Elytania Veiga Menezes

**Seleção positiva sobre o gene do hormônio do crescimento e sua
associação com a diversificação de tamanho nos Platyrrhini**

São Paulo
2008

Elytania Veiga Menezes

**Seleção positiva sobre o gene do hormônio do crescimento e sua
associação com a diversificação de tamanho nos Platyrrhini**

Tese apresentada ao Instituto de Biociências
da Universidade de São Paulo, para a obtenção
do título de Doutor em Ciências, na área de
Biologia/ Genética

Orientador (a): Prof. Dr. Gabriel H. Marroig
Zambonato

São Paulo
2008

Ficha Catalográfica

Menezes, Elytania Veiga

Seleção positiva sobre o gene do hormônio do crescimento e sua associação com a diversificação de tamanho nos Platyrrhini

Número de páginas:

Tese (Doutorado)-Instituto de Biociências da Universidade de São Paulo.

Departamento de Biologia/Genética

1. Palavras-Chave- Seleção Positiva, homônio do crescimento, Primatas Neotropicais I. Universidade de São Paulo. Instituto de Biociências. Departamento de Biologia/Genética.

Comissão Julgadora:

Prof (a).Dr(a).

Prof (a).Dr(a)

Prof (a).Dr(a).

Prof (a).Dr(a)

Prof. Dr. Gabriel Marroig Zambonato
Orientador

Dedicatória

Aos meus pais

“Uma vida não basta ser apenas vivida: também precisa ser sonhada”

Mário Quintana

Agradecimentos

À Fundação de Amparo à Pesquisa do Estado de São Paulo, pelo auxílio financeiro.

Aos professores e funcionários do Instituto de Biociências da Universidade de São Paulo por estarem dispostos a ajudar sempre que precisei.

Ao Prof. Dr. Gabriel Marroig pelo aceite em me orientar, por toda a paciência e ensinamentos durante estes 4 anos. Muito obrigada Gabriel por sempre estar de portas abertas quando eu precisei.

Ao Prof. Dr. Diogo Mayer pela ajuda imprescindível na análise dos dados.

Aos amigos do Laboratório de Evolução de Mamíferos, Val, Leila, Paula, Fino, Harley, Karina, Artur, Sebastien e Roberta, pela companhia, pela alegria, pelo incentivo.

Ao Prof. Dr. Sérgio Matioli pela ajuda no laboratório.

À Elzi Santos por ser minha companhia constante durante estes anos e ocupar o posto de Mãe sempre que precisei.

Aos amigos do Laboratório de Mosca das Frutas, Érica, Keiko, Severino pelo carinho e companhia.

À amiga Carla Santos, sempre presente mesmo morando longe.

Aos meus pais Eliezer e Catarina, as pessoas mais importantes da minha vida, que mesmo de tão longe sempre foram muito presentes, obrigada por apoiarem minhas decisões mesmo com o coração apertado de saudade.

À todos os amigos que fiz durante estes 4 anos, em especial a Solange por todos os momentos muito bem vividos, por fazer parte da minha vida como se fosse uma irmã. A Vida sem amigos é como o céu sem estrelas, pois são eles que nos iluminam nos dias mais escuros.... Amo vcs!!!!!!

Índice

Introdução Geral	1
Material e Métodos	17
Resultados	32
Discussão	79
Conclusões e Considerações Finais	92
Resumo	94
Abstract	95
Referências Bibliográficas	96
Anexos	98

Lista de Tabelas

Tabela 1	31
Tabela 2	35
Tabela 3	40
Tabela 4	45
Tabela 5	50
Tabela 6	51
Tabela 7	52
Tabela 8	53
Tabela 9	54
Tabela 10	56
Tabela 11	57
Tabela 12	58
Tabela 13	60
Tabela 14	61
Tabela 15	61
Tabela 16	62
Tabela 17	64
Tabela 18	64
Tabela 19	65
Tabela 20	69
Tabela 21	70

Lista de Figuras

Figura 01	8
Figura 02	12
Figura 03	33
Figura 04	67
Figura 05	71
Figura 06	72
Figura 07	73
Figura 08	75
Figura 09	82

Introdução Geral

O Hormônio de Crescimento como Gene Candidato a Diversificação de Tamanho nos Plathyrrini

O tamanho do corpo é uma das características mais marcantes da diversificação biológica, sendo importante para análises comparativas da história de vida (Calder 1984, Harvey & Clutton-Brock 1985), dimorfismo intraespecífico, alometria e também é um importante componente da teoria da seleção sexual e natural (Bernstein *et al.* 2005). A ampla variação no tamanho do corpo de primatas promove oportunidades para se investigar os mecanismos de desenvolvimento associados com mudanças evolutivas (Goldfrey *et al.* 1990, Delson *et al.* 2000). Tais transformações de tamanho podem ser produzidas por meio de diferentes processos, incluindo alterações na taxa ou tempo de crescimento somático, alterações essas que podem ser provocadas por diferentes fatores fisiológicos e/ou ambientais (Shea 1992).

O desenvolvimento craniofacial e sua evolução tem sido um assunto intrigante na morfologia dos vertebrados, sendo que alguns estudos vêm mostrando que restrições evolutivas podem ser importantes na evolução dos crânios, impossibilitando alguns modos de alimentação que ocorrem em certas linhagens (Collin & Janis 1997)

A evolução dos primatas neotropicais é considerada uma radiação adaptativa resultando hoje em mais de 110 espécies viventes (Rylands *et al.* 2000) que ocupam muitas zonas adaptativas relacionadas a dieta, locomoção, sistema de cruzamento e hábitat. Marroig & Cheverud (2005), baseados em dados de morfologia do crânio, sugerem que a radiação adaptativa nos

primatas neotropicais foi caracterizada pela associação da evolução do tamanho com a diversificação da dieta, mostrando que a evolução de tamanho é um ponto importante para a diversificação destes primatas.

A diversificação craniana dependente de tamanho, conforme observada por Marroig & Cheverud (2005), certamente tem uma base genética. Identificar exatamente que genes foram modificados ao longo da evolução dos primatas neotropicais, de maneira a produzir esse padrão fenotípico, é uma tarefa árdua, pois o tamanho é uma característica influenciada por vários genes (herança poligênica). Uma primeira abordagem para desvendar as bases genéticas desse padrão é o enfoque de genes candidatos. Genes candidatos são genes de função biológica conhecida e que estão envolvidos com a manifestação da característica de interesse, podendo tanto ser estruturais como reguladores de vias bioquímicas que influenciem a característica estudada (Rothschild & Soller 1999). Em relação à evolução de tamanho, um possível gene candidato é o do hormônio de crescimento (GH), cuja regulação da secreção e ação no tecido alvo são um dos determinantes fundamentais do crescimento do corpo.

O GH é regulado pela nutrição e pelo complexo hormonal e genético que controla o tempo e a taxa de crescimento. Os efeitos mais significantes do GH são pós-natais, sendo que há pouco ou nenhum efeito do GH na regulação do crescimento fetal. Muitos dos efeitos do GH no crescimento são mediados indiretamente pelo controle da síntese dos fatores de crescimento, sendo que o IGF-I (*insulin-like growth factor I*) é o mais importante deles, intermediário na ação do GH (Mullis 2005).

O hormônio de crescimento placentário é produto do gene GH5 em humanos, predominantemente expresso nas células do sinciotrofoblasto da placenta. Durante a gestação, a expressão do GH5 substitui gradualmente a

expressão do GH1 que não é detectado no final da gestação (Lacroix et al. 2002). Contudo, o gene GH5 não parece ter efeito direto no crescimento fetal, pois não é detectado na circulação durante essa fase do desenvolvimento (Lacroix et al. 2002). No fígado materno e em outros órgãos, esse hormônio estimula a gliconeogênese e lipólise, aumentando a disponibilidade de nutrientes na unidade fetoplacentar (Lacroix et al. 2002), o que sugere que o GH5 pode ter um efeito indireto no crescimento fetal. Em primatas neotropicais, existe mais um gene que é expresso na placenta, o GH6, mas sua função é desconhecida.

O papel do GH na regulação do tamanho também é evidenciado por ensaios em laboratório. A administração sistemática de GH induz o balanço de nitrogênio positivo e estimula a síntese protéica em músculos, além do crescimento dos ossos. Em humanos, diferentes doenças associadas à falhas no crescimento devido a deficiência de GH são acompanhadas por anormalidades no metabolismo de lipídios, carboidratos e proteínas, e pelo retardamento da puberdade (Forsyth & Wallis 2002).

O hormônio do crescimento (GH) tem diversas ações fisiológicas e entre elas a regulação da expressão tecido-específica de vários genes envolvidos no crescimento, metabolismo e diferenciação. Vários dos efeitos do GH no crescimento somático e expressão de genes associados ao sexo são regulados pelo padrão de secreção do GH na pituitária (também chamada de hipófise), que em mamíferos é sexualmente diferenciada. Podemos citar como exemplo os roedores, nos quais os machos tem pulsos de secreção de GH, enquanto nas fêmeas a secreção é contínua; é exatamente esse padrão o maior determinante nas diferenças sexuais da taxa de crescimento do corpo e da expressão sexualmente dimórfica de numerosos genes no fígado, como os

receptores de prolactina, vários citocromos P450 e uma gama de enzimas (Lupo *et al.* 2001).

Estudar o gene do GH como candidato para explicar o padrão de evolução de tamanho também se justifica pelo fato de que ele influencia muitas outras moléculas associadas ao crescimento, como o IGF-I (insulin-like growth factor). A transcrição do IGF-I é induzida pelo GH em muitos tecidos incluindo o fígado, o qual supre a maioria do IGF-I circulante. Apesar da predominância do IGF-I hepático no plasma, a indução do IGF-I nos condrócitos da placa de crescimento ósseo contribui substancialmente para o crescimento linear (Yakar *et al.* 2002). O IGF-I é usado como substituto clínico do GH na promoção do crescimento linear e é usado clinicamente no tratamento de insensibilidade ao GH causada pela perda de função por mutações no receptor do GH (GHR). Contudo, o GH é mais efetivo que o IGF-I, como exemplificado pela comparação do padrão de crescimento de ratos realizada por Lupo *et al.* 2001, onde se nota que os roedores *igf-i*^{-/-} são consideravelmente maiores quando comparados com roedores *ghr*^{-/-}/*igf-i*^{-/-}, porque o GH exerce promoção de crescimento adicional, ação essa dependente do IGF-I (Lupo *et al.* 2001).

O início dos estudos *in vitro* identificaram a STAT5a/b como o fator de transcrição chave na sinalização do GH. Udy *et al.* (1997) observou que camundongos machos *stat5b*^{-/-} tinham o crescimento prejudicado e os níveis de IGF-I no soro diminuídos e Teglung *et al.* (1998) observou que machos e fêmeas duplamente deletados (*stat5a/b*^{-/-}) tinham o crescimento prejudicado mais severamente, mas não tão severamente quanto os ratos *ghr*^{-/-}. Os ratos *stat5b*^{-/-} e *stat5a*^{-/-}/*5b*^{-/-} não mostram dimorfismo sexual e crescimento pós natal evidente em ninhadas do tipo selvagem. Além disso ratos *stat5b*^{-/-} não induzem a expressão do IGF-I em resposta a administração de GH (Davey *et*

al. 2001). Inversamente enquanto ratos com deleção Soc-2 (Supressor de sinalização da citocina, importante regulador negativo da ação do GH) tem acentuado crescimento, já ratos Stat5b e Socs-2 deletado não, demonstrando assim que STAT5b está envolvida na ação de crescimento pelo GH (Greenhalgh *et al.* 2002).

Em 2003, foi mostrado que a expressão de STAT5b constitutivamente ativa aumenta a expressão do transcrito hepático IGF-I. Além disso a STAT5b quando transformada experimentalmente em negativo dominante, aumenta a expressão IGF-I induzida pelo GH (Woelfle *et al.* 2003a). Logo depois, dois sítios de ligação a STAT5 adjacentes no intron 2 do IGF-I foram identificado; estes sítios de ligação estão localizados dentro de uma região conservada que sofrem mudanças extremas e reversíveis na estrutura da cromatina após o tratamento com GH *in vivo* (Woelfle *et al.* 2003b). Pouco depois mutações inativando o domínio *encoding* da STAT5b foi encontrado em pacientes com deficiência severa de crescimento (Kofoed *et al.* 2003). Este paciente apresentava concentração de IGF-I circulante muito baixa e alta de GH, indicando a insensibilidade do GH, embora a expressão do GHR fosse normal. Assim, a STAT5b parece ser um elemento crítico no crescimento pós-natal mediado pelo IGF-I.

Em complemento aos seus efeitos de crescimento pós-natal, o GH é um importante regulador da composição do corpo, particularmente aumentando a massa muscular e diminuindo a porcentagem de gordura corporal. Essa relação é bastante evidente no aumento da obesidade em ratos com deficiência de GH. As ações antiobesidade do GH são acentuadas pela secreção em pulsos, evidente em machos, conforme observado em ratos machos jovens transgênicos que mostram obesidade substancial expressando

níveis baixos ou constantes de GH (Takahashi *et al.* 1999). A esses animais falta a ativação em pulsos da STAT5 que é importante na expressão de genes hepáticos para o dimorfismo sexual. A importância da ativação da STAT5 na regulação da composição do corpo é apoiada pela descoberta de que machos com deleção da STAT5 madura foram associados com notável obesidade (Vidasdorttir *et al.* 2006). Os mutantes GHR também exibem obesidade (mais de 50% de gordura corporal) em função da perda da habilidade de ativação da STAT5.

Em relação à regulação do metabolismo de lipídios pelo GH, vários processos parecem ser afetados com base em dados de expressão de transcritos alterados. Assim, genes para o metabolismo de lipídios (Fabp, Cd36, ApoA4), síntese de ácidos graxos (Cpt1, Ech1, Acot1, Dci) e expressão dos receptores nucleares Ppar γ são todos afetados por alterações nos níveis de sinalização do GH (Rinn & Snyder 2005).

Outra evidência da importância das implicações evolutivas de eventuais mudanças no gene do GH está no seu papel na determinação do dimorfismo sexual. Muitos genes conhecidos que têm expressão sexualmente dimórfica são regulados pelo GH em resposta ao seu padrão de secreção contínuo, exclusivo de fêmeas (Rinn & Snyder 2005). Por exemplo, de 37 genes identificados por microarray como sendo predominantemente expressos em fêmeas, no fígado de ratos, 73% foram induzidos depois de 7 dias por administração contínua de GH em machos adultos, inibindo a expressão de cerca de 90% dos 49 genes predominantemente expressos em machos (AhluWalia *et al.* 2004). Inversamente, alguns genes macho-específicos, são induzidos por administração em pulsos do GH. Em alguns casos, a diferença na maneira como o GH é administrado gera diferenças na expressão do

transcrito da ordem de 500 ou mais vezes entre machos e fêmeas (Waxman & O'Connor 2006). Como muitos destes genes são membros da grande família do citocromo P450 monooxigenases, que metabolizam esteróides e ácidos graxos, essas diferenças têm significado fisiológico; tais diferenças são também aparentes em humanos, embora em menor grau.

Existe uma forte correlação entre a secreção em pulsos do GH e os pulsos que ativam a fosfo-STAT5b ambos temporalmente e em magnitude, tal que, em machos, existem grandes pulsos de fosfo-STAT5b resultando de cada pulso secretor, enquanto em fêmeas, a ativação STAT5b é baixa e constante (Tannenbaum *et al.* 2001).

O GH é uma molécula alongada formada por quatro hélices firmemente empacotadas de maneira paralela, sendo que duas delas (1 e 4) contêm aproximadamente 30 resíduos de aminoácidos e as outras duas (2 e 3) cerca de 20 resíduos. Assim, 54% dos resíduos de aminoácidos da proteína estão contidos nas quatro hélices. Um longo cruzamento de conexão, formado pelos resíduos 35 ao 71, liga as hélices 1 e 2 e uma conexão similar é encontrada entre as hélices 3 e 4 nos resíduos 129 a 154. Quatro resíduos de cisteína (Cys) nas posições 53, 164, 181 e 189 formam duas pontes dissulfeto entre os resíduos 53 e 164 e 181 e 189, sendo que os resíduos Cys 164 e Cys 181 estão dentro da porção carboxi-terminal da α -hélice (De Vos *et al.* 1992).

O conjunto GH humano consiste de cinco genes altamente conservados aparentemente resultado de uma série de eventos de duplicação. Geralmente são expressos em vários tecidos e sua expressão é controlada por uma região anterior LCR (Locus Control Region) (Barsh *et al.* 1983, Chen *et al.* 1989). Em primatas não humanos o conjunto é formado por oito genes, representados pelo hormônio do crescimento pituitário (GH1) e seus variantes placentários

GH5 e GH6 e mais cinco pseudogenes ver figura 1 (Wallis 2002). Durante a evolução dos primatas vários eventos de duplicação parecem ter levado ao desenvolvimento de um complexo conjunto de genes e a evolução divergente parece ter levado ao surgimento de membros expressos diferentemente na pituitária e placenta, havendo modificações na função fisiológica (Harvey *et al.* 2000). A organização geral do conjunto GH humano difere consideravelmente da dos Primatas do Novo Mundo (Wallis & Wallis 2002). A análise filogenética dos oito genes GH encontrados na espécie *Callithrix jaccus* sugere que eles são mais semelhantes entre eles do que com os cinco encontrados em humanos, sugerindo que o conjunto GH em primatas do Novo Mundo e grandes macacos tiveram origem independente após sua separação (Wallis & Wallis 2002).

Figura1. Comparação entre o conjunto de genes GH em primatas neotropicais e homem.

Callithrix jaccus

Homo sapiens

Fonte: Wallis & Wallis 2001

O gene GH-1 é fortemente conservado em muitos mamíferos, incluindo primatas prosímios, mas nos primatas antropóides parece ter ocorrido episódios de evolução adaptativa, seguida por dois eventos independentes de duplicação gênica e desenvolvimento de um grande conjunto de genes (Wallis & Wallis, 2006). Sua divergência depois da duplicação implicou num complexo e dinâmico processo evolutivo (Chen *et al.* 1989, Wallis & Wallis 2002, Revol De Mendoza *et al.* 2004, Ye *et al.* 2005).

O GH mostra especificidade biológica paralela marcada pela diferença de aminoácidos. Isso não reflete simplesmente a alta taxa de evolução, mas sim um padrão episódico de evolução molecular em que prolongados períodos nos quais as sequências são altamente conservadas são interrompidas por ocasionais mudanças rápidas (Wallis 1996, Otha 1993). Para a evolução do GH em mamíferos os períodos de estase predominaram ocupando aproximadamente 90% do tempo evolutivo e muitas mudanças nas sequências de aminoácidos ocorreram durante períodos rápidos (Wallis 1994). Para a evolução do GH em mamíferos, dois períodos de rápidas mudanças foram indentificados: o primeiro durante a evolução dos artiodáctilos e o segundo no início da evolução dos primatas. Evolução episódica desse tipo parece ter ocorrido também em outras proteínas como, por exemplo, a insulina e a prolactina (Zelus *et al.* 2000).

Os períodos de mudanças rápidas do GH na linhagem que levou aos primatas antropóides foi particularmente marcado pela substituição de 35% de todos os resíduos de aminoácidos. Estudos anteriores estabeleceram que esses episódios ocorreram antes da separação das linhagens que levaram ao

homem e aos primatas do velho mundo, e depois da divergência de muitas ordens de mamíferos placentários (Wallis 1994).

A ação do GH em seus tecidos-alvo depende fundamentalmente do seu encaixe no receptor do hormônio do crescimento (GHR). O GHR pertence à família dos receptores de citocina, que são polipeptídeos trans-membrana com domínios de ligação extracelular e com atividade catalítica (tirosina quinase) em seus domínios citoplasmáticos. Esses receptores podem ser divididos em quatro diferentes subtipos de acordo com as características estruturais comuns. O GHR pertence a primeira classe de receptores de citocinas, também chamado de família de receptores de citocinas hematopoéticas (Bazan 1990). Esses receptores são caracterizados por quatro resíduos Cys conservados na parte amino-terminal de seu domínio extracelular e a presença de um motivo WSXWS (Trp-Ser-Xaa-Trp-Ser) na parte C-terminal. Estruturalmente, o domínio extracelular é composto de dois módulos de fibronectina tipo III, cada um consistindo de aproximadamente 100 aminoácidos e contendo sete fitas β que são dobradas numa forma anti-paralela formando uma estrutura em forma de barril (Moutoussamy *et al.* 1998). Do ponto de vista da localização no genoma, o gene GHR humano está localizado no cromossoma 5p13. 1-p12 e contém dez éxons.

Os GHR maduros estão localizados na superfície celular formando dímeros; a ligação a uma molécula de GH resulta em mudanças conformacionais, levando a estabilização do dímero e resultando na transdução de sinal (Waters 2006). Os GHRs formam homodímeros funcionais (Figura 2), cada subunidade do GHR dimérico associa-se não covalentemente com uma molécula de Tirosina quinase Citosólica II da família Janus (JAK2). Seguindo a

ligação de uma molécula de GH, o GH dimérico submete-se a mudanças conformacionais que induzem a transfosforilação da JAK2 e inicialização da sinalização do GHR. A atividade ligante da JAK2 fosforila múltiplas tirosinas no domínio intracelular do GHR o qual serve de sítios embarcadouros de múltiplas tirosinas de componentes citosólicos das três distintas vias de sinalização: a transdução de sinal e ativação da transcrição (STAT), a proteína quinase de atividade mitogênica (MAPK) e a via fosfoinositídeo-3 quinase (PI3K). Uma região rica em prolina na porção N-terminal do domínio parecem ser necessários para ativação da JAK2 e MAPK, bem como pela fosforilação da STAT1 e STAT3; tirosinas na porção C-terminal no domínio intracelular são essenciais para a ativação da STAT5. Estas cascatas de sinalizações culminam na regulação de múltiplos genes (Zu *et al.* 2001).

Tendo toda essa base teórica em mente, o GH é uma molécula-chave na determinação do tamanho em mamíferos e, portanto, eventuais mudanças ao longo da evolução no gene que a codifica poderiam estar relacionadas aos padrões fenotípicos detectados em primatas do Novo Mundo (Marroig & Cheverud, 2005). Trata-se, assim, de um gene candidato adequado para este estudo.

Figura 2. Ligação do GH ao receptor e ativação de várias proteínas.

Fonte: Agnieszka *et al.* 2007

Detecção de seleção natural em seqüências nucleotídicas

Sob um modelo de evolução neutra, a taxa de substituições nas posições nucleotídicas que possam alterar o aminoácido codificado (chamado de taxa de substituição nucleotídica não-sinônima, dN), deve ser igual à taxa de substituições em posições nucleotídicas que não alteram o aminoácido codificado (taxa de substituição sinônima de nucleotídios, dS). A maioria dos genes codificantes de proteínas, em uma comparação de espécies proximamente relacionadas, apresenta uma razão dN/dS significativamente menor do que 1, ou seja, o número de mutações sinônimas presentes é muito superior ao número de mutações não sinônimas. Este padrão é indicativo de seleção purificadora ou de que o gene está evoluindo de maneira conservativa. A maioria das proteínas foi selecionada, ao longo de milhões de anos de evolução, para determinadas funções. Dessa maneira, é mais provável que qualquer mutação não sinônima que ocorra diminua a adaptação do indivíduo. Consequentemente, a seleção natural estabilizadora atua contra essa mutação, reduzindo a taxa de substituição nesses sítios. No entanto, em alguns casos, podem ocorrer mutações não sinônimas que modificam a função da proteína de maneira a otimizar sua atividade. Dessa forma, nos eventos em que a mutação aumentar a adaptabilidade dos indivíduos e, portanto, sejam selecionadas ($dN/dS > 1$) tem-se a chamada seleção positiva (direcional). Porém, devido ao fato de que a maioria das mudanças de aminoácidos são possivelmente deletérias, o valor dN é normalmente muito menor do que o dS (Fay *et al.* 2001).

A pressão de seleção positiva, em que mudanças particulares dão ao organismo vantagem sobre outros organismos, pode rapidamente alterar as

sequências encontradas em certos sítios do genoma. O estudo da combinação de seleção positiva e deriva genética pode ser utilizado para avaliar o efeito da seleção no genoma e detectar qual nucleotídeo ou códon sofreu qual tipo de seleção em qual ponto da história evolutiva (Liberles 2002).

Segundo Liberles (2002), a maioria dos estudos moleculares de sequências está focalizada na evolução das sequências codificadoras. Isso se deve, obviamente, às várias funções dos polipeptídeos, que vão desde atividades reguladoras estruturais até interações de superfície com solventes e outras atividades catalisadoras. Dessa maneira, em proteínas que estão tendo suas funções modificadas, provavelmente apenas uma parte dos polipeptídeos (dependendo da natureza da modificação) estarão sobre pressão de seleção positiva, enquanto que os restantes não. Dentre os principais mecanismos que tornam essas mutações funcionais destacam-se: mudanças na expressão gênica, modificações no *splicing* do RNA (retirada dos íntrons de um RNA precursor, de forma a produzir um RNAm maduro funcional) e duplicação gênica.

Vários métodos têm sido propostos para detectar a presença de seleção positiva em proteínas homólogas, desde a contagem do número observado de substituições sinônimas e não sinônimas entre pares de seqüências (Li *et al.* 1985, Nei & Gojobori 1986 MacDonald & Kreitman 1991) até a mais sofisticada técnica que preserva a relação filogenética entre muitas sequências e permite um teste estatístico estruturado por estas relações (Nielsen & Yang 1998, Suzuki & Gojobori 1999). Este último método extrai mais informação por considerar as relações evolutivas entre as sequências analisadas.

Detectar seleção positiva geralmente é difícil porque a seleção usualmente atua em poucos sítios e em curtos períodos de tempo evolutivo,

tendo o sinal mascarado pela seleção purificadora. Recentemente foram propostos métodos que testam a seleção positiva num ramo individual ou em um conjunto de ramos da árvore filogenética (Yu and IrWin 1996, Messier & Stuart 1997, Zhang, Kumar & Nei 1997, Yang 1998, Zhang, Rosenberg & Nei 1998) ou em sítios de códons individuais (por ex. Modelo de sítios) (Nielsen & Yang 1998; Sizuki & Gojobori 1999, Yang *et al.* 2000), ou mesmo utilizam uma combinação de sítios individuais em ramos específicos da filogenia (Yang *et al.* 2000).

Apesar das dificuldades em detectar seleção positiva, alguns estudos apresentam evidências claras de que ela ocorre no nível molecular. A lisozima, por exemplo, é uma enzima bacteriolítica que está presente em várias espécies. Os macacos da subfamília Colobinae são os únicos primatas em que existe um estômago modificado, onde as bactérias realizam uma pré-digestão de vegetais, antes que o alimento siga para o estômago verdadeiro, rico em lisozima (Messier & SteWart 1997). Os outros primatas têm um estômago simples com uma anatomia diferente. Em vez de apenas comparar a razão dN/dS do gene da lisozima entre espécies de macacos existentes, Messier & SteWart (1997) também cacularam as sequências ancestrais em vários nós da árvore filogenética e calcularam a taxa dN/dS sobre os ramos da árvore. Eles chegaram a conclusão que não era apenas o ramo que levava aos macacos colobíneos Colobine que estava sobre pressão de seleção positiva durante a evolução dos primatas, mas também, inesperadamente, o ramo que leva aos hominídeos. Os eventos que levaram a essa seleção positiva durante o surgimento do grupo não estão esclarecidos, mas podem estar relacionados com mudanças na dieta (Yang 1998).

Outra proteína que tem sido bastante estudada devido seu envolvimento no controle da obesidade é a leptina. Duas análises da leptina foram feitas em mamíferos (Benner *et al.* 1998 & 2000) e ambas demonstraram episódios de rápida evolução de sequência, tanto em primatas como em roedores. Em relação aos primatas, observou-se ainda a coevolução da leptina e o seu receptor, uma vez que foram observadas altas taxas dN/dS em vários ramos durante a diversificação desse grupo. Essas duas proteínas parecem estar evoluindo novos padrões de expressão durante esse período, padrões esses que são determinados por transposições nas regiões reguladoras de *splicing* e *enhancer* (sequências de DNA que aumentam a afinidade da maquinaria de transcrição por um certo promotor) dos respectivos genes (Libertes 2005).

Outro bom exemplo é o estudo da miostatina, que é a proteína responsável pelo aparecimento de musculatura dupla em mamíferos, especialmente bovinos, em que algumas raças possuem o dobro do número de um tipo de fibra muscular específico em razão de uma mutação no gene dessa proteína. Tellgren *et al.* (2004), estudaram esse gene no grupo monofilético Ruminantia, aplicando uma análise baseada em verossimilhança em que as taxas dN/dS estiveram livres para variar independentemente entre os ramos da filogenia. Tal análise detectou três ramos positivamente selecionados, levando à conclusão de que a pressão de seleção sobre a miostatina teve efeitos fenotípicos marcantes durante a época de divergência de Bovidae e Atilopinae. Uma hipótese é que a pressão de seleção positiva promoveu um aumento na massa corporal e dos músculos esqueléticos associado principalmente a mudanças ambientais nos habitats desses grupos (Tellgren *et al.* 2004).

Marroig e Cheverud (2004), estudando características fenotípicas, mostraram que a evolução do crânio de primatas neotropicais foi

predominantemente adaptativa. Em um trabalho posterior (Marroig & Cheverud 2005), levantaram a hipótese que esta irradiação adaptativa ocorreu principalmente por meio de uma mudança do tamanho geral dos macacos e a associação do tamanho com tipos de dietas diferentes. Baseado nesses achados, o objetivo principal desta tese é identificar processos de evolução molecular correlacionados aos padrões de evolução morfológica já observados em macacos neotropicais. A evolução do tamanho possivelmente está associada à evolução do hormônio do crescimento, cujo gene apresenta uma taxa evolutiva acelerada e foi alvo de seleção positiva em uma série de organismos (Ryynänen et al. 2006, Bernstein et al. 2006, Yuri et al. 2007, Padhi & Verghese 2007). Nesse contexto, procuramos associar aqui a evolução molecular dos genes GH aos dados de evolução morfológica de tamanho para entender os processos de diversificação nos Platyrrhini. Para atingir estes objetivos sequenciamos os genes GH1, GH5 e GH6 nos 16 gêneros de primatas neotropicais e realizamos análises visando detectar seleção positiva sobre esses genes.

Material e Métodos

Amostragem

Amostramos uma espécie de cada um dos 16 gêneros de primatas neotropicais, sendo que para a maioria das amostras o DNA foi cedido pela Prof. Paula Schneider, da Universidade Federal do Pará. Além disso, para obter amostras de alguns gêneros foi necessário realizar uma coleta de tecido sanguíneo no Centro Nacional de Primatas, Ananindeua/Pará. As espécies amostradas são mostradas no anexo 1.

Análises Laboratoriais

Isolamento de DNA

O Isolamento de DNA foi realizado através do método Fenol Clorofórmio como segue:

- ▶ Distribuir 300µl de sangue total ou 100µl de leucócitos em eppendorf de 1,5 ml
- ▶ Adicionar 300µl de tampão de homogeneização e igual volume de tampão de lise
- ▶ Acrescentar 15µl de proteinase K (10mg/ml) e 15µl de RNase (10mg/ml), Agitar 10 minutos e submeter a incubação a 56⁰C por 30 minutos
- ▶ Retirar as amostras do aquecimento
- ▶ Adicionar 600µl de Fenol clorofórmio álcool isoamil na proporção de 25:24:1 e agitar manualmente por 10 minutos
- ▶ Centrifugar por 10 minutos a 13.000 RPM e transferir o sobrenadante para um tubo limpo
- ▶ Repetir a extração com um volume de Clorofórmio álcool isoamil (24:1), agitar manualmente por 10 minutos
- ▶ Centrifugar por 10 minutos a 13.000 RPM e transferir o sobrenadante para um tubo limpo
- ▶ Precipitar este sobrenadante com 100µl de acetato de sódio 3M pH 4.8 e 500µl de álcool isopropílico
- ▶ Agitar gentilmente até a visualização do DNA
- ▶ Centrifugar por 2 minutos a 13.000 RPM para que o pellet precipite no fundo do tubo
- ▶ Descartar o sobrenadante cuidadosamente para não perder o pellet
- ▶ Lavar com 500µl de etanol a 70 %, centrifugando por 5 minutos a 13.000 RPM
- ▶ Descartar o sobrenadante e secar o pellet em estufa a 37 C
- ▶ Adicionar 100 a 300µl de TE para diluir o DNA

Reação em Cadeia da Polimerase (PCR)

As amplificações foram realizadas no termociclador Robocycler gradiente da Stratagene com o Sistema de PCR TripleMaster (eppendorf) nas seguintes condições: 200 nMoles de cada primer, 10ng de DNA, 2,5mM de $MgCl_2$, 200 μ M de dNTP e 0,05U/ μ l de polimerase, perfazendo um volume final de 50 μ l. O programa de ciclagem da reação foi de 35 ciclos com temperatura 94° por 20 s para desnaturação, temperatura de anelamento variando de 60° a 65° por 20s e extensão final 72° por 8min, todas precedidas por temperatura de 94° por 2min para desnaturação inicial. A sequência dos iniciadores externos e internos. encontra-se abaixo:

GENE	Seqüência dos Iniciadores	
	Externos	
	5'	TGGCTATCCTGACATCCTTTCCCGC 3'
GH 1	3'	CCACCCCATAATATTAGAGAAGGACAC 5'
GH5	5'	TGGCTATTCTGACATCCTTTCCCTC 3'
	3'	CACACTCCATAATATTTGAGAAGGACAC 5'
GH6	5'	TGGCTATTCTGACATCCTTTCCCTC 3'
	3'	CCACCCCATAATATTACAGAAGGGCAC 5'
	Internos	
	5'	TGCTGTAGATCTGCCTGAA 3'
GH 1	3'	GCGATGTGAGGGTTTGAG 5'
GH5	5'	TGCTGTGATCCGCGCCTATC 3'
	3'	ACACTGGAGTGCCACCTTC 5'
GH6	5'	GTCCCGAAGGAACAGAAG 3'
	3'	CATAGAAACGGTTGTCGGT 5'

Clonagem dos genes GH:

Após a amplificação, as amostras foram submetidas à purificação utilizando-se o protocolo do sistema GFX PCR DNA & Gel Band Purification Kit (Amersham Biosciences). Durante o processo de purificação, o produto da PCR foi aplicado em um gel de agarose na concentração de 2%, juntamente com um padrão de peso molecular conhecido para que possamos identificar o tamanho do fragmento esperado e assim descartar os produtos de amplificação inespecífica, caso existam.

Uma vez purificadas, as amostras foram clonadas em *E. coli* competentes da cepa DH5 α , segundo o protocolo do sistema de Pgem T Vector (Promega). Em seguida foi realizado o isolamento do plasmídio, utilizando o sistema Wizard Plus Minipreps (Promega).

Sequenciamento

As amostras foram enviadas ao Centro de Estudos do Genoma Humano para o sequenciamento, que seguiu o seguinte protocolo:

Para reação:

DYEnamic ET Terminator reagent premix – 4uL

Primer (5uM) – 1uL

plasmídeo – 1,5uL de plasmídeo (200ng/uL)

H₂O ultrapura – q.s.p. 10uL

Total = 10uL

Programa de Ciclagem

95°C – 20 seg

50°C ** - 15 seg (a mesma utilizada na PCR)|→30 ciclos

60–C - 1 min e 15 seg

A purificação pós-reação foi feita por precipitação com acetato de amônio e etanol, seguindo o protocolo abaixo:

►Adicionamos 1µL de acetato de amônio 7,5M e 28µL de etanol 100% a cada reação, homogeneizando a mistura à temperatura ambiente. A concentração final do etanol foi de 70%. Os tubos foram mantidos à temperatura ambiente por 15min no escuro.

►Centrifugamos os tubos à temperatura ambiente em microcentrífuga por 15 min em rotação máxima (aproximadamente 13.000 rpm),

►Removemos o sobrenadante por aspiração (utilizando micropipeta). Adicionamos 170 µL de Etanol 70% para lavar o pellet. Centrifugamos nas mesmas condições descritas anteriormente.

►Removemos o sobrenadante por aspiração e deixe o pellet secar ao ar,

►Ressuspendemos o pellet em 20uL de Loading Solution.Homogeneizamos e mantivemos a 4°C protegido da luz.

O cálculo do número de substituições sinônimas e não sinônimas foi computada por dois métodos diferentes, explicados a seguir:

O modelo de Nei e Gojobori

Por esse método, dS e dN são estimados computando o número de substituição sinônimas e não sinônimas e o número de sítios potencialmente sinônimos e não sinônimos. Consideraremos primeiro o número de sítios potencialmente sinônimos e não sinônimos, que são computados para cada códon sob a suposição de probabilidades iguais para todas as mudanças nucleotídicas. Consideremos f_i a proporção de mudanças sinônimas e i -th a posição de um códon ($i=1, 2, 3$). O número de sítios potencialmente sinônimos (S) e não sinônimos (N) para este códon é dado por $s = \sum_{i=1}^3 f_i$ e $n=3-s$, respectivamente. Por exemplo, no caso do códon de fenilalanina (TTT), temos:

$$S=0+0+1/3$$

Nenhuma mudança nucleotídica na primeira e segunda posição pode resultar em códons sinônimos e, na terceira posição, apenas uma das três possibilidades de mudança resultaria em um códon sinônimo (TTC). Assim, todas as outras mudanças são não sinônimas, e N é dado por $3-1/3=8/3$. Quando alguma mudança nucleotídica resultar num códon de terminação, esta mudança é desconsiderada. Por exemplo, uma mudança de T para A na terceira posição no códon de cisteína (TGT) resultaria num códon de terminação, mas resultaria em um códon sinônimo quando T muda para C e não sinônimo quando T muda para G (TGG = triptofano). Assim, neste caso $f_3=1/2$ e $f_1 = f_2 = 0$; para este códon, temos $S=0.5$ e $N=2.5$.

Para obter o número de sítios sinônimos (S) e não sinônimos (N) para toda a sequência do gene do GH, nós usamos as fórmulas $S = \sum_{j=1}^C s_j$ e $N = 3C - S$, onde s_j é o valor de s para o códon j -th e C é o número total de códons.

Para calcular o número de diferenças nucleotídicas sinônimas e não sinônimas entre um par de sequências, nós comparamos as duas sequências, códon por códon, e contamos o número de diferenças nucleotídicas para cada par de códons comparados. Quando há somente um nucleotídio diferente, pode-se rapidamente decidir se a mudança é sinônima ou não sinônima; contudo, quando há duas diferenças entre os dois códons comparados, há dois caminhos possíveis para obter a diferença. Comparando, por exemplo, TTT e GTA, há dois possíveis caminhos de mudanças que levem de um códon a outro, que são:

1) TTT (Phe) $\rightarrow$ GTT(Val) $\rightarrow$ GTA (Val)

2) TTT (Phe) $\rightarrow$ TTA (Leu) $\rightarrow$ GTA (Val)

O caminho 1 envolve uma substituição sinônima e uma não sinônima, e o 2 envolve duas substituições não sinônimas. Nós consideramos que os dois caminhos têm igual probabilidade. O número de diferenças sinônimas e não sinônimas seriam $s_d = 0.5$ e $n_d = 1.5$ respectivamente. Em algumas comparações de códons, há caminhos nos quais códons de terminação estão envolvidos, mas esses caminhos são eliminados da computação.

Quando há três diferenças nucleotídicas entre os códons comparados, há seis diferentes possíveis caminhos entre eles e, para cada caminho, três passos mutacionais. Considerando todos esses caminhos e passos, contamos novamente o número de diferenças sinônimas e não sinônimas da mesma

maneira do que com duas diferenças. O número de substituições sinônimas nos caminhos 1, 2 e 5 são 1, 0, 1 e 1 respectivamente e os números de substituições não sinônimas são 2, 3, 2 e 2. Assumimos então que os quatro passos são igualmente prováveis, então temos $S_d = 3/4$ e $N_d = 9/4$. O número total de diferenças sinônimas e não sinônimas para a comparação de sequências pode ser obtido pela soma dos valores de todos os códons. Esse valor é $S_d = \sum_{j=1}^C S_{dj}$ e $N_d = \sum_{j=1}^C N_{dj}$, onde S_{dj} e N_{dj} são o número de diferenças sinônimas e não sinônimas para j -th códon e C é o número de códons comparados.

O modelo de Goldman e Yang

O modelo básico para a análise de máxima verossimilhança é a versão do modelo de substituição de códons de Goldman & Yang (1994), que leva em consideração a estrutura do código genético, a taxa de transição/transversão e a frequência de bases diferentes em cada posição do códon.

Para cada um dos 61 códons, a probabilidade relativa que um códon i mude para j é:

$$Q_{ij} = \begin{cases} 0, & \text{se os dois códons diferem em mais de uma posição,} \\ \pi_j, & \text{para transversões sinônimas,} \\ k\pi_j, & \text{para transições sinônimas,} \\ W\pi_j, & \text{para transversão não sinônima,} \\ Wk\pi_j, & \text{para transições não sinônimas} \end{cases}$$

Onde o parâmetro K é a taxa de transição/transversão, W (dN/dS) é a taxa não sinônimas/sinônimas e π_j é a frequência de equilíbrio do códon j , calculada usando a frequência nucleotídica nas três posições do códon.

Modelo de Sítios

Neste enfoque, vários modelos de substituições de códon são calculados a partir dos dados e suas respectivas *performances* são avaliadas usando o LRTs (*likelihood ratio test*). Evidências estatísticas de seleção positiva são obtidas quando um modelo que considera essa seleção é significativamente mais explicativo do que o modelo sem seleção. O nível de significância é calculado como duas vezes a diferença nos valores de likelihood ($2\Delta L$) para cada modelo, o qual é comparado com a distribuição de χ^2 ; o número de graus de liberdade calculado como a diferença no número de parâmetros estimados entre os modelos nulo e alternativo. Três diferentes LRTs foram calculados. A primeira comparação é entre M_0 (uma taxa), um modelo que calcula um W médio para todos os sítios, e M_3 (discreto), que possui três classes discretas de sítios, cada uma com um diferente W . O segundo é uma comparação entre M_{1a} (modelo próximo a neutralidade), que calcula dois sítios de classe ($0 < W < 1$ ou $W = 1$), e M_{2a} (seleção), que possui um sítio de classe adicional ($W > 1$). O terceiro teste é uma comparação entre um modelo de pressão seletiva de distribuição beta com 10 sítios de classe, cada com $W < 1$ (beta, M_7), com outro modelo com 11 sítios de classe, um dos quais com $W > 1$ (M_8). Esta última comparação é o mais completo teste de seleção positiva (Anisimova *et al.* 2001). O resumo dos modelos e parâmetros encontra-se abaixo:

Modelos	Parâmetros Livres
M0 (one-ratio)	W
M1a (Nearly Neutral): $p_0, (p_1=1-p_0)$ $W_0<1, W_1=1$	$p_0, W_0<1$
M2a (positive selection): $p_0, p_1 (p_2=1-p_0-p_1)$ $W_0<1, W_1=1, W_2>1$	p_0, p_1 $W_0<1, W_2>1$
M3 (discrete): $p_0, p_1 (p_2=1-p_0-p_1)$ W_0, W_1, W_2	p_0, p_1 W_0, W_1, W_2
M7 (beta): p, q	P, q
M8 (beta & W): $p_0, (p_1=1-p_0)$ $P, q, W_s>1$	$P_0, p, q, W_s>1$

W_0, W_1 e W_2 indicam os valores de W dos grupos 0, 1 e 2

p_0, p_1 e p_2 indicam as proporções dos grupos 0, 1 e 2 para cada modelo

Modelo de Ramo e Sítio-Ramo

Modelo de Ramo

Para esta análise, dividimos previamente os ramos da árvore filogenética dos Platyrrhini em ramos de primeiro plano (*foreground*), que são os ramos que *a priori* esperamos estarem sobre seleção positiva, e ramos de segundo plano (*background*), os ramos restantes da árvore. Para determinar que ramos seriam de primeiro plano, utilizamos como critério a quantidade de mudança morfológica dividida pelo comprimento do ramo; essa medida indica a magnitude de mudança morfológica e os ramos com maior magnitude foram considerados de primeiro plano.

Após classificarmos os ramos e marca-los na árvore, passamos então para construção do teste de hipóteses: a hipótese nula fixa dN/dS nos ramos *foreground* =1; a hipótese alternativa presume que $W(dN/dS)$ nos ramos marcados como de primeiro plano diferem do dN/dS médio entre os ramos de segundo plano. Um teste de LRT foi usado para determinar se o modelo alternativo representa significativamente melhor os dados que o modelo nulo. A probabilidade associada foi calculada pela comparação de duas vezes a diferença nos valores de *log likelihood* à distribuição de X^2 , com o grau de liberdade igual à diferença do número de parâmetros entre as duas hipóteses comparadas. Caso o dN/dS dos ramos de primeiro plano seja >1 e o LRT seja significativo (rejeitando a hipótese nula), inferimos então seleção positiva nos ramos de primeiro plano.

Modelo Sítio-Ramo

O modelo sítio-ramo foi desenvolvido para possibilitar o teste de seleção positiva num subconjunto de sítios (códon) num ramo (ou ramos) da árvore. Procedemos aqui como no modelo de ramos, escolhendo previamente os ramos de primeiro plano, mas nesse caso a construção das hipóteses é diferente. Temos uma hipótese alternativa que define quatro classes de sítios, como se segue: esse modelo assume sítios de classe 0, onde estão os sítios sob seleção purificadora em todos os ramos ($0 < W_0 < 1$). Sítios de classe 1, que inclui códon que estão evoluindo de forma neutra na árvore, com $W_1 = 1$ e códon de classe 2a e 2b, que são conservados ou neutros nos ramos de segundo plano, mas estão sob seleção positiva nos ramos de primeiro plano ($W_2 > 1$). A hipótese nula é construída exatamente da mesma forma que a alternativa, com a diferença de que o $W_2 = 1$; ou seja, na hipótese nula não se admite sítios sob seleção positiva nas linhagens de primeiro plano. A significância do teste é calculada da mesma forma que no modelo de ramos.

Teste Z

Uma maneira de testar se a seleção positiva está agindo em um gene é comparar a abundância relativa de substituições sinônimas e não sinônimas dentro das seqüências em questão. Para um par de seqüências, isso é realizado através da estimativa do número de substituições sinônimas por sítios sinônimos (dS) e o número de substituições não sinônimas por sítio não sinônimo (dN), e suas variâncias. Com essa informação, podemos testar a hipótese nula $H_0: dN = dS$, onde

$$Z = (dN - dS) / \sqrt{\text{var}(dS) + \text{var}(dN)}$$

O nível de significância no qual a hipótese nula é rejeitada depende da hipótese alternativa (H1):

H0= dN=dS

H1=

- A) $dN \neq dS$ (teste de neutralidade)
- B) $dN > dS$ (seleção positiva)
- C) $dN < dS$ (seleção purificadora)

Estes três testes podem ser conduzidos diretamente para pares de seqüências. Para essa análise, utilizamos o programa MEGA (Kumar *et al* 2007).

Estimativa do Estado Ancestral

Utilizamos o método de custo quadrático (*Squared cost*), implementado no programa Mesquite 2.01 (Maddison 2007). O programa calcula o estágio ancestral mais parcimonioso nos nós da árvore assumindo um custo $(x-y)^2$ no ramo, caso o ramo contenha uma mudança de X para Y. Essas reconstruções foram feitas para cada mudança de aminoácidos e nucleotídeos, assim como para os dados morfológicos.

Reconstruímos as seqüências do gene GH1 para cada nó ancestral da filogenia dos primatas neotropicais, alinhando as seqüências resultantes. Em seguida, estimamos as taxas de substituições nucleotídicas sinônimas e não sinônimas par a par para cada relação nó ancestral→ramo terminal, utilizando para isso o programa DNAsp.

Análise de Correlação entre os dados morfométricos e moleculares

Marroig & Cheverud (2005), em seu trabalho de análise do tamanho craniano e variação da alometria nos primatas neotropicais, compararam a direção da evolução, dada pelo vetor da média das diferenças para um conjunto de caracteres morfológicos, com a linha de menor resistência evolutiva (eixo com a maior variação genética e fenotípica). Com isso, mapearam a correlação da direção que a evolução segue na linha de menor resistência fenotípica ao longo da filogenia dos primatas neotropicais. Nesse trabalho, foi detectada uma associação de fatores ecológicos e de tamanho com a quantidade e a taxa de mudança evolutiva, num contexto filogenético.

Nós utilizamos algumas variáveis (Tabela 1) do referido trabalho para testar a possível associação (via correlação de Pearson) com os dados de dN/dS, dN e dS obtidos da análise do modelo de ramos. Buscamos também correlacionar essas variáveis com os resultados das análises par a par entre os nós ancestrais e terminais da filogenia. Com isso, testamos se existe associação linear entre as taxas morfológicas e moleculares, utilizando para isso o programa SYSTAT 11.

Tabela 1 Variáveis morfométricas utilizadas nas correlações e seu significado biológico.

Medidas Morfométricas	Significado Biológico
Correlação entre o vetor das diferenças das médias entre dois grupos e o primeiro componente principal da matriz genética	Mede o quão próximo a direção da evolução segue a linha de menor resistência evolutiva (LLER) genética
Correlação entre o vetor das diferenças das médias entre dois grupos e o primeiro componente principal da matriz fenotípica	Mede o quão próximo a direção da evolução segue a LLER fenotípica
Logarítimo natural da soma do quadrado das diferenças na média dos caracteres fenotípicos entre dois grupos	Quantidade de diferenciação morfológica entre dois grupos
Soma do quadrado das diferenças na média dos caracteres fenotípicos entre dois grupos dividido pelo tamanho do ramo	A taxa de mudança morfológica entre dois grupos
Logarítimo natural do tamanho absoluto do crânio	Tamanho absoluto do crânio

RESULTADOS

Amplificação por PCR

Todos os fragmentos dos genes GH1, GH5 e GH6 amplificaram com especificidade e com bom rendimento nas condições utilizadas, sendo que o tamanho das bandas foi verificado pela comparação com o padrão de peso molecular de 2Kb e eram equivalentes aos descritos na literatura.

Sequenciamento

Seqüenciamos uma espécie de cada um dos 16 gêneros de primatas neotropicais, obtendo um fragmento de aproximadamente 1.9 kb, que inclui os cinco éxons. Nós analisamos 654 sítios referentes a região dos éxons, o número de sítios polimórficos foram 138, com um total de 49 mudanças sinônimas e 89 não sinônimas para o conjunto total de seqüências do gene GH1. Para o gene GH5 o número de sítios polimórficos foram 158, com um total de 43 mudanças sinônimas e 107 não sinônimas e para o gene GH6 o número de sítios com mutações foram 163, com um total de 59 mudanças sinônimas e 104 não sinônimas conjunto total de seqüências. A fig 3 (a, b e c) mostra uma parte do alinhamento produzido pelo programa BioEdit 7.0 para alguns gêneros.

Figura 3 Alinhamento dos genes GH1 (a), GH5 (b) e GH6 (c), para cinco gêneros de Platyrrhini alguns gêneros.

(a)

(b)

(c)

```

      210      220      230      240      250      260      270      280
Callithrix  CAGAAATATTTCTTCCTGCGTAACCCCGAGACCTTCGTCTGTTCTCAGAGTCTATCCCAACACCCCTCCACAAGGAGGAAATGCTGG
CEBUELLA    CAGAAATATTTCTTCCTGCGTAACCCCGAGACCTTCGTCTGTTCTCAGAGTCTATCCCAACACCCCTCCACAAGGAGGAAATGCTGG
SAIMIRI      CAGAAATATTTCTTCCTGCGTAACCCCGAGACCTTCGTCTGTTCTCAGAGTCTATCCCAACACCCCTCCACAAGGAGGAAATGCTGG
CALLICEBUS   CAGAAATATTTCTTCCTGCGTAACCCCGAGACCTTCGTCTGTTCTCAGAGTCTATCCCAACACCCCTCCACAAGGAGGAAATGCTGG
AOTUS        CAGAAATATTTCTTCCTGCGCAACCCCAAGACCTTCCTTCTGTTCTCAGAGTCTATCCCGACACCCCTCCAAACAAGGAGGAAATGCTGG
ATELES       CAGAAATATTTCTTCCTGCGCAACCCCAAGACCTTCCTTCTGTTCTCAGAGTCTATCCCGACACCCCTCCAAACAAGGAGGAAATGCTGG
lagothrix    CAGAAATATTTCTTCCTGCGCAACCCCAAGACCTTCCTTCTGTTCTCAGAGTCTATCCCGACACCCCTCCAAACAAGGAGGAAATGCTGG

      310      320      330      340      350      360      370      380
Callithrix  TAGACCTGCTCCACATCTCCCTGCTGCTCATCCAGTCCTGGCTTGAACCCCATGCAAGCCTCGGCAGTATCTTTGCCAACAGCCAACT
CEBUELLA    TAGACCTGCTCCACATCTCCCTGCTGCTCATCCAGTCCTGGCTTGAACCCCATGCAAGCCTCGGCAGTATCTTTGCCAACAGCCAACT
SAIMIRI      TAGACCTGCTCCACATCTCCCTGCTGCTCATCCAGTCCTGGCTTGAACCCCATGCAAGCCTCGGCAGTATCTTTGCCAACAGCCAACT
CALLICEBUS   TAGACCTGCTCCACATCTCCCTGCTGCTCATCCAGTCCTGGCTTGAACCCCATGCAAGCCTCGGCAGTATCTTTGCCAACAGCCAACT
AOTUS        TAGACCTGCTCCACATCTGCTGCTTCTCTCATCCAGTCCTGGCTTGAACCCCATGCAAGCCTCGGCAGTATCTTTGCCAACAGCCAACT
ATELES       TAGACCTGCTCCACATCTGCTGCTTCTCTCATCCAGTCCTGGCTTGAACCCCATGCAAGCCTCGGCAGTATCTTTGCCAACAGCCAACT
lagothrix    TAGACCTGCTCCACATCTGCTGCTTCTCTCATCCAGTCCTGGCTTGAACCCCATGCAAGCCTCGGCAGTATCTTTGCCAACAGCCAACT

```

Análises de Seleção Positiva

Taxa dN/dS pelo método Nei e Gojobori (NG)

O dados das estimativas de dN, dS e dN/dS para pares de gêneros pelo método de Nei e Gojobori estão apresentados nas tabelas 2, 3 e 4. Essas tabelas incluem dados para cada par de gênero o número de substituições sinônimas, não sinônimas, a taxa dN/dS, o número de diferenças sinônimas e o número de diferenças não sinônimas para os genes GH1, GH5 e GH6 respectivamente.

As análises por este método encontraram a taxa dN/dS > 1 para 16 comparações entre gêneros no gene GH1, são elas: Brachyteles-Leontopithecus (1,01), Ateles-Callimico (1,01), Brachyteles-Saguinus (1,03), Callithrix-Ateles (1,06), Cebus-Callimico (1,08), Callithrix-Cebus (1,08), Ateles-Cebuella (1,08), Callimico-Leontopithecus (1,13), Saguinus-Callimico (1,16), Cebuella-Cebus (1,16), Cebus-Saguinus (1,18), Callithrix-Leontopithecus (1,18), Callithrix-Saguinus (1,20), Cebuella-Leontopithecus (1,22), Cebuella-Saguinus (1,25) e Cebus-Leontopithecus (1,30).

As análises por este método não encontraram nenhuma taxa $dN/dS > 1$ para as comparações entre gêneros nos genes GH5 e GH6.

Tabela 2 Comparação par a par entre gêneros de Platyrrhini, mostrando dS, dN, dS/dN, o número de diferenças sinônimas e o número de diferenças não sinônimas para o gene GH1. Os valores em negrito representam a taxa $W > 1$.

Pares de Gêneros	dS	dN	dN/dS	Diferenças Sinônimas	Diferenças não sinônimas.
Callicebus-Saimiri	0,013	0	0	2	0
Chiropotes-Lagothrix	0,0066	0	0	1	0
Chiropotes-Cacajao	0	0	0	0	0
Cebuella-Callimico	0,0066	0	0	1	0
Saguinus-Leontopithecus	0	0,002	0	0	1
Lagothrix-Cacajao	0,0066	0	0	1	0
Callicebus-Alouatta	0,1045	0,0101	0,096651	15	5
Alouatta-Callimico	0,1088	0,0111	0,102022	15,5	5,5
Alouatta-Cebuella	0,1013	0,0111	0,109576	14,5	5,5
Alouatta-Saimiri	0,0897	0,0101	0,112598	13	5
Aotus-Brachyteles	0,0824	0,0101	0,122573	12	5
Alouatta-Aotus	0,097	0,0121	0,124742	14	6
Callicebus-Brachyteles	0,0897	0,0121	0,134894	13	6
Callithrix-Alouatta	0,1086	0,0151	0,139042	15,5	7,5
Brachyteles-Callimico	0,0938	0,0131	0,139659	13,5	6,5
Callithrix-Brachyteles	0,0937	0,0131	0,139808	13,5	6,5
Callicebus-Callimico	0,097	0,0141	0,145361	14	7
Cebuella-Brachyteles	0,0865	0,0131	0,151445	12,5	6,5
Callicebus-Cebuella	0,0896	0,0141	0,157366	13	7
Brachyteles-Saimiri	0,0751	0,0121	0,161119	11	6

Saimiri-Callimico	0,0823	0,0141	0,171324	12	7
Aotus-Callimico	0,0896	0,0162	0,180804	13	8
Callithrix-Aotus	0,0895	0,0162	0,181006	13	8
Cebuella-Saimiri	0,0751	0,0141	0,18775	11	7
Callithrix-Callicebus	0,0969	0,0182	0,187822	14	9
Cebuella-Aotus	0,0823	0,0162	0,196841	12	8
Callithrix-Saimiri	0,0822	0,0182	0,221411	12	9
Callithrix-Callimico	0,0132	0,004	0,30303	2	2
Callicebus-Aotus	0,0065	0,002	0,307692	1	1
Ateles-Saguinus	0,013	0,004	0,307692	2	2
Aotus-Saimiri	0,0065	0,002	0,307692	1	1
Pithecia-Lagothrix	0,027	0,01	0,37037	4	5
Pithecia-Callicebus	0,2964	0,1272	0,42915	37,33	58,67
Pithecia-Aotus	0,2867	0,1248	0,435298	36,33	57,67
Callicebus-Chiropotes	0,2615	0,1168	0,446654	33,83	54,1
Callicebus-Cacajao	0,2615	0,1168	0,446654	33,83	54,17
Chiropotes-Aotus	0,2523	0,1145	0,453825	32,83	53,17
Aotus-Cacajao	0,2523	0,1145	0,453825	32,83	53,17
Alouatta-Brachyteles	0,0132	0,006	0,454545	2	3
Pithecia-Saimiri	0,2771	0,1272	0,45904	35,33	58,67
Ateles-Leontopithecus	0,013	0,006	0,461538	2	3
Callicebus-Lagothrix	0,2523	0,1168	0,462941	32,83	54,17
Pithecia-Alouatta	0,2886	0,1341	0,464657	36,33	61,67
Aotus-Lagothrix	0,2432	0,1145	0,470806	31,83	53,17
Chiropotes-Saimiri	0,2432	0,1168	0,480263	31,83	54,17
Saimiri-Cacajao	0,2432	0,1168	0,480263	31,83	54,17
Alouatta-Chiropotes	0,2539	0,1236	0,486806	32,83	57,17
Alouatta-Cacajao	0,2539	0,1236	0,486806	32,83	57,17
Pithecia-Brachyteles	0,2695	0,1317	0,488683	34,33	60,67
Pithecia-Chiropotes	0,0201	0,01	0,497512	3	5
Pithecia-Cacajao	0,0201	0,01	0,497512	3	5

Saimiri-Lagothrix	0,2342	0,1168	0,498719	30,83	54,17
Alouatta-Lagothrix	0,2448	0,1236	0,504902	31,83	57,17
Chiropotes-Brachyteles	0,2358	0,1213	0,514419	30,83	56,17
Brachyteles-Cacajao	0,2358	0,1213	0,514419	30,83	56,17
BrachytelesLagothrix	0,2268	0,1213	0,534832	29,83	56,17
Pithecia-Callithrix	0,2412	0,1365	0,56592	31,33	62,67
Pithecia-Cebuella	0,2325	0,1365	0,587097	30,33	62,67
Callithrix-Chiropotes	0,2089	0,126	0,603159	27,83	58,17
Callithrix-Cacajao	0,2089	0,126	0,603159	27,83	58,17
Callithrix-Cebuella	0,0066	0,004	0,606061	1	2
Pithecia-Callimico	0,2235	0,1365	0,610738	29,33	62,67
Chiropotes-Cebuella	0,2005	0,126	0,628429	26,83	58,17
Cebuella-Cacajao	0,2005	0,126	0,628429	26,83	58,17
Callithrix-Lagothrix	0,2003	0,126	0,629056	26,83	58,17
Chiropotes-Callimico	0,192	0,126	0,65625	25,83	58,17
Cebuella-Lagothrix	0,192	0,126	0,65625	25,83	58,17
Callimico-Cacajao	0,192	0,126	0,65625	25,83	58,17
Lagothrix-Callimico	0,1836	0,126	0,686275	24,83	58,17
Pithecia-Cebus	0,2173	0,155	0,7133	28,83	70,17
Callicebus-Ateles	0,1736	0,1239	0,71371	24	57
Callicebus-Cebus	0,1534	0,111	0,723598	21,5	51,5
Aotus-Cebus	0,1455	0,1087	0,747079	20,5	50,5
Ateles-Aotus	0,1655	0,1262	0,762538	23	58
Callicebus-Leontopithecus	0,1578	0,1215	0,769962	22	56
Chiropotes-Cebus	0,1863	0,1442	0,77402	25,33	65,67
Cebus-Cacajao	0,1863	0,1442	0,77402	25,33	65,67
Pithecia-Saguinus	0,2224	0,1734	0,779676	29,33	77,67
Pithecia-Ateles	0,222	0,1735	0,781532	29,33	77,67
Callicebus-Saguinus	0,1578	0,1238	0,784537	22	57
Ateles-Saimiri	0,1575	0,1239	0,786667	22	57
Pithecia-Leontopithecus	0,2224	0,176	0,791367	29,33	78,67

Cebus-Saimiri	0,1377	0,111	0,8061	19,5	51,5
Cebus-Lagothrix	0,178	0,1442	0,810112	24,33	65,67
Saguinus-Lagothrix	0,1995	0,1625	0,814536	26,83	73,17
Ateles-Lagothrix	0,1992	0,1626	0,816265	26,83	73,17
Aotus-Leontopithecus	0,1499	0,1238	0,825884	21	57
Lagothrix-Leontopithecus	0,1995	0,165	0,827068	26,83	74,17
Alouatta-Ateles	0,1505	0,126	0,837209	21	58
Aotus-Saguinus	0,1499	0,1262	0,841895	21	58
Chiropotes-Saguinus	0,1911	0,1625	0,85034	25,83	73,17
Saguinus-Cacajao	0,1911	0,1625	0,85034	25,83	73,17
Ateles-Chiropotes	0,1907	0,1626	0,852648	25,83	73,17
Ateles-Cacajao	0,1907	0,1626	0,852648	25,83	73,17
Saimiri-Leontopithecus	0,142	0,1215	0,855634	20	56
Chiropotes-Leontopithecus	0,1911	0,165	0,863422	25,83	74,1
Leontopithecus-Cacajao	0,1911	0,165	0,863422	25,83	74,17
Alouatta-Cebus	0,1308	0,1132	0,865443	18,5	52,5
Saimiri-Saguinus	0,142	0,1238	0,871831	20	57
Alouatta-Leontopithecus	0,135	0,1236	0,915556	19	57
Ateles-Cebus	0,0163	0,0152	0,932515	2,5	7,5
Alouatta-Saguinus	0,135	0,1259	0,932593	19	58
Ateles-Brachyteles	0,1348	0,1284	0,952522	19	59
Brachyteles-Cebus	0,1155	0,1108	0,959307	16,5	51,5
Brachyteles-Leontopithecus	0,1196	0,1212	1,013378	17	56
Ateles-Callimico	0,1193	0,1213	1,016764	17	56
Brachyteles-Saguinus	0,1196	0,1236	1,033445	17	57
Callithrix-Ateles	0,1192	0,1261	1,057886	17	58
Cebus-Callimico	0,1004	0,1085	1,080677	14,5	50,5
Callithrix-Cebus	0,1003	0,1086	1,082752	14,5	50,5
Ateles-Cebuella	0,1118	0,1213	1,084973	16	56
Callimico-Leontopithecus	0,1045	0,1189	1,137799	15	55,00
Saguinus-Callimico	0,1045	0,1213	1,160766	15	56

Cebuella-Cebus	0,0931	0,1085	1,165414	13,5	50,5
Cebus-Saguinus	0,0163	0,0193	1,184049	2,5	9,5
Callithrix-Leontopithecus	0,1043	0,1237	1,186002	15	57
Callithrix-Saguinus	0,1043	0,126	1,208054	15	58
Cebuella-Leontopithecus	0,097	0,1189	1,225773	14	55
Cebuella-Saguinus	0,097	0,1213	1,250515	14	56
Cebus-Leontopithecus	0,0163	0,0213	1,306748	2,5	10,5

Tabela 3. Comparação par a par entre gêneros para o gene GH5, mostrando dS, dN, dS/dN, número de diferenças sinônimas e número de diferenças não sinônimas.

Pares de Gêneros	dS	dN	dN/dS	Diferenças sinônimas	Diferenças não sinônimas
Callithrix - Cebuella	0	0	0	0	0
Callithrix- Callimico	0	0	0	0	0
Callithrix - Leontopithecus	0	0	0	0	0
Callithrix - Saguinus	0	0	0	0	0
Callithrix - Cebus	0	0	0	0	0
Callithrix- Aotus	0	0	0	0	0
Callithrix - Saimiri	0	0	0	0	0
Callicebus - Ateles	0	0	0	0	0
Callicebus - Lagothrix	0	0	0	0	0
Callicebus - Brachyteles	0	0	0	0	0
Callicebus - Chiropotes	0	0	0	0	0
Callicebus - Cacajao	0	0	0	0	0
Callicebus - Pithecia	0	0	0	0	0
Cebuella - Callimico	0	0	0	0	0
Cebuella - Leontopithecus	0	0	0	0	0
Cebuella - Saguinus	0	0	0	0	0
Cebuella - Cebus	0	0	0	0	0
Cebuella - Aotus	0	0	0	0	0
Cebuella - Saimiri	0	0	0	0	0
Callimico - Leontopithecus	0	0	0	0	0
Callimico - Saguinus	0	0	0	0	0
Callimico - Cebus	0	0	0	0	0
Callimico - Aotus	0	0	0	0	0
Callimico - Saimiri	0	0	0	0	0

Leontopithecus- Saguinus	0	0	0	0	0
Leontopithecus -Cebus	0	0	0	0	0
Leontopithecus - Aotus	0	0	0	0	0
Leontopithecus - Saimiri	0	0	0	0	0
Saguinus - Cebus	0	0	0	0	0
Saguinus - Aotus	0	0	0	0	0
Saguinus - Saimiri	0	0	0	0	0
Cebus - Aotus	0	0	0	0	0
Cebus - Saimiri	0	0	0	0	0
Aotus - Saimiri	0	0	0	0	0
Ateles - Lagothrix	0	0	0	0	0
Ateles - Brachyteles	0	0	0	0	0
Ateles - Chiropotes	0	0	0	0	0
Ateles - Cacajao	0	0	0	0	0
Ateles - Pithecia	0	0	0	0	0
Lagothrix - Brachyteles	0	0	0	0	0
Lagothrix - Chiropotes	0	0	0	0	0
Lagothrix - Cacajao	0	0	0	0	0
Lagothrix - Pithecia	0	0	0	0	0
Brachyteles - Chiropotes	0	0	0	0	0
Brachyteles - Cacajao	0	0	0	0	0
Brachyteles - Pithecia	0	0	0	0	0
Chiropotes -Cacajao	0	0	0	0	0
Chiropotes - Pithecia	0	0	0	0	0
Cacajao -Pithecia	0	0	0	0	0
Cebuella - Lagothrix	0,268	0,21	0,783582	34,4	93
Callicebus - Cebus	0,2685	0,216	0,804469	34,42	93,5
Cebuella - Pithecia	0,2685	0,216	0,804469	34,42	93,5
Callimico - Ateles	0,2685	0,216	0,804469	34,42	93,5
Callimico - Brachyteles	0,2685	0,216	0,804469	34,42	93,5
Callimico - Chiropotes	0,2685	0,216	0,804469	34,42	93,5

Leontopithecus - Lagothrix	0,2685	0,216	0,804469	34,42	93,5
Saguinus - Ateles	0,2685	0,216	0,804469	34,42	93,5
Saguinus - Brachyteles	0,2685	0,216	0,804469	34,42	93,5
Saguinus - Chiropotes	0,2685	0,216	0,804469	34,42	93,5
Aotus - Lagothrix	0,2685	0,216	0,804469	34,42	93,5
Callithrix - Callicebus	0,2685	0,2161	0,804842	34,42	93,5
Callithrix - Ateles	0,2685	0,2161	0,804842	34,42	93,58
Callithrix - Lagothrix	0,2685	0,2161	0,804842	34,42	3,58
Callithrix- Brachyteles	0,2685	0,2161	0,804842	34,42	3,58
Callithrix - Chiropotes	0,2685	0,2161	0,804842	34,42	3,58
Callithrix - Cacajao	0,2685	0,2161	0,804842	34,42	3,58
Callithrix - Pithecia	0,2685	0,2161	0,804842	34,42	3,58
Callicebus- Leontopithecus	0,2685	0,2161	0,804842	34,42	93,5
Callicebus - Aotus	0,2685	0,2161	0,804842	34,42	93,58
Callicebus - Saimiri	0,2685	0,2161	0,804842	34,42	93,5
Cebuella - Ateles	0,2685	0,2161	0,804842	34,42	93,58
Cebuella - Brachyteles	0,2685	0,2161	0,804842	34,42	93,5
Cebuella - Chiropotes	0,2685	0,2161	0,804842	34,42	93,5
Cebuella - Cacajao	0,2685	0,2161	0,804842	34,42	93,58
Callimico - Cacajao	0,2685	0,2161	0,804842	34,42	93,5
Leontopithecus - Ateles	0,2685	0,2161	0,804842	34,42	3,58
Leontopithecus- Brachyteles	0,2685	0,2161	0,804842	34,42	3,58
Leontopithecus- Chiropotes	0,2685	0,2161	0,804842	34,42	3,58
Leontopithecus - Cacajao	0,2685	0,2161	0,804842	34,42	3,58
Leontopithecus - Pithecia	0,2685	0,2161	0,804842	34,42	3,58
Saguinus - Cacajao	0,2685	0,2161	0,804842	34,42	93,5
Cebus - Ateles	0,2685	0,2161	0,804842	34,42	3,58
Cebus - Brachyteles	0,2685	0,2161	0,804842	34,42	3,58
Cebus Chiropotes	0,2685	0,2161	0,804842	34,42	3,58
Cebus Cacajao	0,2685	0,2161	0,804842	34,42	3,58
Cebus Pithecia	0,2685	0,2161	0,804842	34,42	3,58

Aotus	Ateles	0,2685	0,2161	0,804842	34,42	3,58
Aotus	Brachyteles	0,2685	0,2161	0,804842	34,42	3,58
Aotus	Chiropotes	0,2685	0,2161	0,804842	34,42	3,58
Aotus	Cacajao	0,2685	0,2161	0,804842	34,42	3,58
Aotus	Pithecia	0,2685	0,2161	0,804842	34,42	3,58
Saimiri	Ateles	0,2685	0,2161	0,804842	34,42	3,58
Saimiri	Brachyteles	0,2685	0,2161	0,804842	34,42	3,58
Saimiri	Chiropotes	0,2685	0,2161	0,804842	34,42	3,58
Saimiri	Cacajao	0,2685	0,2161	0,804842	34,42	3,58
Saimiri	Pithecia	0,2685	0,2161	0,804842	34,42	3,58
Callicebus	Cebuella	0,268	0,216	0,80597	34,4	93,5
Callicebus	Callimico	0,268	0,216	0,80597	34,4	93
Callicebus	Saguinus	0,268	0,216	0,80597	34,4	93
Callimico	Pithecia	0,268	0,216	0,80597	34,4	93,5
Saguinus	Pithecia	0,2685	0,2161	0,80597	34,4	93,5
Cebus	Lagothrix	0,268	0,216	0,80597	34,4	93,5
Saimiri	Lagothrix	0,268	0,216	0,80597	34,4	93,5
Callimico	Lagothrix	0,26	0,21	0,807692	34	93
Saguinus	Lagothrix	0,26	0,21	0,807692	34	93
Alouatta	Cebuella	0,125	0,101	0,808	18	47
Alouatta	Callimico	0,125	0,101	0,808	18	47
Alouatta	Saguinus	0,125	0,101	0,808	18	47
Callithrix	Alouatta	0,1251	0,1016	0,81215	18	7
Alouatta	Leontopithecus	0,1251	0,1016	0,81215	18	47
Alouatta	Cebus	0,1251	0,1016	0,81215	18	47
Alouatta	Aotus	0,1251	0,1016	0,81215	18	47
Alouatta	Saimiri	0,1251	0,1016	0,81215	18	47
Alouatta	Ateles	0,2735	0,2252	0,8234	35,33	6,67
Alouatta	Brachyteles	0,2735	0,2252	0,8234	35,33	96,67
Alouatta	Chiropotes	0,2735	0,2252	0,8234	35,33	96,67
Alouatta	Cacajao	0,2735	0,2252	0,8234	35,33	6,67

Alouatta	Pithecia	0,2735	0,2252	0,8234	35,33	96,6
Callicebus	Alouatta	0,273	0,225	0,824176	35,33	96,6
Alouatta	Lagothrix	0,273	0,225	0,824176	35,3	96,6

Tabela 4. Comparação par a par entre gêneros, mostrando dS, dN, dS/dN, o número de diferenças sinônimas e o número de diferenças não sinônimas para o gene GH6.

Pares de Gêneros		dS	dN	dN/dS	Diferenças sinônimas	Diferenças não sinônimas
Callithrix	Cebuella	0	0	0,000	0.00	0.00
Callithrix	Saimiri	0	0	0,000	0.00	0.00
Callithrix	Callicebus	0	0	0,000	0.00	0.00
Callithrix	Leontopithecus	0	0	0,000	0.00	0.00
Cebuella	Saimiri	0	0	0,000	0.00	0.00
Cebuella	Callicebus	0	0	0,000	0.00	0.00
Cebuella	Leontopithecus	0	0	0,000	0.00	0.00
Saimiri	Callicebus	0	0	0,000	0.00	0.00
Saimiri	Leontopithecus	0	0	0,000	0.00	0.00
Callicebus	Leontopithecus	0	0	0,000	0.00	0.00
Aotus	Saguinus	0	0	0,000	0.00	0.00
Aotus	Ateles	0	0	0,000	0.00	0.00
Aotus	Lagothrix	0	0	0,000	0.00	0.00
Aotus	Alouatta	0	0	0,000	0.00	0.00
Aotus	Brachyteles	0	0	0,000	0.00	0.00
Aotus	Callimico	0	0	0,000	0.00	0.00
Aotus	Cacajao	0	0	0,000	0.00	0.00
Aotus	Chiropotes	0	0	0,000	0.00	.00
Aotus	Pithecia	0	0	0,000	0.00	0.00
Saguinus	Ateles	0	0	0,000	0.00	0.00
Saguinus	Lagothrix	0	0	0,000	0.00	0.00
Saguinus	Alouatta	0	0	0,000	0.00	0.00
Saguinus	Brachyteles	0	0	0,000	0.00	0.00
Saguinus	Callimico	0	0	0,000	0.00	0.00

Saguinus	Cacajao	0	0	0,000	0.00	0.00
Saguinus	Chiropotes	0	0	0,000	0.00	0.00
Saguinus	Pithecia	0	0	0,000	0.00	0.00
Ateles	Lagothrix	0	0	0,000	0.00	0.00
Ateles	Alouatta	0	0	0,000	0.00	0.00
Ateles	Brachyteles	0	0	0,000	0.00	0.00
Ateles	Callimico	0	0	0,000	0.00	0.00
Ateles	Cacajao	0	0	0,000	0.00	0.00
Ateles	Chiropotes	0	0	0,000	0.00	0.00
Ateles	Pithecia	0	0	0,000	0.00	0.00
Lagothrix	Alouatta	0	0	0,000	0.00	0.00
Lagothrix	Brachyteles	0	0	0,000	0.00	0.00
Lagothrix	Callimico	0	0	0,000	0.00	0.00
Lagothrix	Cacajao	0	0	0,000	0.00	0.00
Lagothrix	Chiropotes	0	0	0,000	0.00	0.00
Lagothrix	Pithecia	0	0	0,000	0.00	0.00
Alouatta	Brachyteles	0	0	0,000	0.00	0.00
Alouatta	Callimico	0	0	0,000	0.00	0.00
Alouatta	Cacajao	0	0	0,000	0.00	0.00
Alouatta	Chiropotes	0	0	0,000	0.00	0.00
Alouatta	Pithecia	0	0	0,000	0.00	0.00
Brachyteles	Callimico	0	0	0,000	0.00	0.00
Brachyteles	Cacajao	0	0	0,000	0.00	0.00
Brachyteles	Chiropotes	0	0	0,000	0.00	0.00
Brachyteles	Pithecia	0	0	0,000	0.00	0.00
Callimico	Cacajao	0	0	0,000	0.00	0.00
Callimico	Chiropotes	0	0	0,000	0.00	0.00
Callimico	Pithecia	0	0	0,000	0.00	0.00
Cacajao	Chiropotes	0	0	0,000	0.00	0.00
Cacajao	Pithecia	0	0	0,000	0.00	0.00
Chiropotes	Pithecia	0	0	0,000	0.00	0.00

Callithrix Aotus	0,2895	0,1818	0,628	39.33	78.67
Callithrix Saguinus	0,2895	0,1818	0,628	39.33	78.67
Callithrix Ateles	0,2895	0,1818	0,628	39.33	78.67
Callithrix Lagothrix	0,2895	0,1818	0,628	39.33	78.67
Callithrix Alouatta	0,2895	0,1818	0,628	39.33	78.67
Callithrix Brachyteles	0,2895	0,1818	0,628	39.33	78.67
Callithrix Callimico	0,2895	0,1818	0,628	39.33	78.67
Callithrix Cacajao	0,2895	0,1818	0,628	39.33	78.67
Callithrix Chiropotes	0,2895	0,1818	0,628	39.33	78.67
Callithrix Pithecia	0,2895	0,1818	0,628	39.33	78.67
Cebuella Aotus	0,2895	0,1818	0,628	39.33	78.67
Cebuella Saguinus	0,2895	0,1818	0,628	39.33	78.67
Cebuella Ateles	0,2895	0,1818	0,628	39.33	78.67
Cebuella Lagothrix	0,2895	0,1818	0,628	39.33	78.67
Cebuella Alouatta	0,2895	0,1818	0,628	39.33	78.67
Cebuella Brachyteles	0,2895	0,1818	0,628	39.33	78.67
Cebuella Callimico	0,2895	0,1818	0,628	39.33	78.67
Cebuella Cacajao	0,2895	0,1818	0,628	39.33	78.67
Cebuella Chiropotes	0,2895	0,1818	0,628	39.33	78.67
Cebuella Pithecia	0,2895	0,1818	0,628	39.33	78.67
Saimiri Aotus	0,2895	0,1818	0,628	39.33	78.67
Saimiri Saguinus	0,2895	0,1818	0,628	39.33	78.67
Saimiri Ateles	0,2895	0,1818	0,628	39.33	78.67
Saimiri Lagothrix	0,2895	0,1818	0,628	39.33	78.67
Saimiri Alouatta	0,2895	0,1818	0,628	39.33	78.67
Saimiri Brachyteles	0,2895	0,1818	0,628	39.33	78.67
Saimiri Callimico	0,2895	0,1818	0,628	39.33	78.67
Saimiri Cacajao	0,2895	0,1818	0,628	39.33	78.67
Saimiri Chiropotes	0,2895	0,1818	0,628	39.33	78.67
Saimiri Pithecia	0,2895	0,1818	0,628	39.33	78.67
Callicebus Aotus	0,2895	0,1818	0,628	39.33	78.67

Callicebus	Saguinus	0,2895	0,1818	0,628	39.33	78.67
Callicebus	Ateles	0,2895	0,1818	0,628	39.33	78.67
Callicebus	Lagothrix	0,2895	0,1818	0,628	39.33	78.67
Callicebus	Alouatta	0,2895	0,1818	0,628	39.33	78.67
Callicebus	Brachyteles	0,2895	0,1818	0,628	39.33	78.67
Callicebus	Callithrix	0,2895	0,1818	0,628	39.33	78.67
Callicebus	Cacajao	0,2895	0,1818	0,628	39.33	78.67
Callicebus	Chiropothes	0,2895	0,1818	0,628	39.33	78.67
Callicebus	Pithecia	0,2895	0,1818	0,628	39.33	78.67
Leontopithecus	Aotus	0,2895	0,1818	0,628	39.33	78.67
Leontopithecus	Saguinus	0,2895	0,1818	0,628	39.33	78.67
Leontopithecus	Ateles	0,2895	0,1818	0,628	39.33	78.67
Leontopithecus	Lagothrix	0,2895	0,1818	0,628	39.33	78.67
Leontopithecus	Alouatta	0,2895	0,1818	0,628	39.33	78.67
Lontopithecus	Brachyteles	0,2895	0,1818	0,628	39.33	78.67
Lontopithecus	Callimico	0,2895	0,1818	0,628	39.33	78.67
Lontopithecus	Cacajao	0,2895	0,1818	0,628	39.33	78.67
Lontopithecus	Chiropotes	0,2895	0,1818	0,628	39.33	78.67
Lontopithecus	Pithecia	0,2895	0,1818	0,628	39.33	78.67
Cebus	Callithrix	0,4777	0,305	0,638	58.00	122.00
Cebus	Cebuella	0,4777	0,305	0,638	58.00	122.00
Cebus	Saimiri	0,4777	0,305	0,638	58.00	122.00
Cebus	Callicebus	0,4777	0,305	0,638	58.00	122.00
Cebus	Leontopithecus	0,4777	0,305	0,638	58.00	122.00
Cebus	Aotus	0,1578	0,1072	0,679	23.42	48.58
Cebus	Saguinus	0,1578	0,1072	0,679	23.42	48.58
Cebus	Ateles	0,1578	0,1072	0,679	23.42	48.58
Cebus	Lagothrix	0,1578	0,1072	0,679	23.42	48.58
Cebus	Alouatta	0,1578	0,1072	0,679	23.42	48.58
Cebus	Brachyteles	0,1578	0,1072	0,679	23.42	48.58
Cebus	Callimico	0,1578	0,1072	0,679	23.42	48.58

Cebus	Cacajao	0,1578	0,1072	0,679	23.42	48.58
Cebus	Chiropothes	0,1578	0,1072	0,679	23.42	48.58
Cebus	Pithecia	0,1578	0,1072	0,679	23.42	48.58

Testando a Significância das taxas

Para testar se os valores de dN e dS diferem significativamente aplicamos um teste Z testando três hipóteses: H_0 : $dN=dS$, H_1 : $dN \neq dS$ (teste de neutralidade), $dN>dS$ (seleção positiva) e $dN<dS$ (seleção purificadora). O resultado dos testes mostra que o número de substituições sinônimas difere significativamente do número de substituições não sinônimas em 51 comparações entre gêneros para o gene GH1. A tabela 5 mostra que a hipótese nula foi rejeitada com a probabilidade de 0,05%, contudo a seleção positiva foi descartada uma vez que a hipótese de $dN<dS$ representou melhor os dados, tabela 6.

Para o gene GH5 o teste mostra que o número de substituições sinônimas não difere significativamente do número de substituições não sinônimas. Este resultado indica que o gene GH5 evoluiu de forma neutra. A tabela 7 reporta as probabilidades do teste.

O resultado dos testes para o gene GH6 mostra que número de substituições sinônimas difere significativamente do número de substituições não sinônimas em 65 comparações, e portanto podemos rejeitar a hipótese nula com a probabilidade de 0,05%, ver tabela 8. Contudo o número de substituições não sinônimas não é maior que o número de substituições sinônimas e os resultados da tabela 9 sugerem que o gene GH6 evoluiu sob seleção purificadora.

Tabela 5. Resultados do teste Z, para o gene GH1 ,Hipótese: H0: dN = dS, Ha: dN≠dS, os números em negrito representam as probabilidades <0,05.

	Pithecia	Calithrix	Callicebus	Alouatta	Ateles	Chiropotes	Cebuella	Aotus	Brachyeles	Cebus	Saimiri	Saguinus	Lagothrix	Callimico	Leontopithecus
Calithrix	0.018														
Callicebus	0.001	0.003													
Alouatta	0.002	0.001	0.000												
Ateles	0.264	0.836	0.170	0.495											
Chiropotes	0.418	0.036	0.002	0.004	0.474										
Cebuella	0.025	0.717	0.003	0.001	0.765	0.051									
Aotus	0.001	0.003	0.498	0.001	0.269	0.002	0.006								
Brachyeles	0.003	0.001	0.001	0.471	0.852	0.007	0.003	0.002							
Cebus	0.144	0.788	0.215	0.603	0.911	0.278	0.607	0.267	0.884						
Saimiri	0.002	0.009	0.144	0.001	0.339	0.004	0.009	0.491	0.005	0.422					
Saguinus	0.258	0.501	0.334	0.793	0.338	0.465	0.433	0.492	0.904	0.785	0.595				
Lagothrix	0.230	0.052	0.003	0.005	0.364	0.294	0.074	0.003	0.010	0.370	0.006	0.358			
Callimico	0.038	0.342	0.002	0.000	0.952	0.077	0.276	0.004	0.001	0.793	0.005	0.599	0.112		
Leontopithecus	0.287	0.547	0.300	0.739	0.470	0.508	0.476	0.447	0.960	0.652	0.546	0.311	0.394	0.649	
Cacajao	0.418	0.036	0.002	0.004	0.474	1.000	0.051	0.002	0.007	0.278	0.004	0.465	0.294	0.077	0.508

Tabela 6. Resultados do teste Z para o gene GH1, teste de hipóteses: $dN < dS$, os números em negrito representam as probabilidades $< 0,05$.

	Pithecia	Calithrix	Callicebus	Alouatta	Ateles	Chiropotes	Cebuella	Aotus	Brachyteles	Cebus	Saimiri	Saguinus	Lagothrix	Callimico	Leontopithecus
Calithrix	0.010														
Callicebus	0.000	0.001													
Alouatta	0.001	0.000	0.000												
Ateles	0.130	1.000	0.084	0.253											
Chiropotes	0.202	0.021	0.001	0.002	0.241										
Cebuella	0.014	0.358	0.001	0.000	1.000	0.030									
Aotus	0.000	0.001	0.247	0.000	0.133	0.001	0.002								
Brachyteles	0.002	0.001	0.000	0.231	0.427	0.004	0.001	0.001							
Cebus	0.070	1.000	0.102	0.303	0.457	0.142	1.000	0.126	0.442						
Saimiri	0.001	0.003	0.070	0.000	0.166	0.002	0.003	0.253	0.002	0.202					
Saguinus	0.126	1.000	0.160	0.399	0.166	0.236	1.000	0.239	1.000	1.000	0.290				
Lagothrix	0.113	0.029	0.001	0.003	0.187	0.155	0.042	0.001	0.005	0.187	0.003	0.183			
Callimico	0.022	0.167	0.001	0.000	1.000	0.045	0.154	0.001	0.001	1.000	0.002	1.000	0.062		
Leontopithecus	0.140	1.000	0.143	0.372	0.231	0.257	1.000	0.216	1.000	1.000	0.265	1.000	0.201	1.000	
Cacajao	0.202	0.021	0.001	0.002	0.241	1.000	0.030	0.001	0.004	0.142	0.002	0.236	0.155	0.045	0.257

Tabela 7. Resultado do teste Z para o gene GH5, hipótese $dN \neq dS$, nenhuma estatística foi significativa.

	Callithrix	callicebus	Alouatta	Cebuella	Callimico	Leontopithecus	Saguinus	Cebus	Aotos	Saimiri	Ateles	Lagothrix	Brachiteles	Chiropotes	Cacajao
callicebus	0.272														
Alouatta	0.471	0.325													
Cebuella	1.000	0.272	0.471												
Callimico	1.000	0.272	0.471	1.000											
Leontopithecus	1.000	0.272	0.471	1.000	1.000										
Saguinus	1.000	0.272	0.471	1.000	1.000	1.000									
Cebus	1.000	0.272	0.471	1.000	1.000	1.000	1.000								
Aotos	1.000	0.272	0.471	1.000	1.000	1.000	1.000	1.000							
Saimiri	1.00	0.272	0.471	1.000	1.000	1.000	1.000	1.000	1.000						
Ateles	0.27	1.000	0.325	0.272	0.272	0.272	0.272	0.272	0.272	0.272					
Lagothrix	0.27	1.000	0.325	0.272	0.272	0.272	0.272	0.272	0.272	0.272	1.000				
Brachiteles	0.27	1.000	0.325	0.272	0.272	0.272	0.272	0.272	0.272	0.272	1.000	1.000			
Chiropotes	0.27	1.000	0.325	0.272	0.272	0.272	0.272	0.272	0.272	0.272	1.000	1.000	1.000		
Cacajao	0.27	1.000	0.325	0.272	0.272	0.272	0.272	0.272	0.272	0.272	1.000	1.000	1.000	1.000	
Pithecia	0.27	1.000	0.325	0.272	0.272	0.272	0.272	0.272	0.272	0.272	1.000	1.000	1.000	1.000	1.000

Tabela 8. Resultado do teste Z para o gene GH6, hipótese $dN \neq dS$, resultados em negrito são as probabilidades significativas $p < 0,05$.

	Cebus	Callithrix	Cebuella	Saimiri	Callicebus	Leontopithecus	Aotus	Saguinus	Ateles	Lagothrix	Alouatta	Brachyteles	Callimico	Cacajao	Chiropotes
Callithrix	0.006														
Cebuella	0.006	1.000													
Saimiri	0.006	1.000	1.000												
Callicebus	0.006	1.000	1.000	1.000											
Leontopithecus	0.006	1.000	1.000	1.000	1.000										
Aotus	0.021	0.027	0.027	0.027	0.027	0.027									
Saguinus	0.021	0.027	0.027	0.027	0.027	0.027	1.000								
Ateles	0.021	0.027	0.027	0.027	0.027	0.027	1.000	1.000							
Lagothrix	0.021	0.027	0.027	0.027	0.027	0.027	1.000	1.000	1.000						
Alouatta	0.021	0.027	0.027	0.027	0.027	0.027	1.000	1.000	1.000	1.000					
Brachyteles	0.021	0.027	0.027	0.027	0.027	0.027	1.000	1.000	1.000	1.000	1.000				
Callimico	0.021	0.027	0.027	0.027	0.027	0.027	1.000	1.000	1.000	1.000	1.000	1.000			
Cacajao	0.021	0.027	0.027	0.027	0.027	0.027	1.000	1.000	1.000	1.000	1.000	1.000	1.000		
Chiropotes	0.021	0.027	0.027	0.027	0.027	0.027	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
Pithecia	0.021	0.027	0.027	0.027	0.027	0.027	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Tabela 9. Resultado teste Z para o gene GH6, Hipótese $dN < dS$, resultados em negrito são as probabilidades significativas $p < 0,05$.

	Cebus	Callithrix	Cebuella	Saimiri	Callicebus	Leontopithecus	Aotus	Saguinus	Ateles	Lagothrix	Alouatta	Brachyteles	Callimico	Cacajao	Chiropotes
Callithrix	0.002														
Cebuella	0.002	1.000													
Saimiri	0.002	1.000	1.000												
Callicebus	0.002	1.000	1.000	1.000											
Leontopithecus	0.002	1.000	1.000	1.000	1.000										
Aotus	0.011	0.011	0.011	0.011	0.011	0.011									
Saguinus	0.011	0.011	0.011	0.011	0.011	0.011	1.000								
Ateles	0.011	0.011	0.011	0.011	0.011	0.011	1.000	1.000							
Lagothrix	0.011	0.01	0.011	0.011	0.011	0.011	1.000	1.000	1.000						
Alouatta	0.011	0.01	0.011	0.011	0.011	0.011	1.000	1.000	1.000	1.000					
Brachyteles	0.011	0.01	0.011	0.011	0.011	0.011	1.000	1.000	1.000	1.000	1.000				
Callimico	0.011	0.01	0.011	0.011	0.011	0.011	1.000	1.000	1.000	1.000	1.000	1.000			
Cacajao	0.011	0.01	0.011	0.011	0.011	0.011	1.000	1.000	1.000	1.000	1.000	1.000	1.000		
Chiropotes	0.011	0.01	0.011	0.011	0.011	0.011	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
Pithecia	0.011	0.01	0.011	0.011	0.011	0.011	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Teste de Seleção pelo método de Máxima Verossimilhança

Modelos de Sítio

Os resultados dos LRTs (teste da taxa de verossimilhança) para seleção positiva nos modelos de sítio são mostrados na tabela 10. Para este modelo todos os testes apoiaram a hipótese de influência da seleção positiva para os três genes estudados, desde o teste mais simples M0-M3 que é mais um teste de pressão seletiva variável entre os sítios ao mais rigoroso teste M1a-M2a e M7-M8. Contudo para o gene GH1 os testes mostram 10 sítios 35R, 37L, 62C, 76K, 95L, 105I, 128E, 136N, 175T, 182T evoluindo sob seleção positiva no gene GH1 pela análise Bayesiana Posterior através da metodologia Bayes Empirical Bayes (BEB). A análise se baseia na comparação de modelos que testam a hipótese nula de que $dN/dS \leq 1$. Isto é feito pela comparação do teste da taxa de verossimilhança (LRT) do modelo que fixa $dN/dS=1$ ao que estima o parâmetro dN/dS . Sob a hipótese nula ($dN/dS \leq 1$), o teste estatístico $2\Delta l$ é esperado ter uma distribuição X^2 , com o número de graus de liberdade igual a diferença do número de parâmetros dos modelos. Portanto nestas análises utilizamos 1 grau de liberdade que comparados a distribuição de X^2 obtemos um valor crítico de 3,84.

Para o gene GH5, os testes mais rigorosos M1a-M2a e M7-M8 mostram 5 sítios 29 S, 34 S, 93 T, 95 Q, 99 L evoluindo sob seleção positiva pela análise Bayesiana Posterior através da metodologia Bayes Empirical Bayes (BEB), ver tabela 11.

Para o gene GH6 todos os modelos de seleção positiva representaram melhor os dados do que os modelos de neutralidade, mostrando 17 sítios 3 T, 5

S, 7 M, 11 P, 12 R, 19 P, 20 A, 21 L, 25 D, 26 W 29 S, 34 S, 93 T, 95 Q, 99 L, 123 C, 176 S, evoluindo sob seleção positiva pela análise Bayesiana Posterior através da metodologia Bayes Empirical Bayes (BEB), ver tabela 12.

Table 10 Teste de Máxima Verossimilhança para o GH1, mostrando os modelos, o valor de verossimilhança para cada modelo a comparação através do LRT (Teste de taxa de verossimilhança) e os sítios sobre seleção positiva.

Modelos	Parâmetros	Valor de Verossimilhança	LRT	Sítios selecionados positivamente
M0	W=0.56	- 2949.93	M0 x M3	Não mostrado
M3	P0=0.59	- 2703.57	492.72	
	P1=0.28			
	P2=0.11			
	W0=0			
	W1=1.13			
	W2=5.12			
M1a	P0=0.59	-2746.44	M1a x M2a	Não mostrado
	P1=0.40			
	W0=0		85.12	
	W1=1			
M2a	P0=0.59	- 2703.88		35R, 37L, 76K, 128E, 136N, 175T, 182T
	P1=0.28			
	P2= 0.12			
	W0=0			
	W1= 1			
	W2=4.68			
M7	P=0.005	-2746.44	M7 x M8	Não mostrado
	Q=0.007		84.82	
M8	P0=0.86	-2704.03		35R, 37L, 62C, 76K, 95L, 105I, 128E, 136N, 175T, 182T
	P=0.005			
	Q=0.012			
	P1=0.13			
	WS=4.60			

Todas as estatística são altamente significantes. $P < 0,001$

Tabela 11 Teste de Máxima Verossimilhança para o GH5, mostrando os modelos, a estatística LNL para cada modelo a comparação através do LRT e os sítios sobre seleção positiva.

Modelos	Parâmetros	LNL	LRT	Sítios selecionados positivamente
M0	W= 0.82	-1884,22	M0 x M3	Não mostrado
M3	P0=0.54152 P1=0.41374 P2=0.04474 W0=0.01250 W1=3.496999 W2=99.00000	-1773.29	221.86	2A,3T,4G,5S,7M,10A,11 P,12R,13S,14W,17A,19 P, 20A, 21L,24R,25D,26 W, 27C, 28L, 29S,30K, 31N,32S,33L,35Q,36A, 37F,38W,39R,40R,46I,50 S, 52A,53L,54T, 57R, 63 M,65Q, 66K,74A, 87R,90 P,97Q
M1a	P0=0.49719 P1=0.50281 W0= 0.00000 W1= 1.00000	-1827.04	M1a x M2a 76.94	Não mostrado
M2a	P0=0.44525 P1=0.46867 P2= 0.08608 W0=0.00000 W1= 1.00000 W2=36.01831	-1788.57		93 T 95 Q 99 L 29 S
M7	P=0.00500 Q=0.00503	-1827.04	M7 x M8 76.88	Não mostrado
M8	P0=0.91363 P=0.00524 Q=0.00527 P1=0.08637 WS=35.39309	-1788.60		29 S 34 S 93 T 95 Q 99 L

Todas as estatística são altamente significativas. $P < 0,001$

Table 12 Teste de Máxima Verossimilhança para o GH6, mostrando os modelos, os valores de verossimilhança para cada modelo a comparação através do LRT e os sítios sobre seleção positiva.

Modelos	Parâmetros	Valores de verossimilhança	LRT	Sítios selecionados positivamente
M0	W= 0.89	-3012.95	M0 x M3	Não mostrado
M3	P0=0.63 P1=0.26 P2=0.09 W0=0.04 W1=5.60 W2=0.764	-2674.97	675.96	2A,3T,4G,5S,7M,10A,11 P,12R,13S,14W,17A,19 P, 20A, 21L,24R,25D,26 W, 27C, 28L, 29S,30K, 31N,32S,33L,35Q,36A, 37F,38W,39R,40R,46I,50 S, 52A,53L,54T, 57R, 63 M,65Q, 66K,74A, 87R,90 P,97Q
M1a	P0=0.59 P1=0.40 W0= 0.001 W1= 1.00000	-2792.90	M1a x M2a 202.04	Não mostrado
M2a	P0=0.56 P1=0.26 P2= 0.16 W0=0.00000 W1= 1.00000 W2=10.45	-2691.88		3 T, 5 S, 7 M, 11 P, 12 R, 19 P, 20 A,21 L,25 D,26 W, 29 S,30 K,33 L,35 Q
M7	P=0.007 Q=0.009	-2807.79	M7 x M8 194.62	Não mostrado
M8	P0=0.75 P=0.005 Q=0.01P1=0.24 WS=6.77	-2710.48		3 T, 5 S, 7 M, 11 P,12 R, 19 P, 20 A, 21 L, 25 D, 26 W 29 S, 34 S, 93 T, 95 Q,99 L, 123 C, 176 S

Todas as estatística são altamente significativas. $P < 0,001$

Modelos de Ramos

Para esta análise dividimos previamente os ramos da árvore filogenética em ramos de primeiro plano (que são os ramos que esperamos estarem sobre seleção positiva) e ramos segundo plano (os ramos restantes da árvore). Após classificarmos os ramos e marcá-los na árvore, passamos então para construção do teste de hipóteses: Hipótese alternativa presume que W (dN/dS) nos ramos marcados como primeiro plano diferem do dN/dS médio entre os ramos de segundo plano. Hipótese nula fixa dN/dS nos ramos de primeiro plano =1.

Os resultados da análise de ramos para o gene GH1 encontra-se na tabela 13, sendo que os números identificados como ramos são referentes a relação dos ramos de um nó interno a um nó terminal ou a outro nó interno

como apresentado na figura 5. A Hipótese alternativa presume que W (dN/dS) nos ramos marcados como primeiro plano diferem do dN/dS médio entre os ramos de segundo plano. Um teste de LRT foi usado para determinar se o modelo alternativo representa significativamente melhor os dados que o modelo nulo. A probabilidade associada foi calculada pela comparação de duas vezes a diferença nos valores de log likelihood à distribuição de X^2 , com o grau de liberdade igual a diferença do número de parâmetros entre as duas hipóteses comparadas. O teste de hipóteses mostra que existe diferença entre as taxas dN/dS entre os ramos de primeiro plano e segundo plano, contudo as taxas calculadas por este método não são maiores que 1 nos ramos foreground, sendo que estes resultados estão apresentados na tabela 14.

Embora existam alguns sítios sobre seleção positiva apontados pelo teste de sítios, não encontramos diferenças na taxa W entre os ramos de primeiro e segundo plano na árvore. A hipótese nula representou melhor os dados do que a hipótese alternativa para os genes GH5 e GH6. A tabela 15 e 16 mostram os resultados dos testes de taxa de verossimilhança para o gene GH5 e GH6.

Tabela 13 Resultado da análise de Ramos para o gene GH1, mostrando a relação entre ramos, a taxa dN/dS, dN e dS. Os ramos identificados por números referem-se à relação dos ramos de um nó interno a um nó terminal ou a outro nó interno. Os números em negrito representam os ramos marcados como primeiro plano.

Ramos	dN/dS	d N	d S
23...22	0,5585	0	0
Aotus...18	0,5585	0	0
22...18	0,5585	0,1	0,1791
30...28	0,5585	0	0
Cebus...23	0,5585	0,0051	0,0091
Saimiri...23	0,5585	0,0991	0,1774
18...2	0,5852	0,0013	0,0023
Saguinus...24	0,5585	0,0138	0,0248
26...24	0,5585	0	0
Leont...26	0,5585	0	0
26...28	0,5852	0,104	0,1777
Callimico...28	0,5585	0,0013	0,0024
24...22	0,5852	0	0
Callithrix...30	0,5585	0,004	0,0072
Cebuella...30	0,5852	0,109	0,0187
13...12	0,5852	0	0
10...11	0,5852	0,0209	0,0358
11...12	0,5852	0	0
4...5	0,5585	0	0
Brachy...13	0,5585	0,0024	0,0044
Lagothrix...13	0,5585	0,1264	0,2264
Ateles...12	0,5585	0,1141	0,2043
Alouatta...11	0,5585	0,0065	0,0116
10...3	0,5585	0	0
3...4	0,5585	0,1295	0,2319
10...2	0,5852	0	0
Cacajao...5	0,5585	0	0
Chiropotes...5	0,5585	0	0
Pithecia...4	0,5585	0	0
Callicebus...3	0,5585	0,0027	0,0048

Tabela 14. Resultado dos LRTs para o modelo de ramos, com os modelos utilizados, os parâmetros dos testes, os valores de verossimilhança e o teste de taxa de verossimilhança.

Modelos	Parâmetros do teste	Valor de Verossimilhança	LRT
Hipótese alternativa	Ramos de primeiro plano tem $dN/dS \neq$ dos ramos de segundo plano	-2984.36	8.04
Hipótese nula	Ramos de primeiro plano $dN/dS = 1$	-2988.38	P<0,001

Tabela 15. Resultado dos LRTs para o modelo de ramos para o gene GH5.

Modelos	Parâmetros do teste	Valores de verossimilhança	LRT
Hipótese alternativa	Ramos de primeiro plano tem $dN/dS \neq$ dos ramos de segundo plano	-1884,22	0
Hipótese nula	Ramos de primeiro plano $dN/dS = 1$	-1884,22	

Tabela 16. Resultado do teste de ramos para o gene GH6, mostrando os parâmetros do teste, os valores de verossimilhança e LRTs.

Modelos	Parâmetros do teste	Valores de verossimilhança	LRT
Hipótese alternativa	Ramos de primeiro plano tem $dN/dS \neq$ dos ramos de segundo plano	-2996,57	0,48
Hipótese nula	Ramos de primeiro plano $dN/dS = 1$	-2996,81	

Modelo Sítio-Ramo

O modelo sítio-ramo foi desenvolvido para possibilitar o teste de seleção positiva num subconjunto de sítios (códon) num ramo (ou ramos) da árvore. Procedemos aqui como no modelo de ramos, escolhendo previamente os ramos de primeiro plano, contudo o que difere é a construção das hipóteses.

A tabela 17 mostra os modelos com seus respectivos parâmetros, os valores de verossimilhança e LRTs e os sítios selecionados resultado da análise de sítio-ramo para o gene GH1. Neste teste temos uma hipótese alternativa que define quatro classes de sítios como segue: este modelo assume sítios de classe 0 onde estão os sítios sobre seleção purificadora em todos os ramos ($0 < W_0 < 1$). Sítios de classe 1 que incluem códon que estão evoluindo sob neutralidade na árvore, com $W_1 = 1$ e códon de classe 2a e 2b que são conservados ou neutros nos ramos de segundo plano, mas estão sobre seleção positiva nos ramos de primeiro plano, com $W_2 > 1$. Já a hipótese nula é construída exatamente da mesma forma que a alternativa com a única

diferença que o $W_2=1$, ou seja na hipótese nula não admite-se sítios sobre seleção positiva nas linhagens de primeiro plano, e da mesma forma que no modelo de ramos avaliamos a significância do teste através do LRT.

O modelo alternativo (seleção positiva nos sítios dentro dos ramos de primeiro plano) representou melhor os dados do que o modelo nulo. Esta análise também confirma os 10 sítios sobre seleção positiva encontrados pelo modelo de sítios.

A tabela 18 mostra os modelos com seus respectivos parâmetros, os valores de verossimilhança, LRTs e os sítios selecionados resultado da análise de sítio-ramo para o gene GH5. O modelo nulo (as taxas dN/dS são iguais entre ramos foreground e background) representou melhor os dados do que o modelo alternativo, o que indica que o gene evoluiu de forma neutra. Neste teste não confirmamos os 5 sítios encontrados previamente sob seleção positiva.

Para o gene GH6 novamente o modelo nulo representou melhor os dados que o modelo alternativo, não obtendo nenhum sítio selecionado positivamente dentro dos ramos de primeiro plano, confirmando que o gene evoluiu de forma neutra, ver tabela 19.

Tabela 17. Resultado do testes Sítio-Ramo, os números em negrito mostram a significância do teste para o gene GH1.

Modelos	Parâmetros 0 1 2a 2b	Valores de verossimilhança	LRT	Sítios Selecionados
Modelo Alternativo	p 0.59 0.40 0 0	-2797.45		35R, 37L,62C,76K, 95L,105I,128E, 136N,175T, 182T
	bac W 0 1 0 1		110.1	
	for W 0 1 1 1		p<0,001	
Modelo Nulo	p 0.59 0.40 0 0	-2742.40		
	bac W 0 1 0 1			
	for W 0 1 1.5 1.5			

Tabela 18. Resultado do testes Sítio-Ramo para o gene GH5.

Modelos	Parâmetros 0 1 2a 2b	Valores de verossimilhança	LRT	Sítios Selecionados
Modelo Alternativo	p 0.59 0.40 0 0	-1827,04		29S
	bac W 0 1 0 1		0	
	for W 0 1 1 1			
Modelo Nulo	p 0.59 0.40 0 0	-1827,04		
	bac W 0 1 0 1			
	for W 0 1 1.5 1.5			

Tabela 19. Resultado do testes Sítio-Ramo para o gene GH6..

Modelos	Parâmetros					Valor de verossimilhança	LRT	Sítios Seleccionados
	O	1	2a	2b				
Modelo Alternativo	p	0.59	0.40	0	0	-2750,34		nenhum
	bac W	0	1	0	1		0,6	
	for W	0	1	1	1			
Modelo Nulo	p	0.59	0.40	0	0	-2750,25		
	bac W	0	1	0	1			
	for W	0	1	1.5	1.5			

Estimativa do Estado ancestral dos códons sobre seleção positiva

Para mapear as mudanças de aminoácidos sobre seleção positiva no gene GH1, estimamos o estado ancestral por parcimônia das mudanças de códon nos sítios, 35,37,76,105,128,136 estão relacionadas na figura 3. Para os códons 62, 95, 175, 182 não obtivemos certeza na reconstrução do carácter ancestral, por isso eles não estão identificados na figura.

Figura 4 . Esquema Filogenético dos Primatas do Novo Mundo, representando as mudanças de aminoácidos nos respectivos ramos da filogenia.

Legenda: Códons: 35 — 105 —
 37 — 128 —
 76 — 136 —

Análise de Correlação entre dados morfométricos e moleculares

A tabela 20 mostra o resultado da análise da correlação entre as variáveis morfológicas e as taxas dN, dS e dN/dS entre os ramos estimados pelo programa PAML . Existe correlação significativa entre a taxa dN/dS e a direção que a evolução tomou em relação a linha de menor resistência evolutiva fenotípica ($R=0,55$, $P=0,0045$) e genética ($R=0,43$, $P= 0,004$). Todas as correlações restantes não são significativas.

A tabela 21 mostra o resultado das correlações entre as variáveis morfológicas e as taxas dN, dS e dN/dS entre o nó ancestral e os ramos terminais pela análise par a par. Devido a algumas relações nó ancestral→terminal resultarem em estimativas de dN, dS com valores muito baixos decidimos realizar duas correlações: 1) Utilizando todas as taxas estimadas entre nó ancestral→terminal mesmo que esta taxa seja zero. 2) Excluimos da correlação as taxas com valor zero. Observamos correlações marginalmente significativas entre a taxa dN/dS e direção que a evolução tomou em relação a linha de menor resistência evolutiva fenotípica e genética ($R= \mathbf{0.346}$, $P= 0.106$ e $R= \mathbf{0.350}$, $P= 0.101$), tamanho absoluto ($R= \mathbf{0.333}$, $P= 0.120$) e com a taxa de mudança morfológica ($R=\mathbf{0.452}$, $P= 0.03$), contudo esta correlação aumenta quando retiramos as relações onde as taxas são estimadas como zero. As figuras 5, 6 e 7 mostram os valores das taxas dN, dS e dN/dS em cada nó na filogenia dos primatas neotropicais.

Tabela 20. Correlações entre a direção da evolução em relação a linha de menor resistência evolutiva genética e fenotípica, a quantidade de mudança morfológica, taxa de mudança morfológica e tamanho absoluto com as taxas dN/dS, dN e dS, conforme estimadas pelo modelo de ramos do PAML. Os valores em negrito representam as correlações estatisticamente significativas.

Variáveis Morfométricas	dN/dS	dN	dS
Direção da evolução em relação a LLER genética	0,4399	-0,1292	-0,137
Direção da evolução em relação a LLER fenotípica	0,5593	-0,1134	-0,1247
Quantidade de mudança	-0,00365	-0,234	-0,2312
Taxa de mudança	-0,3125	-0,2391	-0,2395
Tamanho absoluto	0,137	-0,3644	-0,3647

Tabela 21 Correlações entre a direção da evolução em relação a linha de menor resistência evolutiva genética e fenotípica, a quantidade de mudança morfológica, taxa de mudança morfológica e tamanho absoluto com a taxa dN/dS, dN e dS estimados pela abordagem de reconstrução de caracteres ancestrais. Os valores em negrito representam as correlações estatisticamente significativas. O asterisco (*) representa as análises que não consideraram os ramos com dN ou dS com valores zero

Variáveis Morfométricas	dN/dS	*dN/dS	dN	dS
Direção da evolução em relação a LLER genética	0.350	0.362	-0.288	-0.547
Direção da evolução em relação a LLER fenotípica	0.346	0.369	-0.360	-0.616
Quantidade de mudança	0.187	0.183	0.290	0.280
Taxa de mudança	0.333	0.334	0.192	0.131
Tamanho absoluto	0.452	0.458	0.027	-0.091

Figura 5 Árvore filogenética dos primatas do Novo Mundo, mostrando entre parênteses os valores da correlação da direção que a evolução segue na linha de menor resistência evolutiva fenotípica e os valores da taxa de substituição não sinônima (dN) para cada ramo. Ramos com taxa zero não são mostrados.

Tamanho do Crânio

-2.50 -2.00 -1.50 -1.00 -0.50 0.00 0.50 1.00 1.50 2.00

Taxon

Phylogenetic tree showing relationships between various primate taxa, with skull size (Tamanho do Crânio) on the x-axis ranging from -2.50 to 2.00. The tree includes taxa such as *Cacajao* (0.73), *Chiropterus* (0.02), *Pithecia* (0.45), *Callicebus* (0.85), *Cebuella* (0.83), *Callithrix* (0.92), *Calimico* (0.81), *Leontopithecus* (0.48), *Saguinus* (0.88), *Saimiri* (0.84), *Aotus* (0.71), *Cebus* (0.86), *Brachyteles* (0.75), *Lagothrix* (0.29), *Ateles* (0.66), *Alouatta* (0.68), and *Cebus* (0.86). Branch lengths and posterior probabilities are indicated. Red arrows indicate shifts in skull size. Small images of skulls are provided for several taxa.

Skull size

-2.50 -2.00 -1.50 -1.00 -0.50 0.00 0.50 1.00 1.50 2.00

Taxon

Cacajao (0.73)

Chirotopotes (0.02)

5 (0.65)

Pithecia (0.45)

0.5

4 (0.71)

Brachyteles (0.75)

Lagothrix (0.29)

13 (0.84)

0.31

Ateles (0.66)

0.71

12 (0.72)

Alouatta (0.68)

11 (0.91)

0.70

10 (0.92)

0.93

3 (0.83)

1.25

Callicebus (0.85)

0.63

Callithrix (0.92)

Cebuella (0.83)

30 (0.93)

Calumico (0.81)

28 (0.97)

1.43

Leontopithecus (0.48)

26 (0.91)

0.66

Saguinus (0.88)

0.72

24 (0.90)

Saimiri (0.84)

23 (0.76)

Cebus (0.86)

0.92

22 (0.92)

0.69

Aotus (0.71)

18 (0.92)

1

2

Análise Filogenética

Para estas análises utilizei o método de Máxima Parcimônia (MP) para diferentes regiões (éxons e sequência total) dos genes GH e obtive topologias similares. A figura 8 mostra a árvore filogenética inferida por MP construída a partir da sequência dos éxons. Podemos notar a formação de 3 grupos monofiléticos fortemente suportados por valores de bootstrap de 99%, para cada gene do conjunto GH estudados neste trabalho. No entanto estes genes não reconstroem as relações filogenéticas esperadas para os primatas do novo mundo, com excessão do gene GH5 que recupera quase totalmente dois clados, os Calithrichineos e Atelidae + Pitheciidae.

Figura 8 Árvore Filogenética, construída pelo método de Máxima Parcimônia, mostrando as relações entre os três genes do conjunto GH.

Discussão

Os resultados dos testes de seleção pelo método par a par foram negativos para seleção positiva nos três genes estudados, sendo que o gene GH1 foi o único gene para o qual obtivemos a taxa $dN/dS > 1$. Contudo, o teste de hipóteses implementado pelo teste Z ($H_0: dN=dS$; $H_1: dN>dS$), aponta para seleção purificadora nos genes GH1 e GH6 e neutralidade para o gene GH5. Comparações par a par têm pouco poder estatístico porque trabalham com as médias da taxa W (dN/dS) sobre sítios e sobre o tempo. Métodos para detectar seleção positiva ao longo de linhagens funcionam somente se a média da taxa W sobre todos os sítios é maior que 1. Outra diferença metodológica que contribui para diferenças acerca do resultado final é que os métodos que ignoram a taxa de transição/transversão levam a uma superestimativa de substituições sinônimas, e uma superestimativa de dS acaba subestimando a taxa W (Yang & Bielawski 2000).

Uma característica importante da evolução de proteínas é que uma grande proporção de aminoácidos pode ser invariável (W aproximadamente zero) devido às fortes consequências funcionais que uma mudança acarretaria. Assim, muitas proteínas parecem estar sob seleção purificadora na maior parte do tempo. No entanto, a evolução adaptativa pode ocorrer em poucos aminoácidos, levando a taxa W a não ser significativamente maior que 1 (Li 1997) mesmo que esteja havendo seleção positiva.

Conclusões acerca da evolução do conjunto de genes GH baseados somente na análise par a par não nos dão evidências estatísticas fortes de seleção positiva, embora qualitativamente (ver tabela 2) alguns pares de comparações apresentaram $W > 1$, o que poderia sugerir seleção

diversificadora atuando no GH1. A utilização de uma abordagem que preserva a relação filogenética entre os grupos, por outro lado, nos permitiu extrair mais informação, uma vez que, pelo menos no caso do gene GH1, os resultados obtidos pela análise de máxima verossimilhança apontam em outra direção.

Todos os modelos de seleção positiva para o gene GH1 representaram melhor os dados do que os modelos de evolução neutra, sendo que as comparações obtidas entre os modelos M0-M3 (LRT= 492.72), M1a-M2a (LRT=85.12) e M7-M8 (LRT=84.82) foram altamente significativas $p < 0,001$ (ver tabela 10). Esses testes evidenciaram 10 sítios selecionados positivamente. Consideramos esses resultados uma boa indicação de seleção positiva, uma vez que muitos trabalhos que testaram a eficiência e o poder da análise (Wong *et al.* 2004) observaram que a taxa de falsos positivos pelo Bayes Empirical Bayes para a comparação dos modelos M8 (com seleção positiva) *versus* M7 (sem seleção positiva) ,usando o limite de 0.95, se mostrou abaixo de 5%.

Baseados na análise morfológica, selecionamos ramos que sofreram maior diferença em tamanho como potenciais candidatos à seleção positiva, marcando estes ramos na filogenia como grupo de primeiro plano para aplicar o teste de ramos. Os resultados desse teste sugerem que existe diferença entre as taxas dN/dS entre os ramos de primeiro plano e segundo plano; contudo, as taxas calculadas por esse método não são maiores que 1 nos ramos de primeiro plano (ver tabelas 13 e 14). Um problema particular do teste de ramos é que, nos casos em que a divergência de alguns gêneros e seus ancestrais ocorreu em um curto período evolutivo, o número de substituições sinônimas e não sinônimas não é estimado, resultando numa taxa dN/dS estimada a partir das calculadas para outros ramos da árvore. O teste de ramos estima o valor de W para a proteína inteira e, por isso, não é

considerado um bom teste, particularmente porque a seleção positiva é mais provável de ocorrer em alguns sítios (Nickel *et al.* 2008). Historicamente, isso levou a construção do teste de hipótese que considera sítios e ramos ao mesmo tempo (Yang & Nielsen 2002)

As estatísticas obtidas para o teste sítio-ramo confirma os 10 sítios sob seleção positiva nas linhagens de primeiro plano (ver tabela 17). Os 10 sítios positivos para seleção encontram-se na maioria dispostos nas hélices-quatro, sendo que apenas dois sítios estão posicionados nas hélices dois e três, onde se localizam resíduos que participam da interação com o receptor. Trabalhos recentes como o de Braun *et al* (2007) também mostraram seleção positiva no gene GH. Contudo, nesses trabalhos a duplicação gênica é considerada o mecanismo mais importante para a divergência funcional e é considerada uma forte explicação para a seleção positiva (Ohno 1970, Otha 1993). No entanto, indícios de seleção positiva no GH nunca foram detectados num contexto em que mudanças no organismo podem ser influenciadas pela forte restrição evolutiva encontrada em um gene que faz parte da via metabólica responsável pelo crescimento, como no presente trabalho.

A variação dos valores de W entre diferentes sítios durante a evolução, que fica evidente em nossos resultados, sugere que existe uma pressão seletiva diferencial entre sítios, possivelmente indicando que há diferenças em aminoácidos importantes no que diz respeito à função do hormônio do crescimento.

Nossos resultados mostram aceleração na taxa de evolução do gene GH1, corroborando os trabalhos de Otha 1993 e Wallis 1996 e 1997, que relatam episódios de aceleração na taxa de substituições para o gene GH em

mamíferos, inclusive para primatas. Nossos resultados sugerem que esta aceleração é dirigida por seleção positiva.

Quando caracteres sob seleção natural são completamente independentes um do outro, a resposta evolutiva pode ocorrer livremente e certo caráter pode rapidamente adaptar-se ao ambiente. Contudo, quando um caráter depende de um mecanismo comum a outro caráter para sua expressão, ele pode evoluir como uma unidade, potencialmente pressionada em resposta à seleção natural (Maynard Smith *et al.* 1985; Arnold 1992; Schluter 1996). É provável que muito da evolução adaptativa dependa da relação entre a força seletiva externa e a restrição interna, sendo que os hormônios possivelmente atuem em ambos os lados dessa interação.

Marroig & Cheverud (2005) demonstraram, por meio de análises morfométricas do crânio de primatas neotropicais, que existe uma alta correlação entre as diferenças de tamanho absolutas e a quantidade de diferença morfológica entre grupos; em outras palavras, muito da diversificação morfológica que ocorreu foi relacionada ao tamanho e dirigida por seleção natural. Além disso, a irradiação adaptativa em primatas neotropicais ocorreu principalmente através de uma mudança do tamanho geral associada a tipos de dietas diferentes. Quando correlacionamos nossos resultados de dN/dS com algumas variáveis utilizadas no trabalho de Marroig & Cheverud (2005), obtivemos correlações marginalmente significativas entre a taxa dN/dS e a direção que a evolução tomou em relação a linha de menor resistência evolutiva fenotípica e genética (figura 9a e b). Isso significa que a taxa dN/dS é maior quando a evolução morfológica ocorreu no sentido de maior variação genética e fenotípica. Os gráficos c e d na figura 9 mostram ainda que o aumento da taxa dN/dS segue o sentido da mudança de tamanho

absoluto e da taxa de mudança morfológica. Nossos resultados sugerem que a diversificação morfológica durante a irradiação adaptativa dos primatas do Novo Mundo é acompanhada pela seleção positiva sobre o gene do GH. Esse tipo de seleção atuou diretamente sobre sítios funcionalmente importantes para a interação com o receptor do hormônio (localizados nas hélices 2 e 3), provavelmente influenciando a função de uma molécula fundamental para o crescimento do corpo.

Figura 9. Gráficos de correlação entre a taxa dN/dS e a direção que a evolução tomou em relação a linha de menor resistência evolutiva fenotípica (a) e genética (b), com o a mudança de tamanho absoluto (c) e a taxa de mudança morfológica (d). Os números dentro dos gráficos referem-se aos nós da figura 7. Os gráficos referem-se à taxa dN/dS calculada retirando-se as estimativas com valor zero de dN e/ou de dS.

(a)

(b)

(c)

(d)

As análises de seleção positiva nos genes GH5 e GH6 mostraram resultados semelhantes entre si. Nenhum teste foi estatisticamente significativo para diferença entre o número de substituição sinônima e não sinônima para cada gene na análise par a par; adicionalmente, o teste de máxima verossimilhança não identificou sítios sob pressão seletiva diferente para nenhum ramo da árvore. Pouco se sabe a respeito da fisiologia desses genes. Dentre os poucos estudos em humanos está o de Barrera-Saldanha *et al* (1998), que detectou ausência do GH5 e GH6 na placenta. Os resultados de seu estudo revelaram que a ausência do transcrito não influencia na

manutenção da gestação e não causa nenhum tipo retardamento no crescimento fetal. Segundo Lacroix *et al* (2002), o hormônio de crescimento placentário não parece ter efeito direto no crescimento fetal, uma vez que não é detectado na circulação fetal. Contudo, o GH placentário tem clara atividade somatotrófica e pode alterar o metabolismo maternal durante a gravidez. Um dos poucos estudos sobre a evolução do conjunto GH em primatas não humanos é o de Revol de Mendoza *et al* (2004), cujas análises demonstram que os genes relacionados ao GH surgiram independentemente em primatas do Novo e Velho Mundo, bem como nas linhagens de humanos e grandes macacos. Esses estudos também demonstraram que somente os genes GH1 e GH5 de humanos e chimpanzés são ortólogos; os demais são todos parálogos. Muitos estudos têm sido realizados a respeito da fisiologia desses hormônios em humanos, mas estudos com outros primatas, sobretudo neotropicais, são inexistentes. Isso impossibilita qualquer visão clara sobre a real atuação do GH placentário no desenvolvimento fetal.

Nossos resultados mostram que membros do conjunto GH apresentam diferentes restrições evolutivas. A seleção positiva é dominante nos genes GH1, em contraste com a seleção purificadora que predomina a evolução do gene GH6 e no gene GH5 (onde os valores de dN e dS não mostram diferenças significativas, em acordo com um modelo de evolução neutra). Ye *et al* (2005) também encontrou membros da família GH evoluindo sob pressão seletiva diferente nas linhagens de primatas do Velho Mundo. Contudo, como não existem dados sobre estrutura, função e expressão desses genes em primatas do Novo Mundo, não podemos determinar o significado biológico dessa variação de pressão seletiva, e nem determinar se os membros do conjunto GH são utilizados de maneira diferente no tempo e espaço.

Todos os métodos de detecção de seleção positiva utilizados neste trabalho parecem ser conservativos, mas a abordagem par a par tem pouco poder porque trabalha com as médias da taxa W sobre os sítios e sobre o tempo. Métodos para detectar seleção positiva entre linhagens funcionam somente se a taxa W média sobre todos os sítios é maior que 1. Se a evolução adaptativa ocorrer somente num curto intervalo de tempo e afetar poucos aminoácidos, nenhum dos métodos terá sucesso (Wong *et al.* 2004)

Outra importante observação é que a constância da pressão seletiva entre sítios parece ser muito mais importante do que a constância entre linhagens, especialmente para genes que sofrem pressão seletiva contínua. Por isso, modelos que assumem pressão seletiva variável entre sítios têm uma performance melhor e maior poder na detecção de seleção positiva (Yang & Bielawski 2000). No entanto, métodos baseados na máxima verossimilhança agrupam os códons em duas ou mais categorias, com diferentes valores de W . Nesse caso, a hipótese nula de $W=1$ é testada indiretamente para os grupos com $W>1$, uma vez que os sítios com valores de W elevados são agrupados em uma mesma categoria, o que explica a observação de alguns sítios selecionados positivamente nos genes GH5 e GH6, mesmo quando os demais testes de seleção são negativos. Apesar desses problemas, esse método de verossimilhança, utilizado em conjunto com todos os testes (Sítio, Ramo e Sítio-Ramo), é considerado o mais eficiente na detecção de seleção positiva (Suzuki & Nei 2004).

Para testar as relações filogenéticas entre os 3 genes do conjunto GH estudados neste trabalho, construímos a árvore baseada em MP. Para discutir esses resultados, é importante ter em mente os conceitos de genes ortólogos e parálogos. Ortólogos são genes que divergiram após um evento de especiação

enquanto parálogos são genes que divergiram após um evento de duplicação (Fitch, 1970). Como Fitch (1970) explicou: “*Where the homology is the result of gene duplication so that both copies have descended side by side during the history of an organism, (for example, alpha and beta hemoglobin) the genes should be called paralogous (para = in parallel). Where the homology is the result of speciation so that the history of the gene reflects the history of the species (for example alpha hemoglobin in man and mouse) the genes should be called orthologous (ortho = exact)*”. Ortólogos são genes similares quanto à sequência de DNA, com funções semelhantes em espécies diferentes; genes parálogos são aqueles resultantes de duplicação e possível diferenciação de função em relação ao gene ancestral. Um ponto importante é nos certificarmos de que o resultado da amplificação não está sendo contaminado pela paralogia. Ou seja, precisamos confirmar a ortologia dos genes verificando se eles formam grupos monofiléticos.

Analisando a árvore, observamos que cada gene forma um grupo distinto, fortemente suportado pelos valores de *bootstrap* de 99%. Além disso, em cada grupo evidenciamos a relação de similaridade entre os genes ortólogos. Esse resultado é semelhante ao encontrado por Revol de Mendoza *et al.* (2004), que mostrou a relação de ortologia entre esses genes para os gêneros *Callithrix* e *Ateles*. Nossos dados mostram que a similaridade entre genes ortólogos é maior do que entre os parálogos (figura 8) como esperado pela divergência após a duplicação (Ohta 1993).

Este resultado reflete um ou mais eventos de duplicação gênica que ocorreu no ancestral de todos os Platyrrhini após sua divergência dos macacos do Velho Mundo. Isso provavelmente também é um reflexo das diferentes pressões seletivas atuando sobre cada grupo, ou seja, o tipo de evolução

atuante em cada gene está sendo refletido no padrão filogenético da árvore. O único gene que recupera quase totalmente dois clados da filogenia dos primatas do Novo Mundo (Calitrichinae e Atelidae + Pitheciidae) é o gene GH5 que, segundo os testes apresentados anteriormente, é o único gene que evoluiu de forma neutra.

Conclusão e Considerações Finais

O presente trabalho apresentou uma proposta de associação dos padrões de diversificação morfológica a padrões de evolução molecular a fim de entender os processos de diversificação de tamanho nos Platyrrhini. Os resultados deste trabalho nos levam a concluir que:

- Os diferentes genes formadores do conjunto GH em Platyrrhini estão evoluindo sobre diferentes pressões seletivas: GH1 (seleção positiva), GH5 (Neutralidade) e GH6 (Seleção Purificadora).

- O gene GH1 apresenta sítios sob seleção positiva em vários gêneros.

- Existem sítios localizados na região de ligação com o receptor.

- A filogenia molecular não reflete as relações sistemáticas já razoavelmente conhecidas entre estes grupos.

- Nossos resultados confirmam a ortologia entre os genes.

Assim, a diversificação de tamanho é influenciada pela forte restrição evolutiva atuando diretamente sobre sítios funcionalmente importantes no gene, que faz parte da via metabólica responsável pelo crescimento. Esta conclusão é de fundamental importância e contribuirá substancialmente para o entendimento da história evolutiva deste grupo tão diverso. Cabe ressaltar que a abordagem realizada neste trabalho com a utilização de ambos dados morfométricos e moleculares para estudar os processos que levam à diversificação nos Neotrópicos tem caráter inédito. No entanto, mais análises da estrutura, função e expressão destes genes em primatas neotropicais são fundamentais para determinar o significado desta grande variação de pressão seletiva, além do conhecimento de se estes genes são diferencialmente

expressos no tempo e espaço durante o desenvolvimento. Estudos sobre as funções de cada gene formador deste conjunto são necessárias para elucidar a evolução desta família de genes neste grupo de primatas.

Resumo

As bases moleculares da diversidade fenotípica dentro de espécies são de grande interesse aos biólogos evolutivos porque a evolução morfológica adaptativa depende da seleção de variantes genéticas. Entretanto, há poucos exemplos da variação fenotípica cuja a base molecular é compreendida, especialmente entre animais vertebrados. Biólogos em geral concordam que o processo da seleção natural é a fonte predominante de diversificação morfológica. Tem sido documentado previamente que existe diversificação adaptativa da morfologia craniana entre os taxa mais elevados de Platyrrhini e também em espécies e em muitos gêneros, que aparentemente diversificaram na morfologia do crânio por seleção natural. O hormônio de crescimento (GH) é um hormônio multifuncional, produzido principalmente pela glândula pituitária de todos os animais vertebrados para regular o metabolismo e para promover o crescimento pos-natal linear. A evolução molecular do GH foi estudada extensivamente em um grande número espécies de vertebrados, incluindo primatas. A evolução rápida do gene do GH em primatas, especialmente em regiões funcionalmente importantes, sugere a seleção darwiniana positiva. Entretanto, o relaxamento da seleção purificadora depois de múltiplas duplicações do locus GH não podem ser descartadas. O objetivo deste estudo era investigar a evolução molecular do gene do GH em primatas neotropicais, tentando identificar diferenças potenciais e funcionais entre taxa. As análises de máxima verossimilhança deste trabalho mostraram que a relação dN/dS no gene GH1 de primatas de Neotropicais é diferente entre gêneros, demonstrando que a evolução do GH1 no Platyrrhini não é compatível com o modelo neutro da evolução molecular. Estes resultados sugerem que o gene GH1 está sobre seleção positiva em alguns sítios na proteína durante o processo de diversificação dos primatas de Neotropical.

Abstract

The molecular bases of phenotypic diversity within and between species are of great interest to evolutionary biologists because the adaptive morphological evolution depends on the selection of genetics variants. However, there are few examples of phenotypic variation whose molecular basis is understood, especially among vertebrates, while most biologists agree that the natural selection process is the predominant source of morphological diversification. It has been previously documented that there is adaptive diversification of cranial morphology among the higher taxa of Platyrrhini and also in species in most genera that apparently diversified in skull morphology by natural selection. The Growth Hormone (GH) is a multifunctional hormone, mainly produced by the pituitary gland of all vertebrates to modulate the metabolism and promote linear postnatal growth. The molecular evolution of GH has been extensively studied in a large number of vertebrate species, including primates. The rapid GH gene evolution on primates, especially on sites that are functionally important, suggest positive Darwinian selection. However, the hypothesis of purifying selection relaxation after multiple duplications of the GH locus cannot be ruled out either. The aim of this study was to investigate the molecular evolution of the GH gene on neotropical primates, trying to identify potential and functional differences between taxa. The analyses of Maximum likelihood of this work has shown that the ratio dN/dS in the GH1 gene of Neotropical primates are different among genera, demonstrating that the evolution of the GH1 in the Platyrrhini is not compatible with the neutral model of molecular evolution. These results suggest that the GH1 gene suffered positive selection at some sites inside the protein during the process of diversification of the Neotropical primates.

Referências Bibliográficas

- Agnieska, M. L. and Waters, M.J.2007. HoW groWth hormone controls groWth, obesity and sexual dimorphism. Trends in Gen. 24: 41-47.
- AhluWalia, A. 2004. Sexual dimorphism of rat liver gene expression: regulatory role of GH revealed by DNA microarray analysis. Mol. Endocrinol.18: 747-760.
- Anisinova, M., BielaWaski, J. P., Yang, Z. 2001. Accuracy and poWer of the likelihood ratio in detecting adaptive molecular evolution. Mol.Biol. Evol. 18, 1585-1592.
- Barrera-Saldana H. A., Beck B. L., Esquivel-Escobedo D., Revol A. and Riggard K. 1998. Absence of Human placental lactogen and placental GroWth Hormone (HGHV) during Pregnancy: PCR analysis of deletion. Human Genetics. 102: 87-92.
- Bazan, J.F. Proc. Natl. Acad. Sci.USA. 1990. 87, 6934.
- Calder WA.1984. Size, function and life history. Cambridge: Harvard Uiniversity Press.
- Chen, E. Y., Liao, Y, C., Smith D, H., Barrera-Saldanha, H, A., Gelinas, R, E. , Seeburg, P, H. 1989. The human groWth hormone locus: nucleotide sequence, biolog and evolution. Genomics. 4: 479-497.
- Davey, H.W. 2001. STAT 5b is required for GH induced liver IGF-1 gene expression. Endocrinology 142: 3836-3841.
- De Vos, A. M. , Ultsch,M. Kossiakoff,A.. A. 1992. Human GroWth hormone and extracellular domain of its receptor: crystal structure of the complex. Science. 255: 306-312.

- Delson E, Terranova CJ, Jungers WL, Sargis EJ, Jablonski NG, Dechow PC. 2000. Body mass in cercopithecidae (Primates, Mammalia): estimation and scaling in extinct and extant taxa. *Anthropol Pap Am. Mus. Nat. Hist.* 83: 1-159.
- Fitch WM. 1970. Distinguishing Homologous from Analogous Proteins *Systematic Zoology* 19: 99-113.
- Forsyth, I. A, and Wallis, M. 2002. Growth hormone and prolactin molecular and functional evolution. *J. Mammary Gland Biol. Neoplasia* 7: 291-312.
- George D.I., Phillips J.A., Francke R.D.U and Seeburg P.H. 1981. The genes for growth hormone and chorionic somatomammotropin are the long arm of human chromosome 17 in region q21 to qter, *Human Genetics* 57 138-141..
- Godfrey LR, Sutherland MR, Petto AJ, Boy DS. 1990. Size, space and adaptation in some subfossil lemurs from Madagascar. *Am. J. Phy. Anthropology* 81: 45-66.
- Gollos G. T., Durning M., Fisher J. M. and Fowler P. D. 1993. Cloning of four Growth Hormone/Chorionic Somatomammotropin – Related Complementary Deoxyribonucleic Acid Differentially Expressed during Pregnancy in Rhesus Monkey Placenta. *Endocrinology.* 133(4) 1744-1752.
- Greenhalgh, C. J. 2002. Growth enhancement in SOCS-2 deficient mice is dependent on STAT5b. *Mol. Endocrinol.* 16: 1394-1406.
- Hall T. A. 1999. BioEdit: a user- friendly biological sequence alignment editor and analysis program for Windows 95/98 NT. *Nuc. Acids. Simp. Ser.* 41: 95-98.
- Harvey PH, Clutton-Brock TH. 1985. Life history variation in primates. *Evolution Int. J. Org. Evolution.* 124:801-810.

- Harvey, S. Azumaya, Y, Hull, K, L. 2000. Pituitary and extrapituitary growth hormone: Pit-1 dependence? *Can. J. Physiol.Pharmacol.* 78: 1013-1028.
- Honacki, J. H., Kinman, K. E., Koeppl, J. W. 1982. *Mammals species of the World: a taxonomic and geographic reference.* Allen Prees and the Association of Systematic Collections.
- Kofoed, E. M. 2003. GH insensitivity associated With STAT5b mutation.*N. Engl. J. Med.* 349: 1139-1147.
- Lacroix, M. C., Guibourdenche, J. L., Muller, F. and Evain-Brion, D. Human 2002. Placental Growth Hormone – A Review. *Placenta*, 23(16) S87-S94.
- Lanning, N. J. and Carter-Su, C. 2006. Recent advances in GH signaling. *Rev. Endocrinol. Metab. Disord.* 7: 225-235.
- Li W.H. 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J. Mol. Evolution.* 36: 96-99.
- Liu, J-C., Makova, K. D., Adkins, R. M and Gibson, S. 2001. Episodic Evolution of Growth Hormone in Primates and Emergence of the Species Specificity of Human Growth Hormone Receptor. *Mol. Biol. Evol.* 18(6): 945-953..
- Lupu, F.2001. Roles of gh and igf-1 in mouse postnatal growth. *Dev. Biol.* 229: 141-162.
- Marroig G., and Cheverud J.M. 2004. Did natural selection or genetic drift produce the cranial diversification of Neotropical monkeys? *Am. Nat.* 163: 417-428.
- Marroig G., Cheverud J.M 2005. Size as a line of least evolutionary resistance: Diet and adaptive morphological radiation in new World monkeys. *Evolution* 59(5) 1128- 1142.

- Marroig G. and Cheverud J.M. 2001. A comparison of phenotypic variation and covariation patterns and the role of phylogeny, ecology and ontogeny during cranial evolution of Neotropical World Monkeys. *Evolution* 55: 2576-2600.
- Ohta, T. 1993. The nearly neutral theory of molecular evolution. *Ann. Rev. Ecol. Syst.* 23: 263- 286.
- Revol de Mendoza. A., Escobedo. D. E., Dávila .I.M. e Saldaña. H.B. 2004. Expansion and divergence of the GH locus between spider monkey and chimpanzee. *Gene*. 336 185-193.
- Rinn, J. L. and Snyder, M. 2005. Sexual dimorphism in mammalian gene expression. *Trends Genet.* 21, 298-305.
- Rylands, A.B, Schneider, H., Langguth, R.A, Mittermeier, C.P. Groves, P. and Rodrigues-Luna, E. 2000. An assessment of the diversity of Neotropical World primates. *Neotropical Primates* 8: 61-93.
- Schneider H. 2000. The Current of the Neotropical World Monkey Phylogeny. *An. Acad. Bras. Ciências.*, 72(2) 165-172.
- Steppan. S. J., Adkins, R. M. and Anderson. J. 2004. Phylogeny and Divergence-Date Estimates of Rapid Radiations in Murine Rodents based on Multiple Nuclear Genes. *Syst. Biol.* 53(4):533-553.
- Takahashi, J. 1999. Characterization of hepatic CYP450 isozyme composition in the transgenic rat expressing low level of human GH. *Xenobiotica*. 29, 1203-1212.
- Tannenbaum, G. S. 2001. Temporal relationship between the sexually dimorphic spontaneous GH secretory profiles and hepatic STAT5 activity. *Endocrinology*. 142: 4595-4598.
- Teglund, S. 1998. STAT 5a and b proteins have essential and nonessential, or redundant, roles in cytokine responses. *Cell* 93: 841-850.

- Udy, G. B. 1997. Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression. *Proc. Natl.Acad.Sci.USA*. 94:7239-7244.
- Vidarsdottir,S. 2006. Clinical and biochemical characteristics of a male patient With a novel homozygous STAT5b mutation.*J. Clin. Endocrinol.Metab*.91: 3482-3485.
- Wallis, M. 1994. Variable evolutionary rates in the molecular evolution of mammalian growth hormones. *J. Mol.Evol*.38: 619-627.
- Wallis, M. 1996. The molecular evolution of vertebrate growth hormone: a pattern of near-stasis interrupted by sustained bursts of rapid change. *J. Mol. Evol*. 43: 93-100.
- Wallis, O. C. and Wallis, M. 2002.Characterisation of the GH gene cluster in a New-World monkey, the marmoset (*Callithrix jacchus*). *Journal of Molecular Endocrinology*. 29: 89-97.
- Waters, M. J. 2006. New insights into growth hormone action. *J. Mol. Endocrinol*. 36: 1-7.
- Waxman, D. J. and O'Connor, C. 2006. Growth hormone regulation of sex-dependent liver gene expression. *Mol.Endocrinol*. 20: 2613-2629.
- Woelfle, J. 2003. Acute control of IGF-1 gene transcription by GH through STATb. *J. Biol. Chem*. 278: 22696-22702.
- Woelfle, J. 2003. Mechanisms of GH action. Identification of conserved STAT5 binding sites that mediate GH induced IGF-1 gene activation. *J. Biol.Chem*. 278: 51261-51266.
- Yakar, S. 2002. The role of circulating IGF-1: lessons from human and animal models. *Endocrine* 19: 239-248

- Yang Z. 2000. Codon-Substitution Models for Heterogeneous Selection Pressure at Amino Acid Sites. *Genetics*. 155: 431- 449.
- Yang. Z. 1998. Likelihood Ratio Test for Detecting Positive Selection and Application to Primate Lysozyme Evolution. *Mol. Biol.Evol.* 15(5): 568 - 573.
- Yang. Z.. 1997. PAML: a program package for phylogenetic analysis by Maximum likelihood Computer Applications in Biosciences. 13: 555 - 556.
- Ye, C. Li, Y., Shoi, P. Zhang, Y. P. 2005. Molecular evolution of growth hormone gene family in old World monkeys and hominoids. *Gene*. 350: 183-192.
- Zelus, .D. Robinson-Rechavi, M. Delacre, M. Auriault, C. Laudet, V. 2000. Fast evolution of interleukin-2 in mammals and positive selection in ruminants. *J. Mol. Evol.* 51: 234-244.
- Zhu, T. 2001. Signal transduction via the growth hormone receptor. *Cell. Signs*. 13: 599-616.

Anexo

Espécies Utilizadas

Callithrix jaccus
Cebuella pygmaea
Saguinus bicolor
Leontopithecus chrisomelas
Callimico goeld
Saimiri sciureus
Cebus apella
Ateles paniscus
Brachyteles arachnoides
Lagothrix lagothricha
Alouatta belzebul
Aotus azarae
Callicebus moloch
Cacajao calvus
Chiropotes satanas
Pithecia pithecia