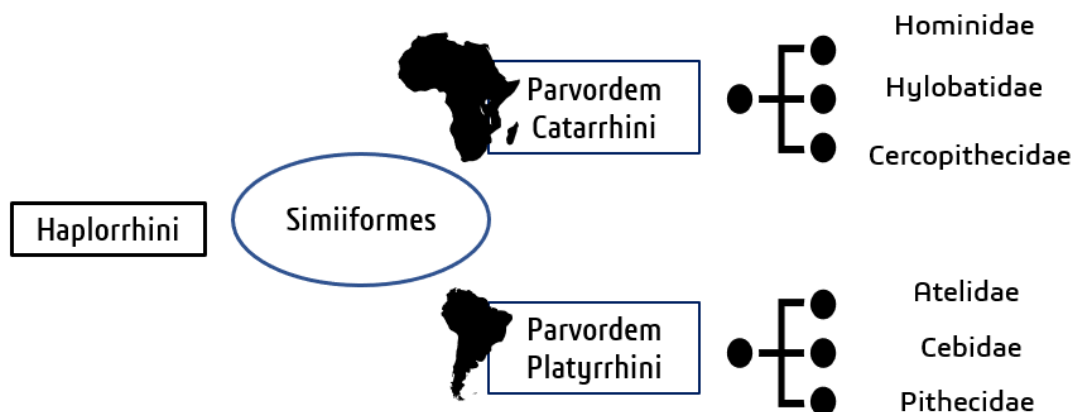


Figura 12. Classificação dos Simiiformes. Primatas antropoides do Velho e do Novo Mundo. Ilustração baseada nos estudos filogenéticos de Perelman et al. (2011), Schneider & Sampaio (2015) e Pozzi et al. (2014).

Os Platyrrhini, por fim, são agrupados em três famílias – Cebidae, Atelidae e Pitheciidae (Schneider and Sampaio, 2013). O grupo monofilético Callitrichinae, uma subfamília de Cebidae, apresenta características incomuns aos primatas antropóides. Além do pequeno tamanho e da baixa massa corporal (que costuma variar entre 115 e 600 gramas), calitriquíneos também são conhecidos pela gemelaridade, resultante da liberação de múltiplos óvulos por ciclo (Harris *et al.*, 2014). Há predomínio de gêmeos dizigóticos. A Tabela 4, abaixo, retrata a classificação da parvordem Platyrrhini, com destaque para a subfamília Callitrichinae (Cebidae).

Tabela 4. Classificação em Platyrrhini

| Famílias | Subfamília | Gêneros | Espécies | |
|-------------|-------------------|-----------------------|--------------|----|
| Cebidae | Callitrichinae | <i>Cebuella</i> | 1 | |
| | | <i>Mico</i> | 15 | |
| | | <i>Callithrix</i> | 6 | |
| | | <i>Callimico</i> | 1 | |
| | | <i>Saguinus</i> | 15 | |
| | | <i>Leontopithecus</i> | 4 | |
| | Saimirinae | <i>Saimiri</i> | 5 | |
| Pitheciidae | Cebinae | <i>Cebus</i> | 4 | |
| | Callicebinae | <i>Sapajus</i> | 8 | |
| | | Aotinae | <i>Aotus</i> | 10 |
| | | <i>Callicebus</i> | 28 | |
| | | <i>Pithecia</i> | 5 | |
| Pitheciinae | <i>Chiropotes</i> | 5 | | |
| Atelidae | Alouattinae | <i>Cacajao</i> | 3 | |
| | | <i>Alouatta</i> | 14 | |
| | | <i>Ateles</i> | 7 | |
| | Atelinae | <i>Brachyteles</i> | 2 | |
| | | <i>Lagothrix</i> | 5 | |

Tabela 4. Três famílias compõem a parvordem Platyrrhini. Em Cebidae, a subfamília Callitrichinae é caracterizada por fenótipos comportamentais e reprodutivos únicos entre os primatas.

1.4. JUSTIFICATIVA DO ESTUDO

Os membros da família do nonapeptídeo OXT exercem funções fisiológicas essenciais em vertebrados e até mesmo em invertebrados. Como vimos até aqui, diversas características reprodutivas e comportamentais em mamíferos são moduladas por OXT. Assim, outros genes envolvidos na biodisponibilidade desse hormônio nos tecidos-alvo devem apresentar padrões evolutivos relacionados. Diante dos elementos que compõem esse repertório genético, é preciso ressaltar a funcionalidade da oxitocinase. Dada a intensa interação de OXT e LNPEP, o gene *LNPEP* tornou-se, enfim, um nítido candidato para esse estudo. A evolução de *LNPEP* e de seu produto funcional em mamíferos continua praticamente inexplorada. Esta pesquisa aborda, então, a evolução molecular do gene *LNPEP* na classe Mammalia, com especial atenção à ordem Primates. Existiria alguma modificação molecular que implicaria diferenças estruturais e potencialmente funcionais em LNPEP? Em Primates, haveria variabilidade *taxa*-específica em *LNPEP* correspondente àquelas encontradas em *OXT* e *AVP*?

2. OBJETIVOS

2.1. OBJETIVO GERAL

Caracterizar a evolução molecular do gene *LNPEP* e da oxitocinase na ordem Primates e em outros grupos da classe Mammalia.

2.2. OBJETIVOS ESPECÍFICOS

Objetivo 1: Verificar as taxas evolutivas da região codificadora do gene *LNPEP*;

Objetivo 2: Inferir modificações estruturais de *LNPEP* a partir de sequências homólogas;

Objetivo 3: Investigar a variação em domínios funcionais de *LNPEP* em Primates, com especial atenção à parvordem Platyrrhini.

3. MÉTODOS

Em linhas gerais, todas as informações que fundamentaram a pesquisa estão publicamente dispostas em bancos de dados especializados. Já as análises foram desenvolvidas em ferramentas de bioinformática apropriadas, considerando sequências de DNA e de aminoácidos, estrutura e função proteicas e taxas evolutivas entre as espécies. Os resultados, exceto aqueles puramente descritivos, foram ponderados matematicamente.

3.1. CONJUNTO DE DADOS INICIAIS

As sequências de DNA da região codificadora do gene *LNPEP* foram obtidas no banco de dados NCBI GenBank (Benson et al., 2013). Quando necessário, as referências canônicas foram consultadas na base de dados UniProt (UniProt Consortium, 2018). Do UniProt e dos dados de DNA compilados também foram derivadas as sequências de aminoácidos da proteína LNPEP. A filogenia consenso das espécies foi verificada na base de dados TimeTree (Kumar S, Stecher G, Suleski M, 2017). Estruturas derivadas de cristalografia por difração de raio-X e depositadas no RCBS Protein Data Bank (<http://www.rcsb.org>) também compuseram os dados iniciais.

3.2. ANÁLISES EVOLUTIVAS

Alinhamentos

O alinhamento das sequências de DNA foi construído com o programa MAFFT (Kato, Rozewicki and Yamada, 2017) por método iterativo baseado na transformação de Fourier (FFT-NS-i), retendo *Homo sapiens* como referência. Obteve-se, também, a tradução das sequências com o programa AliView (Larsson, 2014), no qual o alinhamento foi editado manualmente para aprimorar a correspondência entre os códons e aumentar a acurácia dos resultados das etapas seguintes.

Verificação de seleção positiva

Para estimar as taxas evolutivas, duas abordagens foram implementadas no programa CODEML/PAML 4.9h (Yang, 2007), tendo como parâmetro a taxa de substituições não-sinônimas/sinônimas ($dN/dS = \omega$, sendo $\omega > 1$ indicativo de seleção positiva, enquanto a neutralidade é sugerida por um $\omega \cong 1$): (1) Modelo de seleção positiva em *NSsites*, que permite estimar diferentes valores de ω entre as colunas (ou seja, entre sítios) de um dado alinhamento múltiplo de DNA, enquanto admite uma taxa constante ao longo dos ramos da filogenia correspondente e (2) *Clade Model C*, que verifica a disparidade de ω entre um grupo filogenético estipulado previamente – denominado *foreground* – e o conjunto de táxons restantes, considerados *background*.

Em *NSsites* foram aplicados os modelos M8 (modelo alternativo), M7 e M8a (modelos nulos). O modelo M8 assume 11 categorias de ω para os sítios: 10 delas seguem uma distribuição beta com valores $\omega \leq 1$, enquanto a 11ª categoria permite $\omega > 1$ (isto é, admite seleção positiva). O modelo M8a possui também 11 categorias de valores ω , sendo 10 deles beta distribuídos e menores que 1 ($\omega < 1$), enquanto a 11ª categoria é fixada em 1 ($\omega = 1$). O modelo M7 admite 10 categorias de valores ω com distribuição beta, todos menores ou iguais a 1 ($\omega \leq 1$).

Para determinar se o modelo nulo deveria ser rejeitado em favor da admissão do modelo alternativo, que assume $\omega > 1$, o teste de razão de verossimilhança (LTR) foi aplicado comparando M8 – M7 e M8 – M8a com 2 graus e 1 grau de liberdade, respectivamente. A abordagem *Bayes Empirical Bayes* (BEB) (Yang Z, Wong WS, 2005), que calcula a probabilidade posterior de que um sítio pertença de fato à categoria de códons com $\omega > 1$, identificou os sítios sob seleção positiva.

Em *Clade Model C*, a linhagem correspondente a *Platyrrhini* foi indicada como *foreground*. Havendo dois tipos de clados, o modelo presume cinco parâmetros na distribuição ω ($p_0, p_1, \omega_0, \omega_2$, e ω_3) e $0 < \omega_0 < 1$. O modelo M22 (M2a_rel) foi o parâmetro nulo para o cálculo deste LRT, que assumiu 1 grau de liberdade, conforme recomendado Weadick and Chang (2012).

Filogenia

A filogenia de LNPEP, baseada no alinhamento dos aminoácidos, foi construída no programa MEGA X (Kumar *et al.*, 2018) por *Maximum Likelihood* com o modelo de substituição JTT (Jones, Taylor and Thornton, 1992). Uma distribuição Gamma discreta foi utilizada para modelagem das diferenças nas taxas evolutivas entre sítios com distribuição Gamma (5 categorias (+G, *parameter* = 0.5457)).

Reconstrução de sequências ancestrais

A filogenia e as sequências de aminoácidos foram as informações de entrada para a reconstrução das sequências ancestrais, hipotéticas, com o programa CODEML. O objetivo desta reconstrução foi acessar os eventos evolutivos (substituições de aminoácidos) ao longo das linhagens e na base de Primates. Adicionalmente, substituições foram avaliadas conforme a distância de Grantham (*Grantham score*), que considera três propriedades (composição, polaridade e volume molecular) para classificar as alterações como conservativas (0 – 50), moderadamente conservativas (51 – 100), moderadamente radicais (101 – 150) ou radicais (>151).

3.3. EFEITOS DE MUTAÇÕES SOBRE A ESTABILIDADE DA PROTEÍNA

Mutações referentes aos sítios com indícios de seleção positiva ou localizadas em regiões críticas para a função enzimática foram simuladas no programa FoldX com base em um modelo tridimensional da proteína canônica. A estabilidade de uma proteína é dada pela energia livre (ΔG), expressa em kcal/mol. Quanto menor a energia livre, mais estabilidade é atribuída à proteína. Por sua vez, $\Delta\Delta G$ é a diferença na energia livre entre os tipos selvagem (wt) e mutante (mut). Uma mutação que acrescenta energia ($\Delta\Delta G > 0$) tenderá a desestabilizar a estrutura, enquanto uma mutação que reduza a energia ($\Delta\Delta G < 0$) estabilizará o sistema.

O programa opera em duas etapas. Na primeira, a estrutura original PDB é otimizada pela função *RepairPDB*. Na segunda fase, a função *BuildModel* é utilizada para efetuar as mutações indicadas na estrutura reparada. A precisão reportada do programa FoldX é 0.46 kcal/mol. Ou seja, o desvio padrão da diferença entre $\Delta\Delta G$ s calculada pelo programa e os valores experimentais é de 0.46 kcal/mol. Sendo assim, é possível estimar o efeito de uma dada mutação assumindo sete categorias gradativas, com valores extremos correspondendo a uma mutação Altamente Estabilizadora (-3σ) ou Altamente Desestabilizadora (3σ).

4. RESULTADOS E DISCUSSÃO

4.1. CONJUNTO DE DADOS E CONSIDERAÇÕES INICIAIS

Foram compiladas para as análises as sequências da região codificadora do gene *LNPEP* de 45 táxons da classe Mammalia. As espécies e as famílias a que eles pertencem são apresentadas na Tabela 5, bem como as referências ao GenBank. Um único representante de Prototheria foi incluído. As demais espécies pertencem à infraclasse Theria. Da ordem Primates, estão presentes 26 espécies. A topologia consenso correspondente a todo esse conjunto de táxons é mostrada na FIGURA 13 (à esquerda da Tabela 5, na página seguinte).

O alinhamento dos nucleotídeos teve como referência a sequência de *Homo sapiens* (NM_005575.3). A FIGURA 14 é uma visão geral do alinhamento dos aminoácidos de LNPEP. É possível observar que as sequências são, em grande parte, conservadas. Na parte superior da figura, dois índices gráficos exibem as médias de hidrofobicidade e de identidade de cada sequência em relação àquela de *H. sapiens*.

Duas informações derivadas desses alinhamentos contrapõem o conhecimento corrente sobre LNPEP. A primeira é relativa ao sítio de clivagem para a liberação da forma solúvel da proteína, Phe154↓Ala155. De acordo com Pham *et al.* (2009), esse motivo seria restrito à família Hominidae. Contudo, essa conclusão foi derivada de um conjunto de 11 táxons, dos quais apenas dois são hominídeos (FIGURA 5, seção 1.2). No presente trabalho, conforme os dados detalhados na FIGURA 15, percebe-se que esse padrão não é uma exclusividade de Hominidae.

Tabela 5. Táxons, referências e classificação

| Espécie | Referência | CDS | Família |
|-----------------------------------|-----------------------|-----------------|-------------------|
| <i>Homo sapiens</i> | NM_005575.3 | 77-3154 | Hominidae |
| <i>Pan paniscus</i> | XM_014344937.1 | 35-3343 | |
| <i>Pan troglodytes</i> | XM_009449397.3 | 101-3382 | |
| <i>Gorilla gorilla</i> | XM_019027962.1 | 234-3311 | |
| <i>Pongo abelii</i> | XM_024246950.1 | 531-3608 | |
| <i>Nomascus leucogenys</i> | XM_003277197.3 | 531-3611 | Hylobatidae |
| <i>Mandrillus leucophaeus</i> | XM_011989410.1 | 533-3610 | Cercopithecoidea |
| <i>Cercocebus atys</i> | XM_012049641.1 | 106-3183 | |
| <i>Papio anubis</i> | XM_003899947.3 | 544-3621 | |
| <i>Theropithecus gelada</i> | XM_025387272.1 | 534-3611 | |
| <i>Macaca mulatta</i> | XM_001092287.3 | 534-3611 | |
| <i>Macaca fascicularis</i> | XM_005557431.2 | 1012-4089 | |
| <i>Macaca nemestrina</i> | XM_011745139.1 | 545-3622 | |
| <i>Chlorocebus sabaeus</i> | XM_007979244.1 | 543-3620 | |
| <i>Rhinopithecus roxellana</i> | XM_010357238.1 | 934-4011 | |
| <i>Rhinopithecus bieti</i> | XM_017858842.1 | 529-3606 | |
| <i>Ptilocolobus tephrosceles</i> | XM_023212204.1 | 529-3606 | Cercopithecoidea |
| <i>Colobus angolensis</i> | XM_011944296.1 | 529-3606 | |
| <i>Saguinus midas</i> | NA | NA | Cebidae |
| <i>Callithrix jacchus</i> | XM_002744784.3 | 907-3981 | |
| <i>Saimiri boliviensis</i> | XM_003920729.2 | 995-4072 | |
| <i>Cebus capucinus</i> | XM_017507323.1 | 913-3987 | |
| <i>Carlito syrichta</i> | XM_008055047.2 | 172-3249 | Tarsiidae |
| <i>Otolemur garnettii</i> | XM_003788938.3 | 45-3122 | Galagidae |
| <i>Propithecus coquereli</i> | XM_012638839.1 | 512-3589 | Indriidae |
| <i>Microcebus murinus</i> | XM_012773726.1 | 519-3596 | Cheirogaleidae |
| <i>Cavia porcellus</i> | XM_003479447.4 | 1782-4859 | Caviidae |
| <i>Rattus norvegicus</i> | NM_001113403.2 | 72-3149 | Muridae |
| <i>Mus musculus</i> | NM_172827.3 | 78-3155 | |
| <i>Mus caroli</i> | XM_021186305.1 | 78-3155 | |
| <i>Meriones unguiculatus</i> | XM_021654956.1 | 214-3291 | |
| <i>Nannospalax galili</i> | XM_008830542.2 | 132-3209 | Spalacidae |
| <i>Vicugna pacos</i> | XM_006212956.2 | 54-3134 | Camelidae |
| <i>Camelus dromedarius</i> | XM_010993973.1 | 582-3662 | |
| <i>Sus scrofa</i> | NM_001105291.1 | 598-3678 | Suidae |
| <i>Ovis aries</i> | XM_012121045.2 | 141-3221 | Bovidae |
| <i>Balaenoptera acutorostrata</i> | XM_007197483.1 | 688-3765 | Balaenopteridae |
| <i>Lipotes vexillifer</i> | XM_007466866.1 | 591-3668 | Lipotidae |
| <i>Tursiops truncatus</i> | XM_019929663.1 | 93-3164 | Delphinidae |
| <i>Orcinus orca</i> | XM_004263510.2 | 62-3133 | |
| <i>Desmodus rotundus</i> | XM_024563311.1 | 63-3140 | Phyllostomidae |
| <i>Eptesicus fuscus</i> | XM_008144521.2 | 129-3206 | Vespertilionidae |
| <i>Sarcophilus harrisii</i> | XM_012541125.2 | 196-3273 | Dasyuridae |
| <i>Phascolarctos cinereus</i> | XM_021009147.1 | 206-3286 | Phascolarctidae |
| <i>Ornithorhynchus anatinus</i> | XM_007669122.2 | 22-3057 | Ornithorhynchidae |

Figura 13. Topologia consenso (Esquerda). Representação da topologia consenso dos táxons exibidos na Tabela 4. Topologia baseada no banco de dados TimeTree (<http://www.timetree.org/>).

Tabela 5. Conjunto de dados. Diferentes ordens de Eutheria. Cinza: Primates; amarelo: Rodentia. Verde claro: Artiodactyla; laranja: Cetacea. Verde escuro: Chiroptera. Metatheria: *S. harrisii* e *P. cinereus*. Prototheria: *O. anatinus*. Cebidae, destacada entre os primatas, corresponde ao clado considerado *foreground* (ver Tabelas 8 e 9). CDS: *Coding sequence*; os números indicam início e fim da região codificadora. NA: sequência inferida a partir do exoma SRX039412 – *S. midas*).

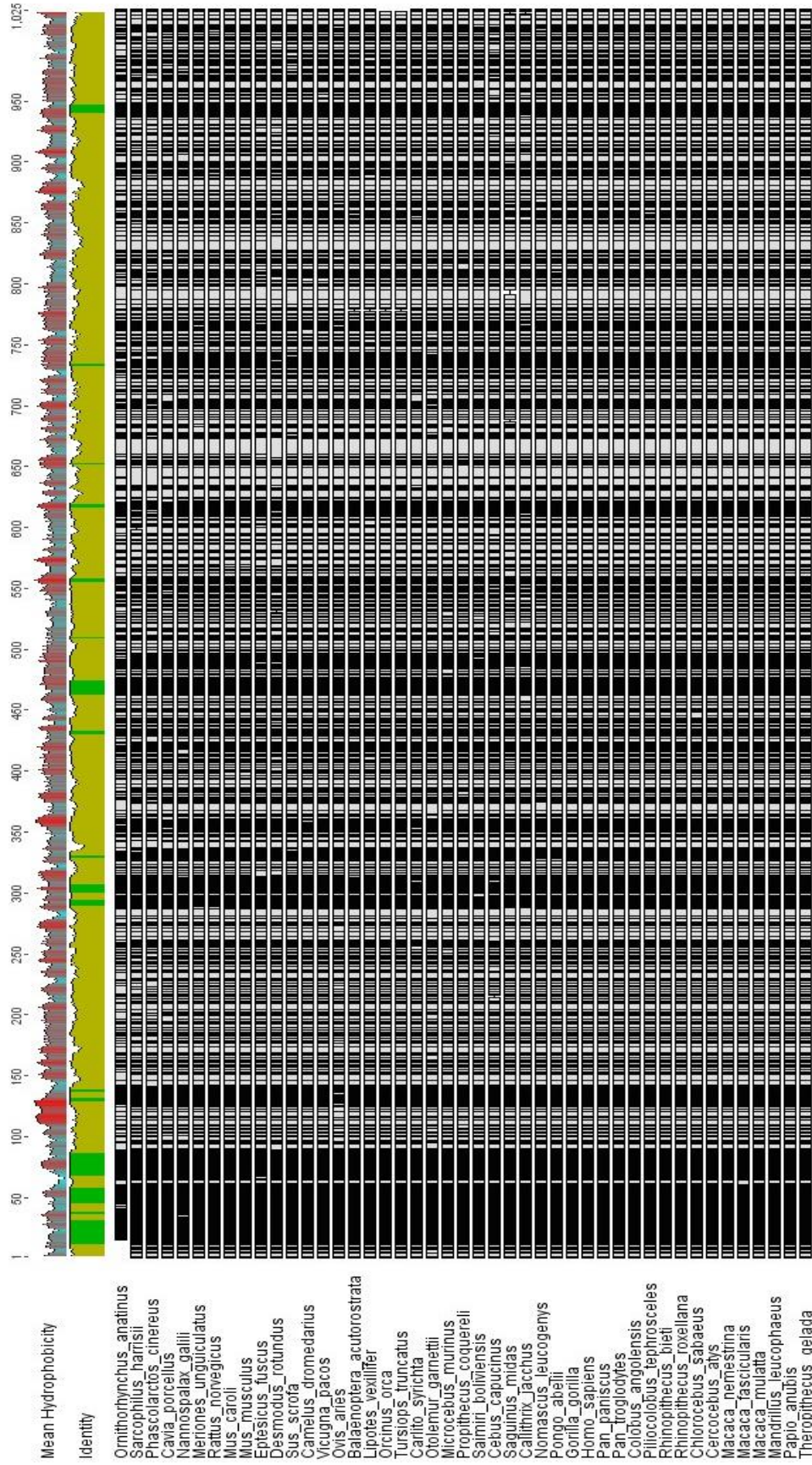


Figura 14. Alinhamento dos aminoácidos de LNPEP. Regiões escuras correspondem a segmentos com similaridade > 85% em relação à sequência de *Homo sapiens*. Hidrofobicidade média: os valores aumentam em uma escala de cores entre azul e vermelho. Identidade: segmentos verdes indicam identidade de 100% nas 45 espécies.

Abaixo, na FIGURA 15, as espécies que também compartilham o motivo de clivagem Phe154↓Ala155, em uma imagem semelhante à FIGURA 5, destacando a presença de Phe154 entre as mais variadas famílias.

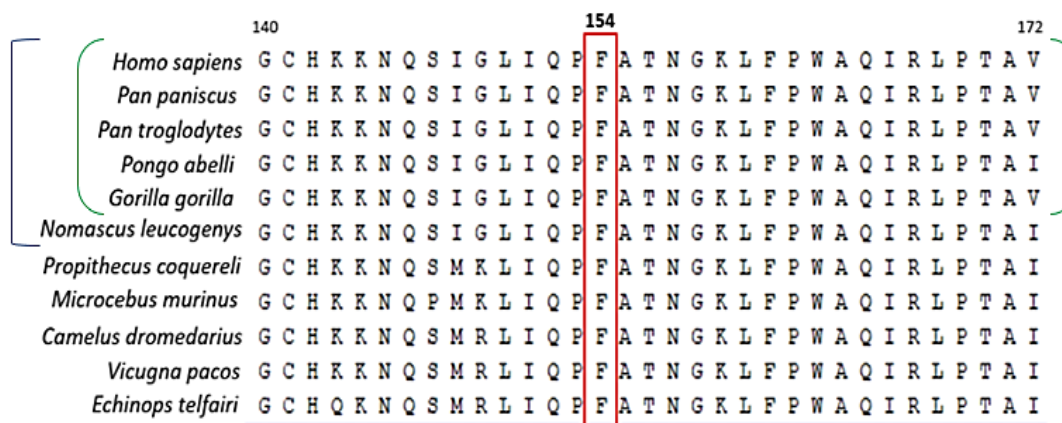


Figura 15. Phe154 não é restrito à família Hominidae. Em vermelho, a posição 154Phe(F). Homínidos estão mostrados em colchetes verdes. Está assinalado à esquerda, em azul, a superfamília Hominoidea.

Além de *Nomascus leucogenys* (família Hylobatidae), os estrepsirrinos *Propithecus coquereli* e *Microcebus murinus* (famílias Indriidae e Cheirogaleidae, respectivamente), e os artiodáctilos *Camelus dromedarius* e *Vicugna pacos* (Camelidae) compartilham o motivo. Uma pesquisa adicional nos bancos de dados mostra que Phe154 está presente também em *Echinops telfairi* (XM_004703474.1, não incluído nas demais análises). *E. telfairi* é uma espécie do clado Afrotheria (família Afrosoricida).

Segundo Pham et al. (2009), a ausência de IRAP/LNPEP na circulação de ratos e de outras espécies no período gestacional seria uma consequência da impossibilidade de conversão à forma solúvel da oxitocinase a partir de sua forma integral ligada à membrana. Por essa hipótese, a clivagem da oxitocinase seria dependente de uma protease que reconheceria apenas e especificamente o motivo Phe154↓Ala155. Não está claro, contudo, se existe de fato essa inabilidade de conversão à forma solúvel caso o motivo não seja conservado ou se todas as espécies listadas na FIGURA 15, e certamente outras mais, também apresentam LNPEP sérica durante a gestação.

A segunda informação relevante é a existência de variantes no motivo de ligação ao substrato, GAMEN, precisamente no sítio 428. Esse motivo é considerado conservado em todos os mamíferos. Assim, dentre a quase absoluta homogeneidade da sequência GAMEN entre as espécies, as dissonâncias se destacaram facilmente no alinhamento. Variantes foram descobertas nas espécies *Callithrix jacchus* e *Saguinus midas*, nas quais há uma substituição da glicina por uma valina (G428V), e em *Saimiri boliviensis*, espécie na qual uma alanina ocupa a posição (G428A). Curiosamente, a variante G428A também é observada no marsupial *Phascolarctos cinereus*, o popular coala. As estruturas são mostradas na FIGURA 16.

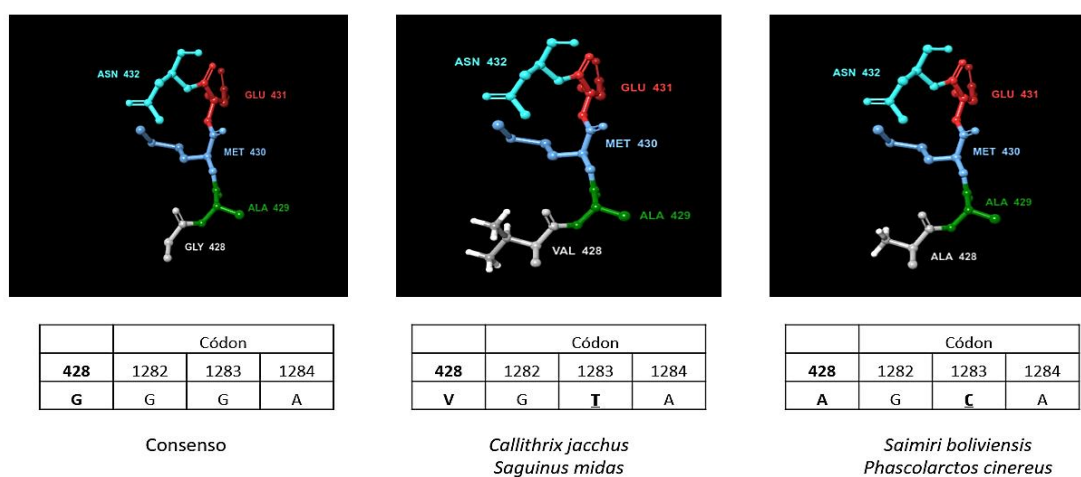


Figura 16. Variantes no motivo GAMEN. G428V em *C. jacchus* e *S. midas* e G428A em *Saimiri boliviensis* e *P. cinereus*.

As modificações nesses motivos são provavelmente funcionais, dadas as particularidades reprodutivas dessas espécies e também suas variantes *taxa*-específicas em peptídeos substratos. Essa questão é um interessante tema para estudos futuros. É também essencial ressaltar que as substituições mostradas na figura anterior são restritas à enzima LNPEP: *C. jacchus*, *S. midas*, *Saimiri boliviensis* e *P. cinereus* mantêm a habitual conservação nas proteínas ERAP1 e ERAP2, que também possuem o motivo de exopeptidase GAMEN. Esse fato dá sustentação à ideia de que a variação em LNPEP está associada àquelas em seus substratos, notavelmente nos peptídeos da família OXT/AVP.

A FIGURA 17 exemplifica a restrição das variantes G428V e G428A à LNPEP. As sequências na imagem são de *Callithrix jacchus* e *Saimiri boliviensis*.

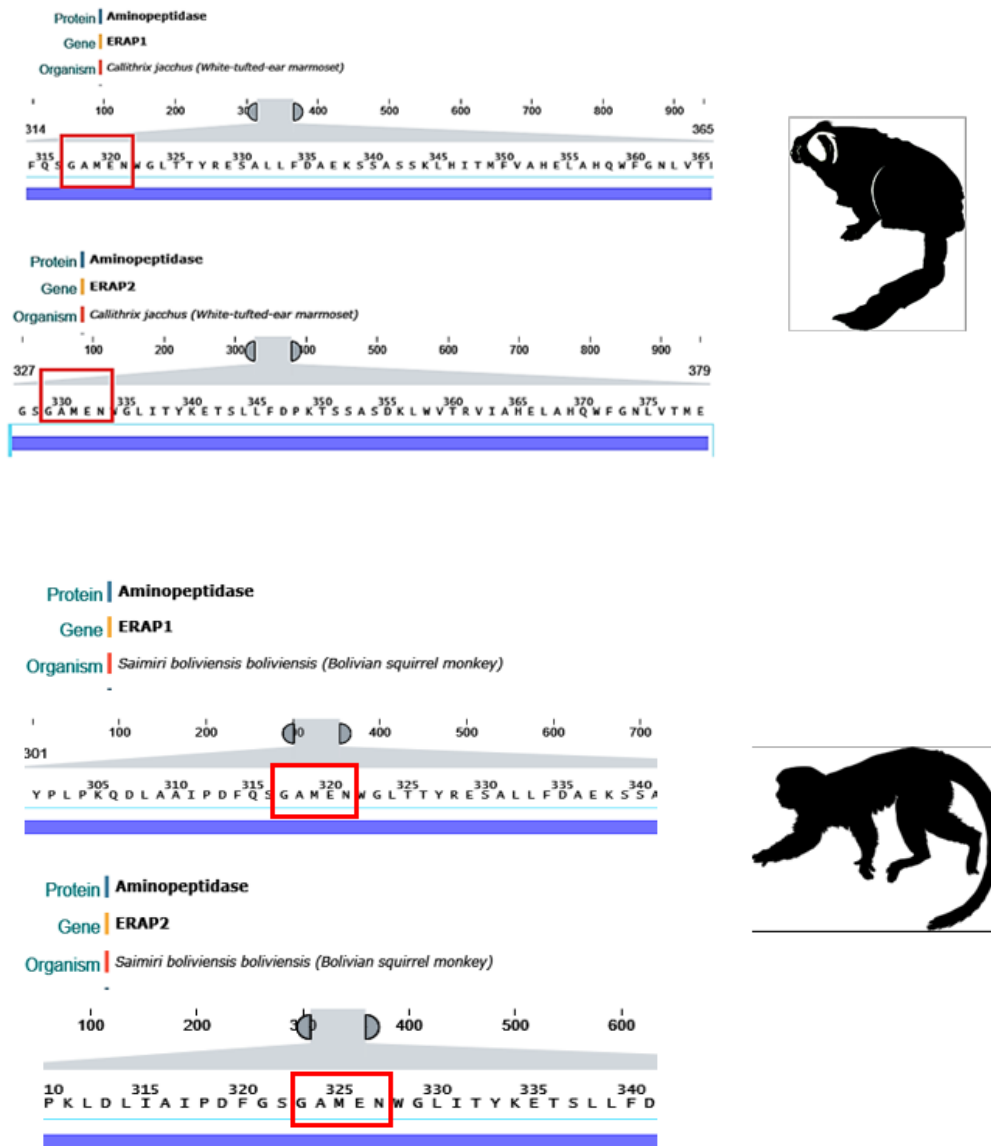


Figura 17. GAMEN em ERAP1 e ERAP2. Superior: *C. jacchus*. Inferior: *Saimiri boliviensis*. A glicina na posição 428 é mantida em ERAP1 e ERAP2 (apresentam o motivo GAMEN conservado). As substituições G428V e G428A só estão presentes em LNPEP (VAMEN em *C. jacchus* e AAMEN em *S. boliviensis*).

4.2. ANÁLISE DAS TAXAS EVOLUTIVAS

A Tabela 6 apresenta os resultados dos testes no programa CODEML para os modelos M7 e M8a (nulos) e M8 (alternativo). Com os valores de significância apresentados, os dois testes permitem admitir os resultados do modelo M8. Os sítios sob seleção positiva indicados por esse modelo estão, então, listados na Tabela 7. O valor médio de dN/dS , considerando todos os sítios, foi $\omega = 0.2687$.

Tabela 6. Resultados dos modelos M8, M7 e M8a em *NSsites*

| | lnL | Número de Parâmetros | Valor P |
|----------------|---------------|-----------------------------|------------------|
| lnL Modelo M8 | -24951.853675 | 93 | |
| lnL Modelo M7 | -24983.914171 | 91 | |
| lnL Modelo M8a | -24966.237970 | 92 | |
| LRT M7 M8 | 64.1210 | 2 | 1.188e-14 |
| LRT M8a M8 | 28.7686 | 1 | 8.156e-08 |

Tabela 6. Modelos M8, M7 e M8a. Valores de cada modelo em *NSsites*, teste da razão de verossimilhança (LRT) e a significância estatística.

Tabela 7. Sítios sob seleção positiva segundo o modelo M8

| Posição e Aminoácido¹ | BEB | dN/dS ± Desvio Padrão |
|---|------------|------------------------------|
| 338 V | 0.977* | 1.486 ± 0.094 |
| 793 T | 0.976* | 1.485 ± 0.098 |
| 831 N | 0.959* | 1.474 ± 0.129 |
| 838 T | 0.958* | 1.474 ± 0.128 |
| 882 I | 0.984* | 1.490 ± 0.080 |
| 884 I | 0.983* | 1.490 ± 0.081 |
| 1023 W | 0.998** | 1.499 ± 0.028 |

Tabela 7. Sítios sob seleção positiva. Valores de cada modelo em *NSsites*, teste da razão de verossimilhança (LRT) entre modelos e a significância estatística. BEB: probabilidade posterior de acordo com a análise Bayes Empirical Bayes. *: $P > 95\%$; **: $P > 99\%$. $\omega = dN/dS$.
¹Aminoácidos correspondentes à sequência de *Homo sapiens*.

A FIGURA 18 mostra os valores de dN/dS estimados para cada sítio. O valor médio de $\omega = 0.2687$ reflete a extrema conservação da enzima.

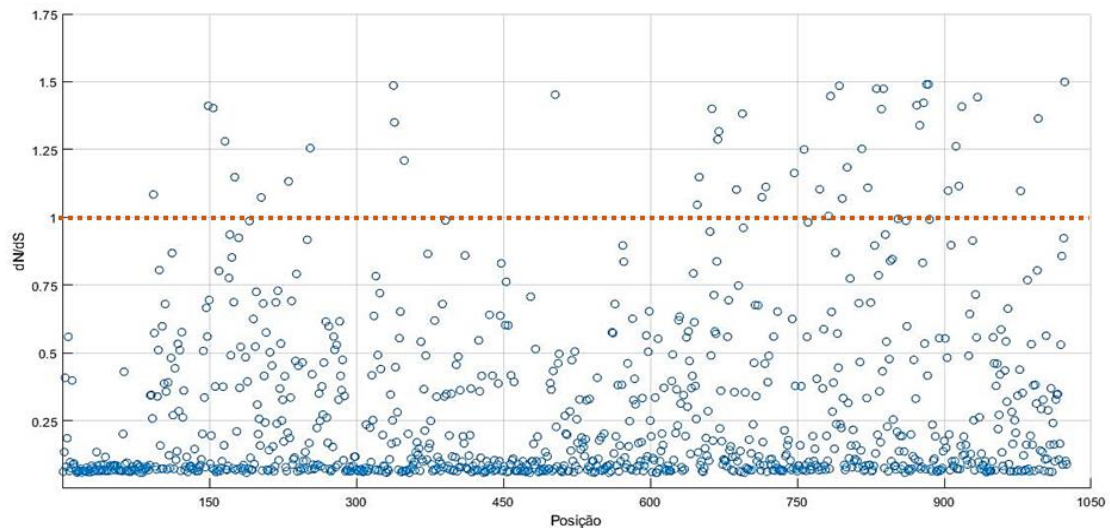


Figura 18. Os valores de ω em LNPEP. A linha laranja indica o valor $\omega = 1$, o que indicaria idealmente a neutralidade. A maior parte dos sítios está próxima a 0.

Considerando a localização dos sítios sob forte indício de seleção adaptativa, chegamos à FIGURA 19. Nela, podemos observar que a maioria dos sítios está localizada no domínio ERAP1_C. Apenas um sítio pertence à região M1_N. Um aspecto interessante, e que ressalta a importância das variantes descobertas em GAMEN, é a ausência de qualquer sítio com valor de $\omega > 1$ no domínio de peptidase M1. Ou seja, são variações em regiões extremamente conservadas. Mais uma vez, podemos deduzir uma forte implicação funcional.

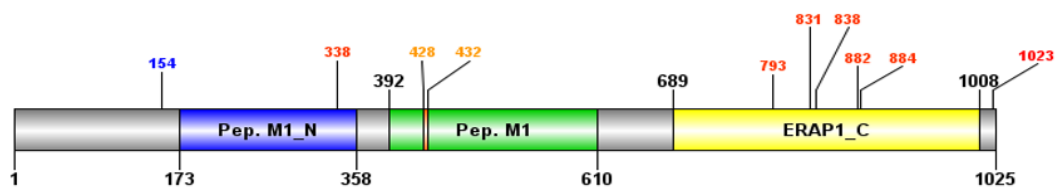


Figura 19. Localização dos sítios sob seleção positiva. A ampla maioria está em ERAP1_C. Um sítio apenas pertence à região M1_N. Sítios com valores de $\omega > 1$ estão ausentes no domínio de Peptidase M1, onde está localizado o motivo GAMEN.

A FIGURA 20 complementa a figura anterior. Os mesmos sítios são exibidos e as cores dos domínios são preservadas.

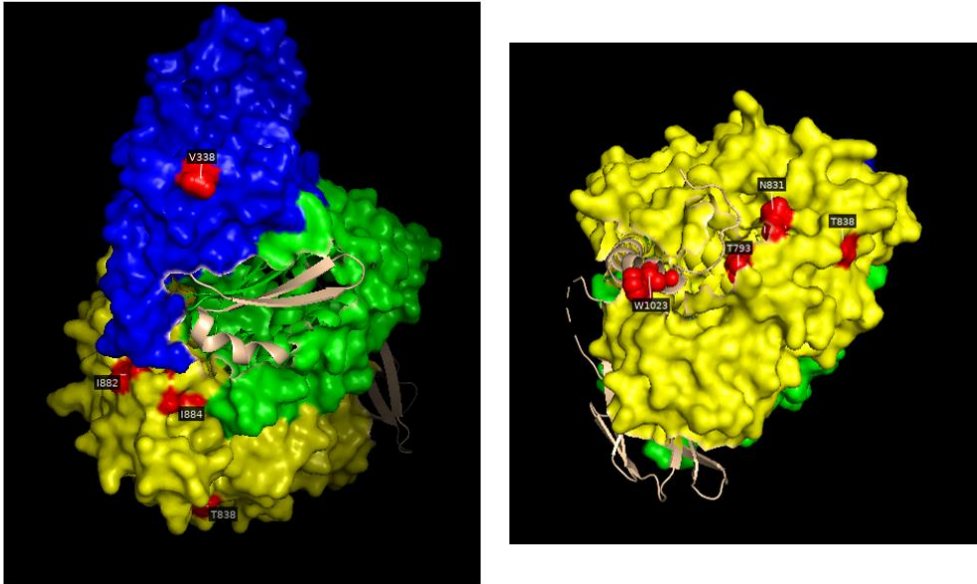


Figura 20. Visão estrutural de LNPEP. Os pontos vermelhos apontam os sítios que mostraram valor ω significativamente alto. Os domínios estão indicados com as mesmas cores da Figura 11.

Os resultados do teste de razão de verossimilhança entre Clade Model C e M22 são exibidos na Tabela 8. A proporção de sítios em cada categoria nas linhagens Platyrrhini (*foreground*) e as demais espécies (*background*) é mostrada na Tabela 9. O valor significativo em LRT entre M22 e Clade Model C permite assumir os resultados da Tabela 9. As linhagens em Platyrrhini possuem uma proporção maior de sítios sob seleção adaptativa em relação às demais espécies de mamíferos. Essa informação corrobora a ideia de que LNPEP compõe o cenário adaptativo desses primatas em relação às funcionalidades do sistema de regulação dos peptídeos OXT/AVP.

Tabela 8. Resultados estatísticos para o modelo Clade C

| | | np | Valor P |
|--------------------------------|---------------|-----------|--------------------|
| lnL Clade Model C | -24892.337161 | 94 | |
| lnL M22 (M2a_rel) | -24967.093093 | 93 | |
| LRT M22 Clade Model C | 149.5119 | 1 | < 0.0001 |

Tabela 8. *Clade Model C*. Valores do teste Clade Model C e M22, teste de razão de verossimilhança (LRT) e significância estatística.

Tabela 9. Proporção de sítios sob diferentes tipos de seleção

| | Classe 0 | Classe 1 | Classe 2 |
|-------------------|-----------------|-----------------|-----------------|
| Proporção | 0.49544 | 0.12360 | 0.38096 |
| Background | 0.02471 | 1.00000 | 0.25165 |
| Foreground | 0.02471 | 1.00000 | 1.60764 |

Tabela 9. Valores para cada classe em *Clade Model C*. Classe 0: sítios conservados ($0 < \omega < 1$); Classe 1: sítios neutros ($\omega = 1$); Classe 2: sítios sob seleção positiva nas diferentes linhagens. *Foreground*: Platyrrhini; *Background*: restante das espécies.

5.3. FILOGENIA EM MAMMALIA E ANCESTRAIS EM PRIMATES

A FIGURA 21 corresponde à filogenia estimada de LNPEP nas 45 espécies relacionadas. Para investigar em quais pontos ao longo da evolução primata provavelmente ocorreram as substituições nos sítios 154, 428 e naqueles sob seleção positiva, o programa CODEML foi utilizado desta vez a fim de reconstruir sequências ancestrais na ordem Primates.

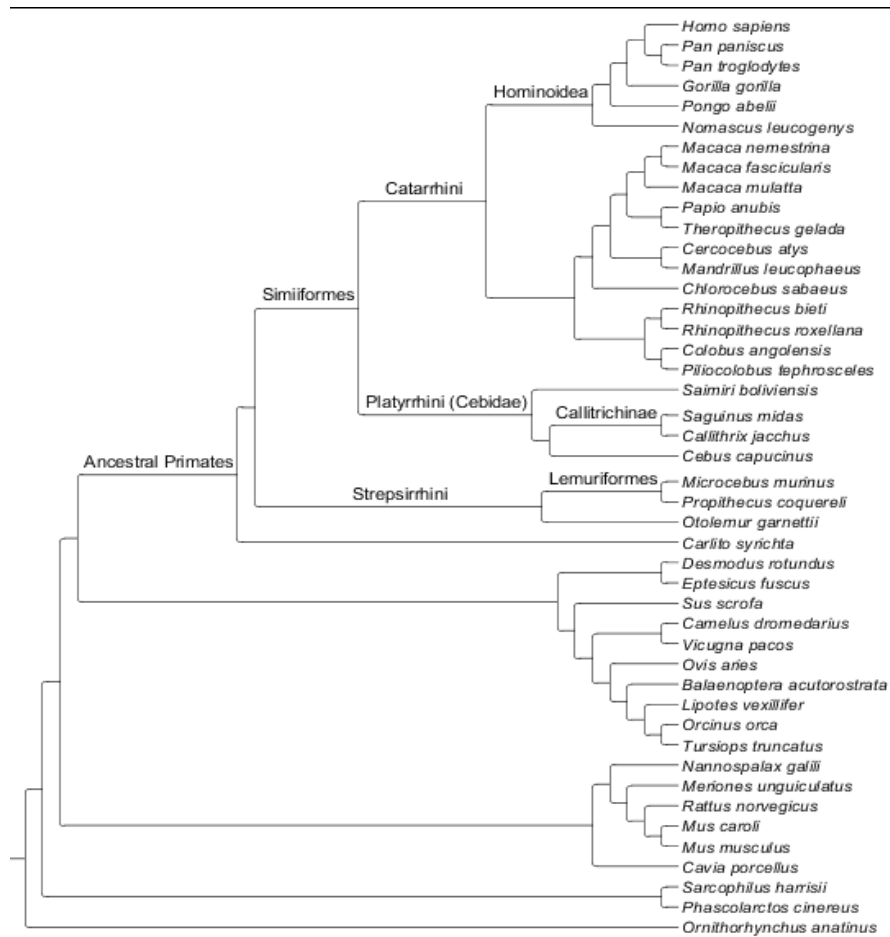


Figura 21. Filogenia de LNPEP. As linhagens ancestrais, denominadas ao longo dos ramos, foram reconstruídas.

A Tabela 10 exibe as prováveis modificações que ocorreram, portanto, nos sítios 154, 428, 338, 831, 838, 882 e 884 durante a evolução das linhagens apontadas na figura anterior. O resíduo W1023 não sofreu variação entre esses táxons; por esse motivo, não está relacionado.

Tabela 10. Substituições não sinônimas entre as linhagens de primatas

| Resíduo | Substituições | Grantham score | Classificação |
|-------------|-------------------------------------|----------------|-------------------|
| L154 | L154F em Strepsirrhini | 22 | Conservativa |
| | L154F em Hominoidea | 22 | Conservativa |
| G428 | G428V em Callitrichinae | 109 | Mod. Radical |
| | G428A em <i>Saimiri boliviensis</i> | 60 | Mod. conservativa |
| L338 | L338V em Catarrhini | 32 | Conservativa |
| T793 | T793V em Callitrichinae | 69 | Mod. conservativa |
| S831 | S831N em Catarrhini | 46 | Conservativa |
| | S831K em Platyrrhini | 121 | Mod. radical |
| T838 | L838A em Strepsirrhini | 96 | Mod. conservativa |
| V882 | V882M em Strepsirrhini | 21 | Conservativa |
| | V882I em Catarrhini | 29 | Conservativa |
| | V882N Platyrrhini | 133 | Mod. radical |
| V884 | V884I em Simiiformes | 29 | Conservativa |
| | I884S em Callitrichinae | 142 | Mod. radical |

Tabela 10. Substituições não-sinônimas e Grantham *score*. As sequências ancestrais foram reconstruídas com o programa CODEML. Classificação das substituições segundo a distância de Grantham: conservativas (0 – 50), moderadamente conservativas (51 – 100), moderadamente radicais (101 – 150) ou radicais (>151).

A distância de Grantham (Grantham *score*) considera composição, polaridade e volume para categorizar as alterações em conservativas (0 – 50), moderadamente conservativas (51 – 100), moderadamente radicais (101 – 150) ou radicais (>151). A substituição de glicina por uma valina na posição 428, ocorrida no ancestral Callitrichinae, foi considerada moderadamente radical. Outra informação mostrada, derivada dessa análise, é que a substituição de leucina por fenilalanina na posição 154 ocorreu duas vezes, independentemente: nos ancestrais Strepsirrhini e Hominoidea.

5.4. EFEITO DE MUTAÇÕES SOBRE A ESTABILIDADE

A finalidade desse experimento foi verificar o efeito de cada mutação sobre a estabilidade da estrutura. Foram consideradas as mutações presentes em sítios que apresentaram valores significativos de seleção positiva. Também foram consideradas as variantes do sítio de ligação ao substrato (G428V e G428A). A estrutura utilizada é identificada pelo PDB ID 5MJ6 (obtida por cristalografia de raio X com resolução de 2.53Å), que primeiramente foi reparada pelo programa. As mutações foram classificadas em sete categorias, de acordo com o efeito sobre a estabilidade:

1. Mutação Altamente Estabilizadora
2. Mutação Estabilizadora
3. Mutação Ligeiramente Estabilizadora
4. Mutação Neutra
5. Mutação Ligeiramente Desestabilizadora
6. Mutação Desestabilizadora
7. Mutação Altamente Desestabilizadora

Tal classificação é baseada nos respectivos valores de $\Delta\Delta G$ e no desvio padrão do programa FoldX (0.46 kcal/mol). Veja a Tabela 11.

Tabela 11. Valores atribuídos a cada categoria de efeito sobre a estabilidade








| | | |
|-------------|---|---|
| Categoria 1 |  | $\Delta\Delta G < -1.84$ kcal/mol |
| Categoria 2 |  | -1.84 kcal/mol $\leq \Delta\Delta G < -0.92$ kcal/mol |
| Categoria 3 |  | -0.92 kcal/mol $\leq \Delta\Delta G < -0.46$ kcal/mol |
| Categoria 4 |  | -0.46 kcal/mol $< \Delta\Delta G \leq +0.46$ kcal/mol |
| Categoria 5 |  | $+0.46$ kcal/mol $< \Delta\Delta G \leq +0.92$ kcal/mol |
| Categoria 6 |  | $+0.92$ kcal/mol $< \Delta\Delta G \leq +1.84$ kcal/mol |
| Categoria 7 |  | $\Delta\Delta G > +1.84$ kcal/mol |

Tabela 11. Para melhor visualização dos resultados expostos, a cada grupo mostrado foi atribuída uma cor, que se repete na página seguinte.

Tabela 12. Resultados do programa FoldX

| Sítio | Aminoácido (ref.) | Substituição | $\Delta\Delta G$ | Categoria |
|-------|-------------------|-----------------|------------------|-----------|
| 338 | Valina | Isoleucina | 0.14193 | 4 |
| | | Leucina | 0.901619 | 5 |
| | | Prolina | 0.925694 | 6 |
| | | Serina | 0.0969644 | 4 |
| | | Treonina | 0.0293147 | 4 |
| 793 | Treonina | Alanina | -0.0214693 | 4 |
| | | Glicina | -0.489127 | 3 |
| | | Isoleucina | 0.499434 | 5 |
| | | Metionina | 0.92715 | 6 |
| | | Asparagina | 0.333147 | 4 |
| | | Serina | 0.24058 | 4 |
| | | Valina | 0.239702 | 4 |
| 831 | Asparagina | Ácido Aspártico | -0.101626 | 4 |
| | | Glicina | 0.842964 | 5 |
| | | Histidina | -0.369588 | 4 |
| | | Lisina | -0.790292 | 3 |
| | | Arginina | 0.534162 | 5 |
| | | Serina | 0.111673 | 4 |
| 838 | Treonina | Alanina | 0.696329 | 5 |
| | | Ácido glutâmico | 0.970002 | 6 |
| | | Lisina | 1.61706 | 6 |
| | | Metionina | 1.58922 | 6 |
| | | Glutamina | 1.41637 | 6 |
| | | Valina | 1.2084 | 6 |
| 882 | Isoleucina | Alanina | -0.373249 | 4 |
| | | Glicina | -0.0560144 | 4 |
| | | Metionina | 0.599108 | 5 |
| | | Asparagina | -0.347045 | 4 |
| | | Serina | -0.0923192 | 4 |
| | | Treonina | -0.300854 | 4 |
| | | Valina | -0.311314 | 4 |
| 884 | Isoleucina | Glicina | -1.49113 | 2 |
| | | Leucina | -0.765225 | 3 |
| | | Metionina | 0.210425 | 4 |
| | | Serina | -1.25114 | 2 |
| | | Valina | -1.06497 | 2 |
| 1023 | Triptofano | Alanina | -0.0129085 | 4 |
| | | Fenilalanina | 0.315636 | 4 |
| | | Lisina | 0.219383 | 4 |
| | | Leucina | 0.171284 | 4 |
| | | Glutamina | 0.120437 | 4 |
| | | Serina | -0.00556833 | 4 |
| | | Tirosina | 0.267592 | 4 |
| 428* | Glicina | Alanina | -4.07949 | 1 |
| | | Valina | -9.83335 | 1 |

Tabela 12. Resultados do programa FoldX. Valores estimados de $\Delta\Delta G$ para cada mutação. As cores correspondem àquelas vistas na Tabela 11.

A estabilidade de uma proteína é definida pela energia livre, expressa em kcal/mol. Quanto menor o valor de ΔG , mais estável é a estrutura. $\Delta\Delta G$ é a diferença de energia entre a forma canônica e aquela resultante de uma determinada mutação. Conforme os resultados dispostos na Tabela 12, a maioria das mutações encontradas nos sítios sob seleção positiva foi considerada neutra em relação ao efeito sobre a estabilidade. Entre os sítios com $\Omega > 1$, 55.81% das mutações são neutras nesse contexto. No entanto, pode-se perceber a disparidade das mutações no sítio 428, pertencente ao motivo de ligação ao substrato GAMEN. As variantes G428A e G428V possuem valores que se distanciam de 99.9% dos achados, e configuram um incremento significativamente alto na estabilidade estrutural. Por outro lado, não foram encontradas variantes consideradas altamente desestabilizadoras.

A evolvabilidade – a capacidade de evolução contida em uma dada proteína – é favorecida pela estabilidade. Embora a seleção continue neutra sobre a função proteica, existe uma grande importância indireta: a estabilidade confere à proteína uma robustez mutacional, ou seja, um “amortecimento” do efeito suscitado por variações entre os aminoácidos e que poderiam refletir na estrutura. Proteínas mais estáveis toleram melhor possíveis mutações. O que decorre desse fato é o que em última análise importa nesse contexto: algumas das mutações que são desestabilizadoras são também, ocasionalmente, benéficas; mas só poderiam ser sustentadas e mantidas por uma proteína relativamente estável.

6. CONSIDERAÇÕES FINAIS E CONCLUSÃO

“Il faut imaginer Sisyphe heureux”.

– Albert Camus, em *Le Mythe de Sisyphe*.

Esta pesquisa assumiu uma abordagem biológica puramente baseada em modelos computacionais. Tal metodologia é justificada perante o ineditismo da temática e da ausência de dados efetivos. Os modelos e análises construídos e verificados neste estudo permitiram que questões relevantes fossem originadas; tais tópicos podem ser temas para estudos futuros. Sendo assim, a partir de diferentes experimentos e de uma amostra maior e mais variada de espécies, as seguintes questões podem ser elucidadas: (i) As variações do gene *LNPEP* e da oxitocinase possuem implicação funcional sobre as particularidades reprodutivas da subfamília Callitrichinae (Primates, família Cebidae)?; (ii) As variantes de *LNPEP* coevoluem com aquelas observadas em seus peptídeos substratos?; (iii) *LNPEP* e padrões morfológicos da placenta (interface, formato e interdigitação) entre as ordens e famílias da classe Mammalia estariam correlacionados?

O objetivo principal deste estudo foi caracterizar o gene *LNPEP* e a enzima oxitocinase entre primatas e demais mamíferos. Para isso, a pesquisa avaliou as taxas evolutivas da região codificadora do gene e construiu um panorama da variabilidade entre suas sequências homólogas. Foi demonstrada a existência de variações em regiões críticas para a funcionalidade catalítica da oxitocinase. Essas modificações implicam alterações relativas à estabilidade estrutural da proteína e indicam uma alteração funcional *taxa*-específica. A exclusividade atribuída à família Hominidae em relação à presença da forma placentária da enzima foi descartada. Por fim, a parvordem Platyrrhini obteve mais uma vez uma relevância evolutiva em proteínas que interagem com os nonapeptídeos oxitocina e vasopressina. Os resultados mostraram que, na subfamília Callitrichinae, ao menos duas modificações não-sinônimas ocorreram no motivo responsável pela ligação ao substrato, o que é no mínimo instigante ao considerar a extrema conservação dessa região não apenas em Mammalia, mas em todo subfilo Vertebrata.

7. REFERÊNCIAS

- Albers, H. E. (2012) 'The regulation of social recognition, social communication and aggression: Vasopressin in the social behavior neural network', *Hormones and Behavior*, 61(3), pp. 283–292. doi: 10.1016/j.yhbeh.2011.10.007.
- Albiston, A. *et al.* (2010) 'Phenylalanine-544 Plays a Key Role in Substrate and Inhibitor Binding by Providing a Hydrophobic Packing Point at the Active Site of Insulin-Regulated Aminopeptidase', *Molecular Pharmacology*, 78. doi: 10.1124/mol.110.065458.
- Albiston, A. L. *et al.* (2003) 'AT(4) receptor is insulin-regulated membrane aminopeptidase: potential mechanisms of memory enhancement', *Trends in Endocrinology & Metabolism*, 14.
- Aristide, L. *et al.* (2015) 'Modeling lineage and phenotypic diversification in the New World monkey (Platyrrhini, Primates) radiation', *Molecular Phylogenetics and Evolution*. Elsevier Inc., 82(PB), pp. 375–385. doi: 10.1016/j.ympev.2013.11.008.
- Babb, P. L. (2012) 'Molecular evolution of vasopressin and oxytocin receptor genes in owl monkeys (*Aotus azarai*) of northern Argentina', *ProQuest Dissertations and Theses*, p. 230.
- Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2013) 'GenBank', *Nucleic Acids Research*, 41(Database Issue), pp. D36–D42. doi: <http://doi.org/10.1093/nar/gks1195>.
- Ebstein, R. P. *et al.* (2012) 'The contributions of oxytocin and vasopressin pathway genes to human behavior', *Hormones and Behavior*. Academic Press, 61(3), pp. 359–379. doi: 10.1016/J.YHBEH.2011.12.014.
- Elkins, E. A. *et al.* (2017) 'Identification of an oxytocinase/vasopressinase-like leucyl-cystinyl aminopeptidase (LNPEP) in teleost fish and evidence for hypothalamic mRNA expression linked to behavioral social status', *General and Comparative Endocrinology*. Academic Press, 250, pp. 58–69. doi: 10.1016/J.YGCEN.2017.06.002.
- Feldman, R. *et al.* (2016) 'Oxytocin Pathway Genes: Evolutionary Ancient System Impacting on Human Affiliation, Sociality, and Psychopathology', *Biological Psychiatry*, 79(3), pp. 174–184. doi: 10.1016/j.biopsych.2015.08.008.
- Fernando, R. *et al.* (2005) 'Distribution and cellular localization of the insulin regulated aminopeptidase (IRAP) in the rat central nervous system', *J. Comp. Neurol.*, 487.
- Fleagle, J. G. (2013) *Primate adaptation and evolution*. 3th ed. Elsevier.
- Frank, E. and Landgraf, R. (2008) 'The vasopressin system — From antidiuresis to psychopathology', *European Journal of Pharmacology*, 583(2–3), pp. 226–242. doi: 10.1016/j.ejphar.2007.11.063.
- Garber, P. a *et al.* (2015) 'Tamarins: Insights into monogamous and non-monogamous single female social and breeding systems.', *American journal of primatology*, (November 2014), pp. 1–17. doi: 10.1002/ajp.22370.

- Gazis, D. (1978) 'Plasma half-lives of vasopressin and oxytocin analogs after iv injection in rats.', *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*, 158(4), pp. 663–5.
- Gimpl, G. and Fahrenholz, F. (2010) 'The Oxytocin Receptor System : Structure , Function , and Regulation', 81(2), pp. 629–683.
- Gundling, W. E. and Wildman, D. E. (2015) 'A review of inter- and intraspecific variation in the eutherian placenta', *Philosophical Transactions of the Royal Society B: Biological Sciences*. The Royal Society, 370(1663), pp. 20140072–20140072. doi: 10.1098/rstb.2014.0072.
- Haas, B. W., Anderson, I. W. and Smith, J. M. (2013) 'Navigating the complex path between the oxytocin receptor gene (OXTR) and cooperation: an endophenotype approach.', *Frontiers in human neuroscience*, 7(November), p. 801. doi: 10.3389/fnhum.2013.00801.
- Harris, R. A. *et al.* (2014) 'Evolutionary genetics and implications of small size and twinning in callitrichine primates.', *Proceedings of the National Academy of Sciences of the United States of America*, 111(4), pp. 1467–72. doi: 10.1073/pnas.1316037111.
- Heller, H. (1972) 'The effect of neurohypophyseal hormones on the female reproductive tract of lower vertebrates', *General and Comparative Endocrinology*, 3, pp. 703–714. doi: 10.1016/0016-6480(72)90200-6.
- Hernández, J. *et al.* (2015) 'Interaction of neuropeptidase activities in cortico-limbic regions after acute restraint stress', *Behavioural Brain Research*. Elsevier, 287, pp. 42–48. doi: 10.1016/J.BBR.2015.03.036.
- Ito, N. *et al.* (2004) 'ADAMs, a disintegrin and metalloproteinases, mediate shedding of oxytocinase', *Biochemical and Biophysical Research Communications*. Academic Press, 314(4), pp. 1008–1013. doi: 10.1016/J.BBRC.2003.12.183.
- Ivell, R. and Walther, N. (1999) 'The role of sex steroids in the oxytocin hormone system', *Molecular and Cellular Endocrinology*, 151(1), pp. 95–101. doi: 10.1016/S0303-7207(99)00025-8.
- Johnson, Z. V. and Young, L. J. (2017) 'Oxytocin and vasopressin neural networks: Implications for social behavioral diversity and translational neuroscience', *Neuroscience and Biobehavioral Reviews*. Elsevier Ltd, 76, pp. 87–98. doi: 10.1016/j.neubiorev.2017.01.034.
- Jones, D. T., Taylor, W. R. and Thornton, J. M. (1992) 'The rapid generation of mutation data matrices from protein sequences', *Computer Applications in the Biosciences*, 8.
- Jongeneel, C. V., Bouvier, J. and Bairoch, A. (1989) 'A unique signature identifies a family of zinc-dependent metalloproteinases', *FEBS Letters*. Wiley-Blackwell, 242(2), pp. 211–214. doi: 10.1016/0014-5793(89)80471-5.
- Katoh, K., Rozewicki, J. and Yamada, K. D. (2017) 'MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization', *Briefings in Bioinformatics*. doi: 10.1093/bib/bbx108.

- Keller, S. R. *et al.* (1995) 'Cloning and characterization of a novel insulin-regulated membrane aminopeptidase from Glut4 vesicles.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 270(40), pp. 23612–8. doi: 10.1074/JBC.270.40.23612.
- Kumar S, Stecher G, Suleski M, H. S. (2017) 'TimeTree: A Resource for Timelines, Timetrees, and Divergence Times.', *Mol Biol Evol*, 34 (7), pp. 1812–1819.
- Kumar, S. *et al.* (2018) 'MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms', *Molecular Biology and Evolution*, 35.
- Larsson, A. (2014) 'AliView: a fast and lightweight alignment viewer and editor for large data sets.' *Bioinformatics*30(22): 3276-3278.
- Lee, A. G. *et al.* (2011) 'A novel form of oxytocin in New World monkeys.', *Biology letters*, 7(4), pp. 584–587. doi: 10.1098/rsbl.2011.0107.
- Lee, H.-J. *et al.* (2010) 'Oxytocin : the Great Facilitator of Life', *Prog. Neurobiol*, 88(2), pp. 127–151. doi: 10.1016/j.pneurobio.2009.04.001.Oxytocin.
- Ludwig, M. and Leng, G. (2006) 'Dendritic peptide release and peptide-dependent behaviours', *Nature Reviews Neuroscience*, 7(2), pp. 126–136. doi: 10.1038/nrn1845.
- Martin, R. D. (2008) 'Evolution of Placentation in Primates: Implications of Mammalian Phylogeny', *Evolutionary Biology*. Springer US, 35(2), pp. 125–145. doi: 10.1007/s11692-008-9016-9.
- Matsumoto, H. *et al.* (2001) 'Expression of placental leucine aminopeptidase/oxytocinase in neuronal cells and its action on neuronal peptides', *European Journal of Biochemistry*, 268(11), pp. 3259–3266. doi: 10.1046/j.1432-1327.2001.02221.x.
- Mittermeier, R.A., Rylands, A.B. & Wilson, D. E. (2013) *Handbook of the Mammals of the World. Vol. 3. Primates*. Barcelona: Lynx Edicions.
- Murphy, W. J. *et al.* (2001) 'Resolution of the early placental mammal radiation using Bayesian phylogenetics.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 294(5550), pp. 2348–51. doi: 10.1126/science.1067179.
- Nakada, T. *et al.* (2011) 'Leucyl/Cystinyl Aminopeptidase Gene Variants in Septic Shock', *CHEST*, 139, pp. 1042–1049. doi: 10.1378/chest.10-2517.
- Nomura, S. *et al.* (2005) 'Gene regulation and physiological function of placental leucine aminopeptidase/oxytocinase during pregnancy', *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. Elsevier, 1751(1), pp. 19–25. doi: 10.1016/J.BBAPAP.2005.04.006.
- Olazábal, D. E. (2014) 'Comparative analysis of oxytocin receptor density in the nucleus accumbens: An adaptation for female and male alloparental care?', *Journal of Physiology-Paris*. Elsevier Ltd, 108(2–3), pp. 213–220. doi: 10.1016/j.jphysparis.2014.10.002.

- Parreiras-E-Silva, L. T. *et al.* (2017) 'Functional New World monkey oxytocin forms elicit an altered signaling profile and promotes parental care in rats.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 114(34), pp. 9044–9049. doi: 10.1073/pnas.1711687114.
- Pham, V. *et al.* (2009) 'Reproduction and maternal behavior in insulin-regulated aminopeptidase (IRAP) knockout mice', *Peptides*, 30(10), pp. 1861–1865. doi: 10.1016/j.peptides.2009.07.019.
- Pond, C. M. (1977) 'The Significance of Lactation in the Evolution of Mammals', *Evolution*, 31(1), pp. 177–199. doi: 10.2307/2407556.
- Pozzi, L. *et al.* (2014) 'Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes.', *Molecular phylogenetics and evolution*, 75, pp. 165–83. doi: 10.1016/j.ympev.2014.02.023.
- Puglia, M. H. *et al.* (2015) 'Epigenetic modification of the oxytocin receptor gene influences the perception of anger and fear in the human brain', *Proceedings of the National Academy of Sciences*, 112(11), pp. 3308–3313. doi: 10.1073/pnas.1422096112.
- Rawlings, N. D. *et al.* (2018) 'The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database', *Nucleic Acids Research*. Oxford University Press, 46(D1), pp. D624–D632. doi: 10.1093/nar/gkx1134.
- Rawlings, N. D. and Salvesen, G. (2013) *Handbook of proteolytic enzymes*. Elsevier. doi: <https://doi.org/10.1016/C2009-1-60990-4>.
- Reiss, K. and Saftig, P. (2009) 'The "A Disintegrin And Metalloprotease" (ADAM) family of sheddases: Physiological and cellular functions', *Seminars in Cell & Developmental Biology*. Academic Press, 20(2), pp. 126–137. doi: 10.1016/J.SEMCDB.2008.11.002.
- Rogi, T. *et al.* (1996) 'Human placental leucine aminopeptidase/oxytocinase: A new member of type II membrane-spanning zinc metallopeptidase family', 271(1), pp. 56–61.
- Saito, A. (2015) 'The marmoset as a model for the study of primate parental behavior.', *Neuroscience Research*. Elsevier Ireland Ltd and Japan Neuroscience Society, 93, pp. 99–109. doi: 10.1016/j.neures.2014.12.011.
- Schneider, H. and Sampaio, I. (2013) 'The systematics and evolution of New World primates - A review', *Molecular Phylogenetics and Evolution*. Elsevier Inc., 82, pp. 348–357. doi: 10.1016/j.ympev.2013.10.017.
- Song, Z. *et al.* (2014) 'Oxytocin induces social communication by activating arginine-vasopressin V1a receptors and not oxytocin receptors', *Psychoneuroendocrinology*. Elsevier Ltd, 50, pp. 14–19. doi: 10.1016/j.psyneuen.2014.08.005.
- Tian, C., Huang, Z. and Wen, Z. (2016) 'Associations between serum placental leucine aminopeptidase and pregnancy outcomes', *International Journal of Gynecology & Obstetrics*. No longer published by Elsevier, 135(3), pp. 255–258. doi: 10.1016/J.IJGO.2016.05.016.

- Tipton, K. (2018) 'Enzyme Nomenclature News: Translocases (EC 7): A new EC Class', *ExplorEnz: the primary source of the IUBMB enzyme list*.
- Tsujimoto, M. et al. (1992) Identification of Human Placental Leucine Aminopeptidase as Oxytocinase, *ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS*.
- Tsujimoto, M. and Hattori, A. (2005) 'The oxytocinase subfamily of M1 aminopeptidases', in *Biochimica et Biophysica Acta - Proteins and Proteomics*. Elsevier, pp. 9–18. doi: 10.1016/j.bbapap.2004.09.011.
- Tuppy, H. (1953) 'The amino-acid sequence in oxytocin', *Biochimica et Biophysica Acta*. Elsevier, 11, pp. 449–450. doi: 10.1016/0006-3002(53)90071-7.
- UniProt Consortium, T. (2018) 'UniProt: the universal protein knowledgebase', *Nucleic Acids Research*. Oxford University Press, 46(5), pp. 2699–2699. doi: 10.1093/nar/gky092.
- Vargas-Pinilla, P. et al. (2015) 'Evolutionary pattern in the OXT-OXTR system in primates: Coevolution and positive selection footprints: Fig. 1.', *Proceedings of the National Academy of Sciences*, 112(1), pp. 88–93. doi: 10.1073/pnas.1419399112.
- Voet, D., Voet, J. and Pratt, C. W. (2016) *Fundamentals of Biochemistry: Life at the Molecular Level*. 5th edn. John Wiley & Sons, Inc.
- Wallis, M. (2012) 'Molecular evolution of the neurohypophysial hormone precursors in mammals: Comparative genomics reveals novel mammalian oxytocin and vasopressin analogues.', *General and comparative endocrinology*. Elsevier Inc., 179(2), pp. 313–8. doi: 10.1016/j.ygcen.2012.07.030.
- Watson, J. D. (2014) *Molecular Biology of the Gene*. 7th edn. Benjamin-Cummings Publishing Company, 2014.
- Weadick, C. J. and Chang, B. S. W. (2012) 'An Improved Likelihood Ratio Test for Detecting Site-Specific Functional Divergence among Clades of Protein-Coding Genes', *Molecular Biology and Evolution*, 29(5), pp. 1297–1300. doi: 10.1093/molbev/msr311.
- Wildman, D. E. et al. (2006) 'Evolution of the mammalian placenta revealed by phylogenetic analysis.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 103(9), pp. 3203–8. doi: 10.1073/pnas.0511344103.
- Worley, K. C. et al. (2014) 'The common marmoset genome provides insight into primate biology and evolution', *Nature Genetics*, 46(8), pp. 850–857. doi: 10.1038/ng.3042.
- Yamashita, K. and Kitano, T. (2013) 'Molecular evolution of the oxytocin-oxytocin receptor system in eutherians', *Molecular Phylogenetics and Evolution*, 67(2), pp. 520–528. doi: 10.1016/j.ympev.2013.02.017.
- Yang Z, Wong WS, N. R. (2005) 'Bayes empirical Bayes inference of amino acid sites under positive selection', *Molecular Biology and Evolution*, 22, pp. 1107–1118.

Yang, Z. (2007) 'Phylogenetic Analysis by Maximum Likelihood (PAML)', *Mol. Biol. Evol.* 24, pp. 1586-1591.

Ye, S. *et al.* (2007) 'Insulin-regulated aminopeptidase: analysis of peptide substrate and inhibitor binding to the catalytic domain', *Biol. Chem.*, 388, pp. 399–403. doi: 10.1515/BC.2007.044.

Yeatman, H. R. *et al.* (2008) 'Identification and characterization of a new cognitive enhancer based on inhibition of insulin-regulated aminopeptidase', *The FASEB Journal*, 22(12), pp. 4209–4217. doi: 10.1096/fj.08-112227.