

ALEX SANDRO SCHIERHOLT

O GENE DA MIOGENINA:
SEQUENCIAMENTO EM SUÍNOS E ANÁLISE FILOGENÉTICA

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de "Magister Scientiae".

VIÇOSA
MINAS GERAIS - BRASIL
2005

ALEX SANDRO SCHIERHOLT

O GENE DA MIOGENINA:
SEQUENCIAMENTO EM SUÍNOS E ANÁLISE FILOGENÉTICA

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de "Magister Scientiae".

Aprovada: 25 de fevereiro de 2005.

Prof. Paulo Sávio Lopes
(Conselheiro)

Prof. Ricardo Frederico Euclides
(Conselheiro)

Prof. Théa Miriam Machado Medeiros

Prof. Aldrin Vieira Pires

Prof. Simone Eliza Facioni Guimarães
(Orientadora)

AGRADECIMENTOS

A Deus, pelas oportunidades, encontros e desencontros que definem minha vida.

À Universidade Federal de Viçosa e ao Departamento de Zootecnia, pela oportunidade de realização do curso.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pela concessão da bolsa de estudos.

À professora Simone pelos ensinamentos, pela orientação, pela amizade, pela confiança e paciência, minha admiração e gratidão.

Ao professor Ricardo Frederico Euclides, pelos aconselhamentos e incentivo, pelas preocupações, conversas e bate papos.

Ao professor Paulo Sávio Lopes, pelo aconselhamento, pelo incentivo, pela amizade e pelo apoio.

Ao professor Robledo, pela amizade e feijoadas da vida.

Ao Samuel por me apontar a direção e pelas valiosas sugestões que contribuíram em muito para o início e que foram fundamentais durante todo este trabalho e pela amizade.

À Dani por suas horas ocupadas com os sequenciamentos tão bem feitos e indispensáveis para a realização desse trabalho e pela amizade

Aos meus pais, Carlos e Nair, pela confiança depositada em mim, a liberdade vigilante, e pelo exemplo de uma vida cheia lutas vitoriosas.

Ao meu irmão, Sergio, por todos os conselhos e conversas de botecos jogadas fora.

À linda Luciara, um presente em minha vida, pelo companheirismo, risos e lágrimas e por encher a minha vida com poesias e sensações maravilhosas.

Aos meus amigos de repúblicas durante minha vida em Viçosa: Breno, Guiverme, Wally, Bujão, Ronaldo, Leandro, Lívio, Cláudio e Patrick.

Aos meus colegas do melhoramento Adriana, André, Aldrin, Dani Serra, Débora, Elizângela, Fábio, Fernanda, Fred, Gisele, Guilherme, Gustavo, Jane,

Joãozinho, Júnior, Kécya, Kleibe, Lindenberg, Marcelinho, Mocarongo, Peloso, Rachel, Ricardinho, Rodolphinho, Zé Pereira, 2mi2.

A todos os colegas que passaram pelo laboratório, pela compreensão pelo uso do termociclador nos momentos de aperto.

À turma da Zôo 98, especialmente àqueles que conseguiram transformar coleguismo em amizade. Você saberá se é um deles.

Aos amigos Baby, Boi, Breno, Bruno Curinga, Bruno Japa, Carlão, Danilão, Fábio e Felipe, Fred, Guilherme Japa, Marcelo Bunda, Insetinho, Ítalo, Josele, Ju, Leidi, Léo, Lilão, Lívio, Luíza, Patbeijo, Pedrita, Rato, Raquel, Valdete, Vivianizinha, Wally.

CONTEÚDO

RESUMO.....	vi
ABSTRACT.....	vii
INTRODUÇÃO.....	1
REVISÃO DE LITERATURA.....	2
REFERÊNCIAS BIBLIOGRÁFICAS.....	19

CAPÍTULO 1

Análise do sequenciamento do gene da Miogenina em suínos da raça nativa piau e de linhagens comerciais

Resumo.....	23
Abstract.....	24
Introdução.....	25
Material e Métodos.....	27
Resultados e Discussão.....	33
Conclusões.....	41
Referências Bibliográficas.....	42

CAPÍTULO 2

Análise filogenética do gene da Miogenina em animais com o gene sequenciado

Resumo.....	43
Abstract.....	44
Introdução.....	45
Material e Métodos.....	47

Resultados e Discussão.....	51
Conclusões.....	65
Referências Bibliográficas.....	66

RESUMO

SCHIERHOLT, Alex Sandro, M.S., Universidade Federal de Viçosa, fevereiro de 2005.

O Gene da Miogenina: Sequenciamento em Suínos e Análise Filogenética.

Orientadora: Simone Eliza Facioni Guimarães. Conselheiros: Robledo de Almeida Torres e Paulo Sávio Lopes.

A capacidade de produção de carne está relacionada com o número de fibras musculares apresentadas pelos animais recém nascidos. As fibras musculares são formadas na miogênese que ocorre durante o desenvolvimento embrionário, evento que é em parte controlado pela família de genes MyoD. O gene da Miogenina é membro dessa família e regula a expressão de genes músculo específicos. Ele desempenha uma função importante porque quando é ativado, as células do tecido muscular param de se multiplicar e o tecido passa a crescer apenas em volume. É considerado então um gene candidato. Os objetivos desse trabalho foram investigar possíveis alterações nas seqüências de nucleotídeos no gene da Miogenina em suínos de raças divergentes, e comparar o sequenciamento obtido em suínos com outras espécies que possuem o gene depositado no Genbank, a fim de estudar a história evolutiva do gene nestas espécies. A estratégia usada foi a divisão do gene em sete fragmentos que continham os três exons, os dois introns e as regiões 3' e 5' do gene. Foram seqüenciados dois varrões da raça nativa Piau e 12 matrizes comerciais. Os resultados obtidos mostraram um gene altamente conservado, com nenhum polimorfismo nas regiões exônicas. O sequenciamento dos introns não apresentou resultados satisfatórios. Métodos de avaliação filogenética também comprovaram o estado conservado do gene. Foram comparadas seqüências das espécies *Sus scrofa* (suínos), *Bos taurus* (bovinos), *Ovis aries* (ovinos), *Homo sapiens* (humanos), *Mus musculus* (camundongos), *Rattus norvegicus* (ratos), *Gallus gallus* (galinhas) e *Meleagris gallopavo* (perus), verificando as taxas de substituição de nucleotídeos não sinônimas (Ka) pela taxa de substituições sinônimas (Ks) nos sítios. Os resultados indicaram que provavelmente o gene tenha sofrido uma evolução adaptativa nas espécies do grupo Ruminantia (*B. taurus* e *O. aries*) após as duas espécies terem divergido do seu ancestral comum. Nas outras espécies, o gene parece estar evoluindo de maneira conservativa.

ABSTRACT

SCHIERHOLT, Alex Sandro, M.S., Universidade Federal de Viçosa, February 2005.
The Myogenin Gene: Sequencing in Swine and Phylogenetic Analysis.
Committee members: Simone Eliza Facioni Guimarães, Robledo de Almeida Torres
and Paulo Sávio Lopes.

Meat production capacity is related to muscle fibers number show in newborn animals. The muscle fibers are formed in the myogenesis that takes place during the embryonic development, an event that is controlled in part by the MyoD gene family. The Myogenin gene is a member of this family and rules the expression of muscle specific genes. It has an important because when it is activated, the cells from the muscle tissue stop to multiply and it will be growing only in volume. It is considered than a candidate gene. The objectives of this study were to verify possible alterations in the nucleotide sequences of the Myogenin gene in swine from divergent breeds, and compare the sequence obtained in swine with others species that have this data in the Genbank, to study the evolutionary story of the gene in these species. The strategy used was to divide the gene in seven fragments that contain the three exons, two introns and the 3' and 5' regions of the gene. Two boars of the native breed Piau and 12 commercial sows were sequenced. The results showed a highly conserved gene, with no polymorphism in the exons regions. The introns sequencing did not shown satisfactory results. Phylogenetic evaluation methods also prove the conservatory state of the gene. It was compared sequences from the species *Sus scrofa* (swine), *Bos Taurus* (cows), *Ovis aries* (sheeps), *Homo sapiens* (humans), *Mus musculus* (mice), *Rattus norvegicus* (rats), *Gallus gallus* (chickens) and *Meleagris gallopavo* (turkeys), verifying the nonsynonymous nucleotide substitution rate by the synonymous substitution rate in the sites. The results point out that probably the gene suffered an adaptive evolution in the Ruminantia group (*B. Taurus* and *O. aries*) after these species diverged from their common ancestral. In the other species, the gene seems to be evolving in a conservative way.

1 - INTRODUÇÃO

O uso de dados moleculares em programas de melhoramento genético auxilia o melhorista a alcançar o objetivo desejado ao fazer a seleção dos animais. Essas vantagens vão desde a seleção precoce de animais, por meio da análise de genes antes que ocorra a sua expressão, ou na seleção de características que são mensuradas apenas após a morte do animal. Esses dados podem ser usados como efeitos fixos em metodologia de modelos mistos, como o BLUP, para que a pressuposição de que a característica seja controlada por genes com efeitos infinitesimais não seja quebrada, como para eliminar um alelo indesejável na população.

A descoberta de um gene candidato depende de uma série de fatores para que o gene possa vir a ser implementado em programas de melhoramento. Primeiramente, o gene deve apresentar um ou mais polimorfismos nos animais parentais que foram usados no delineamento das famílias experimentais, em seguida deve-se verificar se essa mutação é conservativa ou não, ou seja, se a cadeia de aminoácidos codificada pela seqüência de nucleotídeos é diferente nos animais que apresentaram os polimorfismos. O tipo de substituição de aminoácidos também deve ser considerado, já que os aminoácidos podem ter características físico-químicas semelhantes, como as suas estruturas atômicas e cargas residuais semelhantes ou equivalentes. Por último deve-se fazer uma associação estatística com o objetivo de verificar se esse polimorfismo teria um efeito significativo sobre as características de interesse para programas de melhoramento.

Diante desses fatos, percebe-se a necessidade de estudos exaustivos na identificação de genes candidatos que possam vir a ter efeitos em características de interesse econômico. No presente estudo é feita a avaliação do gene da Miogenina como gene candidato por meio de sequenciamento em animais da raça naturalizada Piau e em animais de linhagem comercial.

A fim de verificar o nível de conservação do gene foi realizado um estudo filogenético da Miogenina usando a espécie suína e outras espécies que possuíam informações sobre o gene depositadas no Genbank, juntamente com a seqüência obtida usando os animais citados anteriormente.

2 - REVISÃO DE LITERATURA

2.1 - MARCADORES GENÉTICOS

Os marcadores genéticos são utilizados para marcar variações genéticas cuja expressão seja de difícil identificação. Nesse caso, seleciona-se a variante de interesse de forma indireta. É essencial que o marcador seja herdável e de fácil avaliação, isto é, seu genótipo seja eficientemente identificado, o que equivale possuir um valor de herdabilidade próximo de 1,0. Outro ponto fundamental para que o marcador seja eficiente é que ele esteja intimamente ligado à variante gênica que se deseja selecionar, pois assim, eles tendem a permanecer ligados, e sempre que o indivíduo apresentar o marcador ele deverá também ser portador da variante gênica (Ramalho *et al.*, 2001).

Os marcadores morfológicos foram os primeiros tipos de marcadores genéticos utilizados na seleção. Por exemplo, o empenamento tardio das aves “marcava” o sexo feminino (Guimarães, 2001). Apesar de os marcadores morfológicos serem de fácil identificação existem limitações para o seu uso, pois o número reduzido desses não é suficiente para marcar todos os genes de interesse, além de haverem poucas formas múltiplas (Ramalho *et al.*, 2001). Pensando nisso, foram desenvolvidas técnicas para o uso de marcadores moleculares, como os marcadores bioquímicos e os marcadores de DNA.

Segundo Ferreira e Grattapaglia (1998), por marcador molecular define-se todo e qualquer fenótipo molecular, como no caso de isoenzimas, ou de segmento específico de DNA (correspondente a regiões expressas ou não do genoma).

O princípio básico dos marcadores isoenzimáticos reside na técnica de eletroforese em géis de amido e o método histoquímico de visualização do produto enzimático (Hunter e Markert, 1957). Ainda segundo Ferreira e Grattapaglia (1998), a premissa básica adotada, é que as diferenças na mobilidade de isoenzimas em um campo elétrico são resultantes de diferenças no DNA que codificam tais enzimas. A expressão de isoenzimas é co-dominante, ou seja, em um indivíduo diplóide ambos os alelos de um loco são expressos e visualizados. Ocasionalmente, subunidades de uma enzima podem ser codificadas por locos genéticos distintos, tornando mais complexa a

análise do padrão de bandas enzimáticas. Devido às isoenzimas possuírem função metabólica específica, os polimorfismos são limitados e o pequeno número de sistemas isoenzimáticos polimórficos não permite a cobertura completa do genoma. Apesar da técnica ser relativamente barata, esses fatos comprometem a capacidade de se encontrar associações significantes entre isoenzimas e genes que controlam caracteres de interesse (Ferreira e Grattapaglia, 1998).

O desenvolvimento tecnológico na área de marcadores de DNA tem sido extremamente rápido. As técnicas de DNA recombinante e o desenvolvimento da amplificação de segmentos de DNA por PCR (Polimerase Chain Reaction, ou reação da polimerase em cadeia) abriram o caminho para uma mudança no paradigma genético básico: da inferência do genótipo a partir do fenótipo, em que Mendel foi pioneiro, para a análise genética feita diretamente na variação na seqüência de DNA. Esta mudança de enfoque foi denominada transição da “genética Mendeliana” para a “genética genômica” (Ramalho *et al.*, 2001).

Os marcadores de DNA começaram a ser utilizados na década de 1980 e foram desenvolvidas mais de uma dezena de procedimentos. A maior vantagem do DNA é a grande variabilidade que se observa entre os indivíduos de uma espécie, o que equivale dizer que utilizando a análise da seqüência do DNA consegue-se teoricamente um número de marcadores suficientes para marcar todos os alelos de todos os genes da espécie (Ramalho *et al.*, 2001). Eles também apresentam grande estabilidade, podendo ser analisados anos depois do DNA ser coletado a partir de qualquer tecido orgânico que apresente células nucleadas, tal como sêmen, pele, sangue e folículos pilosos. Além disso, devido a sua grande plasticidade, reflete diretamente as relações filogenéticas entre diversos indivíduos, pois mesmo fragmentos de genes não transcritos são utilizados para análise (Guimarães, 2001).

Com os estudos filogenéticos, surgiram várias dificuldades como o desenvolvimento de conceitos, métodos, algoritmos e programas computacionais mais refinados e com exigência de processadores de performance elevada. O uso dos métodos filogenéticos e seus algoritmos permitem a reconstrução de estados ancestrais de caracteres e, dessa maneira, o estudo da evolução adaptativa desses caracteres.

Na história evolutiva, genes e suas regiões reguladoras estão sujeitos uma variedade de modificações, incluindo substituições de nucleotídeos, duplicação,

recombinação e eventos de reparo do DNA. A substituição nucleotídica em seqüências codificadoras é um mecanismo que gera um diferencial na evolução adaptativa das espécies. A maneira mais comum de identificar ramos de árvores filogenéticas que estejam sob evolução adaptativa é a medida do valor da razão da taxa de substituição não sinônima (onde a substituição de um nucleotídeo por outro gera uma mudança na cadeia de aminoácidos) pela taxa de substituição sinônima (onde essa substituição não gera uma mudança na seqüência de aminoácidos) que ocorre nos sítios dos códons em cada ramo da árvore (Liberles, 2005).

O aumento no número de informações em banco de dados genômicos e o desenvolvimento de ferramentas para analisar esses dados permitem estudos sobre a função de genes e proteínas e a sua importância para as espécies.

Os marcadores de DNA são usados, ainda, em programas de seleção onde são, geralmente, os ligados a QTL (*Quantitative Trait Loci* ou Locos de Característica Quantitativa), como os locos de microsatélites, chamados de indiretos, e os marcadores diretos, conhecidos como Genes Candidatos e *Major Genes* onde variações na seqüência do próprio gene de interesse são usadas como marcadores (Ferreira e Grattapaglia, 1998).

2.2 - GENE CANDIDATO

Os genes candidatos são genes seqüenciados de função biológica conhecida e que estão envolvidos com a manifestação da característica de interesse. Podem ser genes estruturais ou genes reguladores de uma via bioquímica que influencie a característica. Rothschild e Soller (1999) citam que a metodologia de estudo de genes candidatos não é cara e que apresenta muitas vantagens tais como: ampla aplicabilidade, baixo custo, simplicidade operacional e conveniente aplicação a MAS (*Marker Assisted Selection*, ou seleção assistida por marcadores). Como principais passos no estudo de genes candidatos, os autores enumeram: escolha dos genes, obtenção dos *primers* para amplificar o gene, busca dos polimorfismos, desenvolvimento de técnicas práticas de genotipagem em larga escala, busca de associações entre o gene candidato e o fenótipo da característica.

Pesquisas indicam que quando um gene candidato é identificado, o programa ótimo seria aquele que não fixasse o gene rapidamente, pois, neste caso, muitos animais portadores de alelos favoráveis em outros locos seriam descartados. Um índice ótimo a ser usado seria aquele em que o peso dos genes candidatos fosse proporcional ao valor genético do gene e sua correlação com o valor genético total do indivíduo. O peso dado ao gene candidato deveria considerar a proporção da variação genética determinada pelo QTL. Contudo, este peso deverá ser aproximado, uma vez que os parâmetros genéticos associados com o gene candidato mudam rapidamente com a mudança da frequência gênica (Ledur e Schmidt, 2000).

2.3 - A MIOGÊNESE, A FAMÍLIA MYOD E A MIOGENINA

O principal componente da carne é o tecido muscular estriado esquelético. Esse tecido é formado por tecido conjuntivo, vasos sanguíneos, fibras nervosas, fluídos extracelulares e fibras musculares. As fibras musculares são as unidades fundamentais do músculo, representando entre 75 e 92% do volume muscular. As fibras são agrupadas em feixes que reunidos formam um grupo muscular. O aumento no diâmetro (hipertrofia) ou no número das fibras musculares (hiperplasia) resulta em aumento da massa muscular, sendo que na fase pós natal, quase todo aumento provém do fenômeno de hipertrofia (Coutinho *et al.*, 1999). O peso de carne magra e a taxa de crescimento em suínos são correlacionados positivamente com o peso ao nascimento dos leitões. Já o peso dos leitões depende, em parte, da quantidade de miofibras produzidas durante a miogênese que ocorre no desenvolvimento embrionário (te Pas *et al.*, 1999).

O processo de miogênese pode ser dividido em duas fases, a determinação e a diferenciação. A determinação é o evento em que as células pluripotentes que estão se multiplicando são mobilizadas para o processo miogênico, se transformando em mioblastos. Os mioblastos continuam se dividindo, mas agora já destinados a participar da linhagem muscular. A diferenciação ocorre quando os mioblastos se unem para formarem células multinucleadas, chamadas de miotúbulos (te Pas *et al.*, 2000). Isso ocorre quando os genes dos mioblastos iniciam sua expressão músculo-específica e as

células, a partir daí, perdem a capacidade de se dividir e as fibras musculares passam a crescer quase que somente em volume (hipertrofia).

Esse processo de miogênese é controlado por uma família de genes regulatórios chamada MyoD, composta por quatro genes: MyoD1, Myf-5, Miogenina e Myf-6 ou MRF4. Cada um desses genes transcreve um fator de transcrição que regula a expressão dos genes durante o processo da miogênese. MyoD1 e Myf-5 podem ser considerados fatores de determinação ou de especificação, já que eles estão envolvidos na proliferação dos mioblastos. Já a miogenina pode ser considerado um fator de diferenciação, pois está associada à fusão dos mioblastos mononucleados em miofibras multinucleadas, e o Myf-6 é expresso principalmente na fase pós-natal (Bergstrom e Tapscoot, 2001).

A miogenina tem então um papel importante durante a miogênese e sua expressão marca o final da proliferação dos mioblastos. Conseqüentemente, quando o gene da Miogenina é expresso, as fibras musculares se desenvolvem dos mioblastos que foram previamente formados (Coutinho *et al.*, 1999).

Os genes da família de proteínas MyoD possuem estruturas semelhantes e parecem ser capazes de se substituírem com um certo grau de eficiência. Na verdade, segundo Coutinho *et al.* (1999), os fatores regulatórios da família MyoD são bem relacionados e possuem homologia em dois domínios: o HLH e em uma região que possui aminoácidos carregados positivamente, o domínio básico.

A transferência de qualquer um desses genes miogênicos para um grande número de células de cultura convertem essas células em músculos. A expressão do MyoD1 leva à expressão da Miogenina, e a expressão do gene da Miogenina ativa a expressão do MyoD1. Dessa maneira, há um *feedback* positivo e recíproco, de maneira que quando tanto a Miogenina ou o MyoD1 são ativados, também são ativados os outros genes. Enquanto que o MyoD e o Myf-5 podem se substituir, não parecem ter redundâncias com as funções da miogenina, já que, em camundongos que tiveram o gene mutado em homozigose, não ocorre a diferenciação das células em miotúbulos (Gilbert, 2003).

A família de genes MyoD codifica proteínas regulatórias do tipo hélice-alça-hélice básico, chamadas de bHLH (*Myogenic basic helix-loop-helix*), que se ligam aos E-box (CANNTG), nas regiões regulatórias de um grande número de genes que se

expressam no músculo esquelético, estimulando a sua transcrição. Dessa maneira, mecanismos que regulem a expressão ou a atividade das proteínas miogênicas bHLH servem como pontos-chaves de regulação na diferenciação muscular (Coutinho *et al.*, 1999).

Várias vias pós-traducionais têm sido propostas para regular a atividade da proteína bHLH. Por exemplo, dimerização com outros tipos de proteínas bHLH, como E12 ou E47, levando à formação de heterodímeros com atividade transcricional. Já a dimerização com proteínas HLH desprovidas da região básica, como as proteínas Id, geram dímeros inativos. A fosforilação de proteínas miogênicas bHLH também regula a sua atividade, por exemplo, a inibição mediada por PKC (proteína quinase C) requer a fosforilação de um único resíduo de treonina, T87, na sua região básica (Blagden *et al.*, 2004).

O gene da Miogenina em suínos está localizada no cromossomo 9 na posição 9q2.1-q2.6 (Fig. 1), sendo formada de 3 exons e 2 introns (Ernst *et al.*, 1998). O mRNA maduro em suínos possui 675 nucleotídeos (exon 1 com 471 bases nitrogenadas, o exon 2 com 82 e o exon 3 com 122 bases nitrogenadas), codificando uma proteína madura composta de 224 aminoácidos, mais o códon terminador. No exon 1 está contido o motivo bHLH com seus dois domínios funcionais, o domínio básico e o domínio HLH. O domínio básico é rico em aminoácidos com carga residual positiva e é o sítio que se liga ao DNA. Já o domínio HLH é responsável pela ligação a outros HLH formando um dímero funcional (Coutinho *et al.*, 1999).

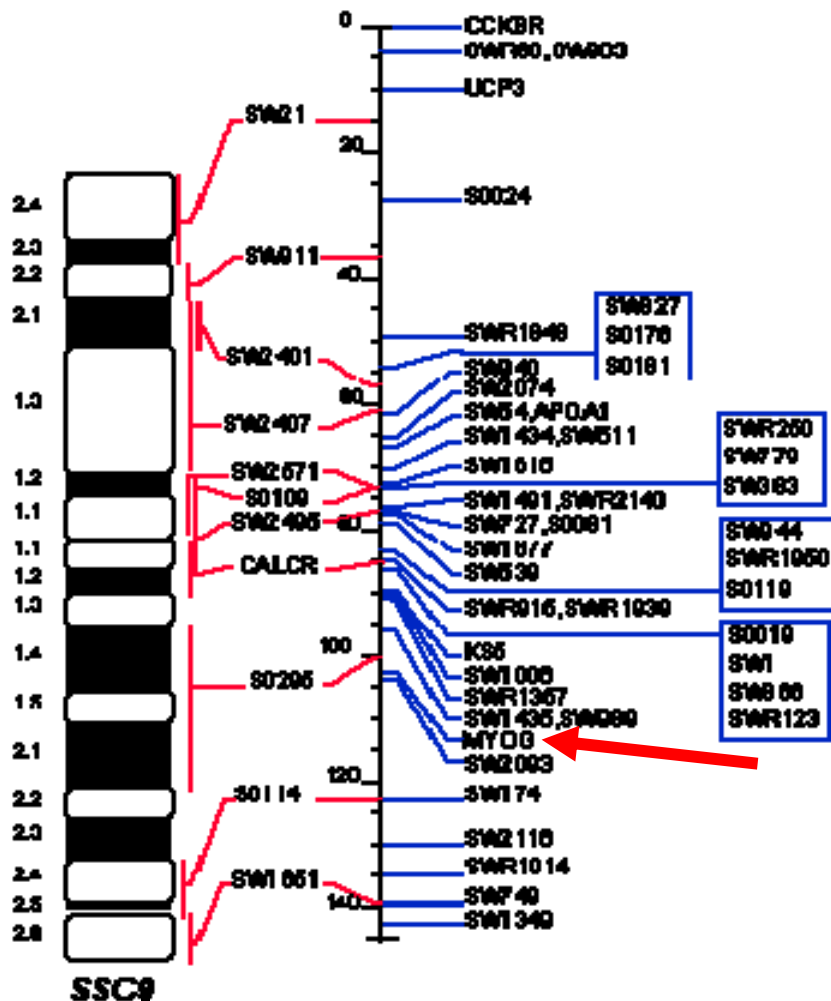


Figura 1 – Representação do cromossomo 9 de suínos e a localização do gene Miogenina. FONTE: Adaptado de <http://www.genome.iastate.edu/maps/marcmmap.htm> (acesso em 15/03/2003)

2.4 - A MIOGENINA COMO GENE CANDIDATO

Quando a miogenina é expressa, as fibras musculares serão geradas a partir dos mioblastos previamente formados na fase de determinação da miogênese. Dessa maneira, diferenças nas funções do gene da Miogenina ou o tempo da sua expressão poderiam ter grande influência no número de fibras musculares que são desenvolvidas durante a miogênese (te Pas *et al.*, 2000).

Hasty *et al.* (1993) desenvolveram um trabalho com camundongos em que inativavam um dos alelos do gene da Miogenina, de maneira que apenas um dos alelos

fosse funcional, e o resultado que obtiveram foram camundongos que apresentavam metade do número de fibras musculares quando comparados com os camundongos selvagens.

te Pas *et al.*, (1999), estudando o gene da Miogenina em suínos da raça Yorkshire usados em linhas comerciais, indentificaram polimorfismos usando a enzima de restrição *MspI*. Apesar desses sítios de restrição se encontrarem na região 3' do gene da Miogenina, e não em regiões codificadoras do gene, os autores encontraram associações estatísticas para peso ao nascimento, peso de carcaça, taxa de crescimento e conteúdo de carne magra para um dos genótipos, que apresentou uma maior produtividade para as características significativamente diferenciadas do outro genótipo. Os autores justificaram que um possível mecanismo para explicar essas diferenças seria que a mutação está localizada na região reguladora do gene, o que poderia ter afetado o tempo e o grau de expressão do gene da Miogenina. Dessa maneira, diferentes funções ou o tempo de expressão do gene poderia ter uma grande influência no desenvolvimento de fibras musculares durante a miogênese.

2.5 - CONCEITOS FILOGENÉTICOS

Segundo Miyaki (2001), o conceito de filogenia surgiu com Darwin, juntamente com o próprio conceito de ancestralidade entre espécies. Assim, as filogenias nada mais são que indicações das relações de ancestralidade supostas para um conjunto de espécies ou de outras unidades taxonômicas. Na prática, o problema fundamental é que muitas das espécies ancestrais não podem ser observadas no presente. Assim, é necessário buscar mecanismos para recuperar as informações a respeito das relações de parentesco entre os grupos analisando os organismos atuais.

A reconstrução filogenética consiste em estimar as relações de ancestralidade para um determinado grupo de organismos (táxons). Uma árvore filogenética representa a história evolutiva dos organismos nela representados. Graficamente, consiste de pontos (nós) ligados por linhas (ramos). Os táxons podem ser famílias, gêneros, espécies ou populações. Com freqüência, são chamados de OTUs (do inglês, *operational taxonomic units*, unidade taxonômica operacional), que representam o nível

taxonômico de amostragem usado pelo pesquisador em uma análise. Os nós externos de uma árvore representam as OTUs estudadas, unidas por ramos cujo nó interno representa o ancestral comum desses táxons (Miyaki, 2001).

Caracteres plesiomórficos ou primitivos e caracteres apomórficos ou derivados são conceitos que indicam a idade relativa de caracteres que são homólogos (possuem semelhanças devidas à herança de um ancestral comum) em uma determinada série de transformações. O estado apomórfico é o mais recente do caráter, que evoluiu por modificações a partir da condição pré-existente, dita plesiomórfica. Desse modo, a apomorfia aparece apenas em um subconjunto do grupo sob estudo, correspondente ao conjunto de populações ou espécies descendentes da ancestral na qual a modificação surgiu. Em uma série de transformações, a condição plesiomórfica é a mais antiga, que é modificada dando origem a outra mais recente (apomórfica). Um outro conceito que a escola filogenética introduziu na sistemática é o de grupos externos. Assim, para afirmar qual condição estudada em um grupo é a mais recente, é necessário saber qual é a mais antiga. Mas essa compreensão da direção da mudança de um caráter em um grupo, no entanto, não pode ser determinada apenas considerando a variação dentro do grupo. Ou seja, é preciso estabelecer grupos externos capazes de dar o sentido do tempo na análise do grupo de interesse (Schneider, 2003).

2.6 - ALGORÍTMOS PARA RECONSTRUÇÃO FILOGENÉTICA

Para a obtenção de uma árvore filogenética existem dois grupos básicos de algoritmos. No primeiro grupo, o princípio do método está embutido no próprio algoritmo que resulta em uma árvore final. No segundo grupo, o princípio do método é o critério usado para a escolha da melhor árvore dentro de um conjunto delas. Tradicionalmente, o primeiro grupo é mais frequentemente aplicado em métodos de distância genética, enquanto os critérios de máxima parcimônia e máxima verossimilhança utilizam algoritmos com base no segundo grupo de métodos (Miyaki, 2001).

No caso do segundo grupo de métodos, o conjunto de árvores pode conter todas as árvores possíveis (algoritmos exatos), em que o método necessariamente

encontrará a árvore (ou as árvores) que satisfaça(m) o critério de otimização. Uma alternativa é procurar uma árvore dentro de um subconjunto de árvores (algoritmos eurísticos) e neste caso, o pesquisador deve ter em mente que pode estar lidando com uma árvore sub ótima para aquele critério escolhido. Por outro lado, recentemente foi demonstrado que algoritmos eurísticos, mais simples e mais rápidos, podem ser tão ou ainda mais eficientes do que algoritmos mais complexos (Miyaki, 2001).

2.7 - MÉTODOS USADOS NA ANÁLISE FILOGENÉTICA

2.7.1 - MÉTODO DE MÁXIMA PARCIMÔNIA

O princípio de máxima parcimônia pode ser assim entendido: se existem diversas explicações para uma determinada observação, adota-se a mais simples, ou seja, a que requer o menor número de pressuposições. Se esse método for encarado sob uma perspectiva probabilística, ela se baseia em um modelo de evolução (implícito) em que uma mudança é mais provável do que duas, ou seja, trata substituições independentes gerando o mesmo resultado como um evento relativamente raro e que, se os táxons compartilham uma característica comum, é porque essa foi herdada de um ancestral comum, ou seja, são homólogas (Miyaki, 2001).

Para seqüências de nucleotídeos e aminoácidos os dados são os alinhamentos de suas seqüências. Cada sítio do alinhamento é um caractere, e cada caractere pode ter diferentes estados em diferentes proporções. Nem todos os caracteres são úteis na construção da árvore parcimoniosa. Caracteres invariantes, aqueles que estão no mesmo estado em todos os táxons, são obviamente inúteis e são ignorados pelo método. Também são ignorados os caracteres em que o estado diferente ocorra em somente uma OTU (Hall, 2001).

Um algoritmo é usado para determinar o número mínimo de passos necessários para que qualquer árvore (qualquer ordem dos ramos) seja consistente com os dados. Esse número é a pontuação da árvore e, a árvore ou as árvores com a menor pontuação, ou seja, onde o menor número de eventos evolutivos ocorreu (a menor soma do comprimento dos braços) serão consideradas as mais parcimoniosas, (Hall, 2001).

2.7.2 - MÉTODO DE MÁXIMA VEROSSIMILHANÇA

O princípio básico do método de máxima verossimilhança consiste em estimar a probabilidade de um conjunto de dados estarem representando um processo que realmente ocorreu, com base em um determinado modelo. Em termos de evolução de seqüências de DNA, o método irá calcular a probabilidade de que aquelas seqüências em estudo tenham sido geradas seguindo premissas do modelo evolutivo escolhido. Nesse caso, a topologia e o comprimento dos ramos de uma árvore são os parâmetros a serem estimados, dadas as seqüências finais nas extremidades dos ramos (Pereira, 2001).

A probabilidade deve ser calculada para todas as topologias possíveis, variando o tamanho dos ramos para um grupo de unidades taxonômicas operacionais (genes, espécies, gêneros, populações ou qualquer outra unidade evolutiva). A árvore (isto é, a topologia mais o comprimento dos ramos) que apresentar a maior verossimilhança (probabilidade) é considerada a melhor estimativa da filogenia (Pereira, 2001).

Calcular a probabilidade de uma árvore envolve o cálculo das probabilidades de ocorrência de todos os possíveis estados ancestrais de caracteres nos nós internos da árvore, isto é, calcular a possível ocorrência de cada um dos nucleotídeos terem estado presente em um nó interno. A maioria dos modelos evolutivos para seqüências de DNA admite reversão de caracteres ao longo do tempo, visto que a mudança de nucleotídeos é independente da mutação reversa.

a-) **Modelo de Substituição (Jukes e Cantor, 1969, ou JC69)**

Os modelos de substituição referem-se àqueles que envolvem a utilização de parâmetros que regem a evolução das seqüências. Jukes e Cantor (1969) propuseram um modelo bastante simples, em que os nucleotídeos A, C, G e T em uma seqüência de DNA ocorrem em freqüências iguais e a probabilidade de substituição de um nucleotídeo i para um nucleotídeo j (P_{ij}) em um intervalo de tempo (dt) depende simplesmente da taxa de substituição u , onde $i \neq j$.

$$P_{ij}(dt) = udt$$

b-) Modelo de Dois Parâmetros (Kimura, 1980, ou K80)

Substituições do tipo transições (entre pirimidinas ou entre purinas) ocorrem mais freqüentemente do que as transversões (entre uma pirimidina e uma purina, ou vice-versa). Kimura (1980) desenvolveu um modelo em que as taxas de transição α e as taxas de transversão β são levadas em consideração, e são apresentadas a seguir:

$P_{ij}(dt) =$	αdt para transição
	βdt para transversão

c-) Modelo Proporcional (Felsenstein, 1981 ou F81)

As freqüências dos quatro nucleotídeos nem sempre é similar. Assim, Felsenstein (1981) elaborou o seguinte modelo para calcular a probabilidade de substituição levando em consideração a freqüência de cada uma das bases:

$$P_{ij}(dt) = u \pi_j dt$$

Em que π_j representa a freqüência do nucleotídeo j .

d-) Modelos HKY85 e F84

Os modelos K80 e F81 foram combinados concomitantemente por Hasegawa (Hasegawa *et al.*, 1984, 1985) e por Felsenstein, em seu pacote de programas Phylip versão 2.6 e posteriores (Felsenstein, 1995), em outros dois modelos, denominados HKY85 e F84, citados por Pereira (2001).

Matematicamente HKY85 é representado por:

$P_{ij}(dt) =$	$A \pi_j dt$ para transição
	$\beta \pi_j dt$ para transversão

A diferença entre esse modelo e o F84 é o emprego de um parâmetro κ , que determina a razão transição/transversão, e do parâmetro Π_j . Este último parâmetro é calculado pela soma das freqüências: $\pi_C + \pi_T$, se o nucleotídeo j for uma pirimidina, ou

$\pi_A + \pi_G$, se j for um purina. Sua representação matemática é:

$P_{ij}(dt) =$	$(\kappa / \Pi_j + 1) u \pi_j dt$ para transição
	$u \pi_j dt$ para transversão

e-) Modelo TN93

A diferença na composição de bases promove modificações não apenas na frequência de transversões e transições, mas também na frequência de transições entre pirimidinas e na frequência de transição entre purinas. Para adequar essa diferença a modelos de evolução de seqüências, Tamura e Nei (1993), elaboraram o seguinte modelo:

$P_{ij}(dt) =$	$\alpha_R \pi_j dt$ para transição entre purinas
	$\alpha_Y \pi_j dt$ para transição entre pirimidinas
	$\beta \pi_j dt$ para transversão

em que α_R e α_Y representam, respectivamente, a taxa de transição entre purinas e a taxa de transição entre pirimidinas.

2.8 - RECONSTRUÇÃO DOS ESTADOS ANCESTRAIS

Parcimônia e máxima verossimilhança são duas estimativas que são usadas e usualmente comparadas entre si. Elas freqüentemente dão resultados similares e quando não, podem complementar uma a outra. Máxima verossimilhança funciona bem quando um bom modelo é disponível. A parcimônia trabalha bem quando não existe um modelo operacional adequado, ou não pode existir, como no caso de processos complexos, e também quando se trata de ramos muito curtos em que a probabilidade de terem ocorrido eventos múltiplos é extremamente rara (Liberles, 2004).

Ambos os métodos podem ser usados para estimar os ancestrais em uma árvore filogenética. Em estudos feitos com dados simulados, a reconstrução das seqüências de ancestrais não foi sempre acurada (até mesmo quando foram usados critérios de máxima

parcimônia), especialmente quando o tamanho dos ramos sob estudo aumenta, já que a probabilidade de ocorrerem substituições múltiplas é maior quanto maior for o tempo de divergência entre dois animais (Koshi e Goldstein, 1996; Zhang e Nei, 1997, citados por Liberles, 2001).

Fitch (1971) desenvolveu um algoritmo para reconstrução parcimoniosa de estados ancestrais de caracteres em uma árvore filogenética enraizada. Um aumento nas estimativas de máxima verossimilhança para determinar as seqüências dos ancestrais também tem sido desenvolvido (Yang, 1998; Yang e Nielsen, 2002). Reconstrução dos caracteres ancestrais baseada em Parcimônia é rápida e pode ser feita em aplicações genômicas em larga escala (Liberles, 2004).

Tanto a reconstrução das seqüências ancestrais por Parcimônia como por Máxima Verossimilhança podem ser usadas para estimar a evolução que ocorreu ao longo de um determinado ramo de uma árvore filogenética. Usando seqüências de reconstrução dos ancestrais, pode-se examinar as diferenças entre os nós conectados por um ramo. Isso reconstrói uma imagem da evolução que possivelmente ocorreu em qualquer dos ramos de interesse em uma árvore filogenética (Liberles, 2005).

2.9 - PRESSÃO DE SELEÇÃO

Sob um modelo de evolução neutra, a taxa de substituições nas posições nucleotídicas que possam alterar o aminoácido codificado (chamado de taxa de substituições nucleotídicas não sinônimas, k_a , ou dN) deve ser igual à taxa de substituições em posições nucleotídicas que não alteram o aminoácido codificado (taxa de substituição sinônima de nucleotídeos, K_s , ou dS). A maioria dos genes codificadores de proteínas em uma comparação de espécies proximamente relacionadas apresenta uma razão K_a/K_s significativamente menor do que 1, ou seja, o número de mutações não sinônimas presentes é muito superior ao número de mutações sinônimas, o que é um indicativo de seleção negativa ou que o gene está evoluindo de maneira conservativa. A maioria das proteínas foi selecionada, ao passar de milhões de anos de evolução, para determinadas funções. Dessa maneira, é mais provável que qualquer mutação não sinônima que ocorra, vá diminuir a adaptação do indivíduo e,

conseqüentemente, ser selecionada contra, reduzindo a taxa de substituição nesses sítios. No entanto, em alguns casos, podem ocorrer mutações 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, serem selecionadas ($Ka/Ks \gg 1$) tem-se a chamada seleção positiva. Porém, devido à maioria das mudanças de aminoácidos serem possivelmente deletérias, o valor de Ka é normalmente muito menor do que o de Ks (Fay *et al.*, 2001).

A pressão de seleção positiva, em que mudanças particulares dão ao organismo vantagem sobre outros organismos, pode rapidamente alterar as seqüências encontradas em sítios selecionados dentro do genoma. Ao mesmo tempo, um processo chamado de deriva genética, na qual as freqüências alélicas de uma população para determinados *loci* que contenham mutações desejáveis ou indesejáveis variam ao acaso ao longo das gerações, o que pode levar à fixação ou à extinção destas mutações. O estudo da combinação de seleção positiva e deriva genética, pode ser utilizado para avaliar o efeito de seleção no genoma e detectar qual nucleotídeo ou códon tem estado sobre diferentes tipos de seleção em diferentes pontos da história evolutiva (Liberles, 2002).

Segundo Liberles (2002), a maioria dos estudos moleculares de seqüências está focalizada na evolução das seqüências codificadoras. Isso se deve, obviamente, às várias funções dos polipeptídeos, que vão desde atividades reguladoras estruturais até em interações de superfície com solventes e outras atividades catalisadoras. Dessa maneira, em proteínas que estão tendo suas funções modificadas, provavelmente apenas um subconjunto desses resíduos (dependendo da natureza da modificação) estarão sob pressão de seleção positiva, enquanto que os restantes não. Dentre os principais mecanismos que tornam essas mutações funcionais destacam-se: mudanças na expressão genética, uso de *splicing* (retirada dos íntrons de um RNA precursor, de forma a produzir um mRNA maduro funcional) específico e duplicação gênica.

2.10 - ESTUDOS DE EVOLUÇÃO EM FAMÍLIAS DE GENES

Lisozima é uma enzima bacteriolítica que está presente em várias espécies. Os macacos *Colobine* são os únicos primatas com uma espécie de estômago modificado onde as bactérias realizam uma pré-digestão de vegetais, antes de estes seguirem para o estômago verdadeiro, rico em lisozima (Messier e Stewart, 1997). Os outros primatas têm apenas um estômago simples com uma anatomia diferente. Ao invés de apenas comparar a razão Ka/Ks do gene da Lisozima entre as espécies de macacos existentes, Messier e Stewart (1997) também calcularam as seqüências ancestrais em vários nós na árvore filogenética e calcularam a taxa Ka/Ks sobre os ramos da árvore. Eles chegaram à conclusão que não era apenas o ramo que levava aos macacos *Colobine* que estava sob pressão de seleção positiva durante esse período histórico (em que os *Colobine* se divergiram das outras espécies), mas também, inesperadamente, o ramo que leva aos Hominídeos. Os eventos que levaram a essa seleção positiva durante a emergência desse grupo não está esclarecida, mas podem estar relacionados com mudanças na dieta (Yang, 1998). A Leptina é outra proteína importante na dieta e parece ter estado sob pressão de seleção positiva nessa mesma época evolutiva (Benner, *et al.*, 1998).

A Leptina recentemente se tornou um gene de interesse farmacêutico, tendo envolvimento no controle da obesidade. Duas análises evolucionárias foram feitas na proteína da Leptina em mamíferos (Benner *et al.*, 1998; Benner *et al.*, 2000). Os dois estudos demonstraram episódios de rápida evolução de seqüência entre primatas e roedores, em que o gene da Leptina implicou em obesidade.

A coevolução intermolecular das proteínas também pode ser estudada filogeneticamente. Tanto a adaptação e a covariância compensatória podem explicar a evolução correlacionada dos resíduos de uma proteína. Um caso interessante é a interação entre a Leptina e o domínio extra celular do seu receptor nos primatas. Ambos mostram períodos de alta taxa Ka/Ks em vários ramos durante a diversificação dos primatas. As duas proteínas parecem estar evoluindo novos padrões de expressão durante esse período, em razão da ação de transposições nas regiões reguladoras de *splicing* e *enhancer* (seqüências de DNA que aumentam a afinidade da maquinaria de transcrição por um certo promotor) dos respectivos genes (Liberles, 2005).

A miostatina é a proteína causadora 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 da mutação no gene da Miostatina. Tellgren *et al.* (2004), ao estudar o gene da Miostatina no grupo monofilético Ruminantia por meio de parcimônia e um método padrão de contagem de sítios sinônimos e não sinônimos, observou que a linhagem que leva ao grupo Bovinae (boi e bisão) tinha uma significativa razão Ka/Ks de $1,17 \pm 0,01$, enquanto que o ramo que levava ao grupo Antilopinae (ovelhas, cabras e parentes próximos) uma razão de $1,00 \pm 0,01$ e o grupo Cephalo-caprini (grupo Antilopinae e parentes mais distantes desse grupo) tinha uma razão significativa de $1,42 \pm 0,01$ e (Tellgren *et al.*, 2004). Uma análise baseada em verossimilhança onde os resíduos Ka/Ks foram permitidos variar independentemente entre os ramos também foi feita por Tellgren *et al.* (2004), analisando o gene da Miostatina no grupo monofilético Ruminantia. Essa análise também mostrou os mesmos três ramos positivamente selecionados, levando ao Bovinae (Ka/Ks = 1,32), Antilopinae (Ka/Ks = 1,32) e ao Cephalo-caprini (Ka/Ks = 1,90). A partir destes resultados, o autor conclui que a pressão de seleção da Miostatina pode ter ocorrido devido a efeitos fenotípicos durante a época de divergência de Bovinae e Antilopinae. Uma hipótese é que a pressão de seleção positiva fez com que ocorresse um aumento na massa corporal e dos músculos esqueléticos, tais mudanças variaram, principalmente, de acordo com mudanças ambientais nos habitats das espécies (Tellgren *et al.*, 2004).

3 – REFERÊNCIAS BIBLIOGRÁFICAS

- BENNER, S.A.; TRABESINGER, N.; SCHEIBER, D. Converting primary sequences into physiological function. *Adv. Enzyme Regul.*, v.38, p.155-190, 1998.
- BENNER, S.A.; CHAMBERLIN, S.G.; LIBERLES, D.A.; GOVINDARAJAN, S.; KNECHT, L. Functional inferences from reconstructed evolutionary biology involving rectified databases: an evolutionary grounded approach to functional genomics. *Res. Microbial.*, v.151, p.97-106, 2000.
- BERGSTROM, D.A.; TAPSCOOT, S.J. Molecular distinction between specification and differentiation in the myogenic basic helix-loop-helix transcription factor family. *Molecular and Cellular Biology*, v.21 n.7, p.2404-2412, 2001.
- BLAGDEN, C.S.; FROMM, L.; BURDEN, S.J. Accelerated Response of the *myogenin* Gene to Denervation in Mutant Mice Lacking Phosphorylation of Myogenin at Threonine 87. *Molecular Cell Biology*, v 24(5), p.1983-1989, 2004.
- COUTINHO, L.L.; GABRIEL, J.E.; ALVARES, L.E. Controle molecular do desenvolvimento da musculatura esquelética. In: SIMPÓSIO INTERNACIONAL DE GENÉTICA E MELHORAMENTO ANIMAL, 1999, Viçosa. Anais...Viçosa: UFV, 1999. p.355-376.
- DE VRIES; SOSNICKI, A.A.; GARNIER, J.P. The role of major gene and DNA technology in selection for meat quality in pigs. *Meat Science*, v. 49, p.245-255, 1998.
- ERNST, C.W.; MENDEZ, A.; ROBIC, A. *et al.* Rapid communication: Myogenin (MYOG) Physically Maps to Porcine Chromosome 9q2.1-12.6. *J. Anim. Sci.*, v.76, p.328, 1998.
- FAY, J.C.; WYCKOFF, G.J.; WU, C. Positive and Negative Selection on the Human Genome. *Genetics*, v.158, p.1227-1234, 2001.
- FELSENSTEIN, J. Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, v.17, p.368-376, 1981.
- FELSENSTEIN, J. PHYLIP, Phylogenetic Inference Package, Versão 3.5c. Washington University. Programa distribuído pelo autor.
- FERREIRA, M.E.; GRATTAPAGLIA, D. Introdução ao Uso de Marcadores Moleculares em Análise Genética. 3. Ed. Brasília: Embrapa-Cenargem, 1998. p.220.

- FITCH, W.M. Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Zoology*, v.20, p.406-416, 1971.
- GILBERT, S.F. Myogenic bHLH Proteins and their Regulators. *Developmental Biology* 7^a ed. Disponível em: <http://www.devbio.com/article.php?ch=14&id=149>. Acesso em: 25 set. 2004.
- GUIMARÃES, S.E.F. Análise de marcadores genômicos e detecção de QTLs e genes candidatos em melhoramento animal. In: PEREIRA, J.C.C. *Melhoramento genético aplicado à produção animal*. 3. ed. Belo Horizonte: FEPMVZ Editora, 2001, p.445-478.
- HALL, B.G. *Phylogenetics Trees Made Easy. A How to Manual for Molecular Biologists*. Sunderland: Sinauer Associates, Inc., 2001, 179 p.
- HASEGAWA, M.; YANO, T.; KISHINO, H. A new molecular clock of mitochondrial DNA and the evolution of hominoids. *Proc. Japan Acad.*, v.60, p.95-98, 1984.
- HASEGAWA, M.; KISHINO, H.; YANO, T. Dating of the humam-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.*, v.32, p. 443-445, 1985.
- HASTY, P.; BRADLEY, A.; MORRIS, J.H. Muscle deficiency and neonatal death in mice with a targeted mutation in the *myogenin* gene. *Nature*, v.364, p. 501-506, 1993.
- HUNTER, R.L.; MARKERT, C.L. Histochemical demonstration of enzymes separated by zone electrophoreses in starch gels. *Science* v.125, p.1294-1295, 1957.
- JUKES, T.H.; CANTOR, C.R. Evolution of protein molecules. In: MUNRO, H.N. *Mammalian protein metabolism*. Academic Press, New York, v.3, p.21-132, 1969.
- KIMURA, M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, v.16, p.111-120, 1980.
- KOSHI, J.M.; GOLDSTEIN, R.A. Probabilistic reconstruction of ancestral protein sequences. *Journal of Molecular Evolution*, v.42, p.313-320, 1996.
- LEDUR, M.C.; SCHMIDT, G.S. *Genética molecular: aplicação de tecnologias moleculares no melhoramento genético de aves*. Revista *Avicultura Industrial*, 2000. Disponível em: <http://www.bichoonline.com.br/artigos/emrapave0007.htm>. Acesso em 19 dec. 2004.

- LIBERLES, D.A. Evaluation of Methods for Determination of a Reconstructed History of Gene Sequence Evolution. *Molecular Biology and Evolution*, v.18, p.2040-2047, 2001.
- LIBERLES, D.A.; WAYNE, M.L. Tracking adaptive evolutionary events in genome sequences. *Genome Biology*, v.3(6), reviews 1018.1-1018.4, 2002.
- LIBERLES, D.A. Detecting and Characterizing Adaptive Evolution in Chordate Proteins in Comparative Genomics of Vertebrates: Concepts and Bioinformatic Tools. INSERM, 2004.
- LIBERLES, D.A. "Using phylogeny to understand genomic evolution" in Parsimony, Phylogeny, and Genomics. Oxford: Oxford University Press, 2005.
- MESSIER, W.; STEWARD, C.B. Episodic adaptive evolution of primate lysozymes. *Nature*, v.385, p.151-154, 1997.
- MIYAKI, C. Y. Reconstrução filogenética. Introdução e o método de máxima parcimônia. In: MATIOLI, S. R. *Biologia Molecular e Evolução*. Ribeirão Preto: Holos Editora, 2001, p 97-107.
- NEI, M. *Molecular evolutionary genetics*. New York: Columbia University Press, 1987.
- NEI, M. Relative efficiencies of different tree-making methods for molecular data. In: MIYAMOTO, M. M.; CRACRAFT, J. *Phylogenetic analysis of DNA sequences*. Oxford University Press, New York, 1991, p.90-128.
- NEI, M.; KAMUR, S. *Molecular evolution and phylogenetics*. New York: Oxford University Press, 2000.
- PEREIRA, S.R. Reconstrução filogenética: Métodos probabilísticos. In: MATIOLI, S. R. *Biologia Molecular e Evolução*. Ribeirão Preto: Holos Editora, 2001, p 117-129.
- RAMALHO, M.A.P.; SANTOS, J.B.; PINTO, C.A.B.P. *Genética na Agropecuária*. 2.a edição. Lavras: Editora UFLA, 2001. 472p.
- ROTHSCILD, M.F.; SOLLER, M. Candidate gene analysis to detect traits of economic importance en domestic livestock. In: SIMPÓSIO INTERNACIONAL DE GENÉTICA E MELHORAMENTO ANIMAL, 1999, Viçosa. *Anais...Viçosa: UFV*, 1999, p.213-242.
- RUSSO, C.A.M. Efficiencies of different statistical tests in supporting a know vertebrate phylogeny. *Molecular Biology Evolution*, v.12, p391-404, 1997.
- RUSSO, C.A.M. Reconstrução filogenética: Métodos geométricos. In: MATIOLI, S. R. *Biologia Molecular e Evolução*. Ribeirão Preto: Holos Editora, 2001, p 108-116.

- SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology Evolution*, v.4, p.406-425, 1987.
- SCHNEIDER, H. Métodos de análise filogenética: Um guia prático. 2a. ed. Ribeirão Preto: Editora Holos e Sociedade Brasileira de Genética. 2003. 113p.
- TAJIMA, F.; NEI, M. Estimation of Evolutionary distance between nucleotide sequences. *Molecular Biology Evolution*, v.1, p.269-285, 1984.
- TAMURA, K. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C content bias. *Molecular Biology Evolution*, v.9, p.678-687, 1992.
- TAMURA, K.; NEI, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology Evolution*, v.10, p.512-526, 1993.
- TELLGREN, Å.; BERGLUND, A.C.; SAVOLAINEN, P. *et al.* Myostatin rapid sequence evolution in ruminants predates domestication. *Molecular Phylogenetics and Evolution*, v.33, p.782-790, 2004.
- TE PAS, M.F.W.; SOUMILLION, A.; HARDERS, F.L. *et al.* Influences of Myogenin genotypes on birth weight, growth rate, carcass weight, backfat thickness, and lean weight of pigs. *J. Anim. Sci.* v.77, p.2352-2356, 1999.
- TE PAS, M.F.W.; VERBURG, F.J.; GERRITSEN, C.L.M. *et al.* Messenger ribonucleic acid expression of the MyoD gene family in muscle tissue at slaughter in relation to selection for porcine growth rate. *J. Anim. Sci.* v.78, p.69-77, 2000.
- UZZEL, T.; CORBIN, K.W. Fitting discrete probability distribution to evolutionary events. *Science* v.172, p.1089-1096, 1971.
- YANG, Z. Likelihood Ratio Tests for Detecting Positive Selection and Application to Primate Lysozyme Evolution. *Molecular Biology Evolution*, v.15(5), p.568-573, 1998.
- YANG, Z.; NIELSEN, R. Códon-Substitution Models for Detecting Molecular Adaptation at Individual Sites Along Specific Lineages. *Molecular Biology Evolution*, v.19(6), p.908-917, 2002.
- ZHANG, J.; NEI, M. Accuracies of ancestral amino acid sequences inferred by parsimony, likelihood, and distance methods. *J. Mol. Evol.*, v.44, p.S139-146, 1997.

CAPÍTULO 1

ANÁLISE DO SEQUENCIAMENTO DO GENE DA MIOGENINA EM SUÍNOS DA RAÇA NATIVA PIAU E DE LINHAGENS COMERCIAIS

RESUMO

O conhecimento de polimorfismos em genes que regulem a formação muscular pode ser usado em programas de melhoramento genético por influenciarem a quantidade de fibras musculares em animais recém nascidos. O gene da Miogenina é um dos membros da família de genes MyoD, que controla a miogênese durante o desenvolvimento embrionário, atuando como regulador da expressão de genes músculo específicos, exercendo efeito direto na quantidade de células que formam as fibras musculares. O objetivo desse estudo foi o de encontrar possíveis polimorfismos no gene da Miogenina que possam vir a ser usados como marcadores em programas de melhoramento de suínos. Foram usadas amostras de DNA extraídos do sangue de dois suínos da raça naturalizada Piau e de 12 matrizes de linhas comerciais. Os fragmentos foram amplificados por PCR usando *primers* específicos que flanqueiam regiões pré-definidas do gene. Os sequenciamentos dos fragmentos amplificados foram feitos por meio da técnica de reação de terminação de cadeia por dideoxinucleotídeo (ddNTP) e analisados usando o programa ClustalX. Os resultados do sequenciamento revelaram que esse gene é altamente conservado, não contendo, nas regiões analisadas, nenhum polimorfismo entre os animais estudados.

Palavras-chave: miogênese, gene candidato, sequenciamento.

CHAPTER 1

SEQUENCING ANALYSIS OF THE MYOGENIN GENE IN THE SWINE PIAU BREED AND IN COMERCIAL LINES

ABSTRACT

Polymorphism identification in genes that regulate the muscle formation may be used in genetic breeding programs because they affect the muscle fibers quantity in new born animals. The Myogenin gene is member of the MyoD gene family, which rules the myogenesis during embryonic development acting as a regulator of specific muscle genes expression, bearing direct effects in the amount of cells that compose the muscle fibers. The objective of this study was to find putative polymorphism in the Myogenin gene that might be used as a marker in swine breeding programs. DNA samples were extracted from blood of two boars of the naturalized Piau breeding and from 12 commercial sows. The fragments were amplified by PCR using specific primers flanking pre-defined regions of the gene. The sequencing of the amplified fragments was made by the dideoxynucleotide terminal chain reaction (ddNTP) and was analyzed using the ClustalX program. The sequences results showed that this gene is highly conserved, having no polymorphism among the studied animals in the analyzed regions.

Keywords: myogenesis, candidate gene, sequencing analysis.

INTRODUÇÃO

O desenvolvimento de pesquisas e tecnologias tem refletido de maneira significativa na produção de suínos no Brasil, com um abate de 33,9 milhões de suínos no ano de 2004. Na região sul do país a taxa de abate teve média de 23 suínos terminados/matriz/ano, com índices de produção e produtividade comparados aos obtidos nos EUA, Canadá, Dinamarca, Alemanha, Holanda e outros (Fonte: ABIPEC, 2005. Disponível em: <http://www.abipecs.com.br/cultura.php>).

A definição de qualidade da carne suína depende muito do mercado a que esta se destina. Para entender sobre rendimento de carcaça e qualidade da carne, primeiramente deve-se reconhecer que este conceito de qualidade é definido em função de um objetivo, que está diretamente relacionado à cadeia de produção e distribuição da carne. Com isto, tem-se a responsabilidade de produzir o melhor produto de acordo com as exigências do mercado específico, sem que se perca de vista o objetivo final (Benevenuto Júnior, 2001).

Do ponto de vista da indústria, têm maior importância as qualidades como magreza da carne, rendimentos de cortes, menores custos no processamento da carne e a integridade do produto após estocagem no frigorífico. Os produtores por sua vez estão interessados em um produto com uma maior eficiência para otimizar o lucro do seu sistema de produção, ou seja, características como peso ao nascimento, conversão alimentar, ganho de peso diário e outras características produtivas, além das características reprodutivas (Benevenuto Júnior, 2001).

O peso ao nascimento e o ganho de peso diário são características associadas ao desenvolvimento muscular. Como o principal componente da carne é o tecido muscular estriado esquelético, os genes envolvidos no fenômeno de formação do tecido muscular podem ser importantes para a indústria de produção de carne (Te Pas *et al.*, 2000).

A família de genes MyoD é responsável pela regulação da miogênese, processo que ocorre durante o desenvolvimento embrionário, responsável pela formação do tecido muscular esquelético. A família MyoD é composta de quatro genes: MyoD1, MYF-5, Miogenina e MYF-6, que se expressam de maneira sincronizada durante a

formação dos músculos. O processo de miogênese pode ser dividido em duas fases, a determinação, onde as células especificadas para a formação do tecido crescem por hiperplasia (divisão celular), e a diferenciação, quando as células se unem para formar as fibras musculares e, a partir daí, só crescem por hipertrofia (volume celular). Especificamente, a expressão do gene da Miogenina desencadeia o início da fase de diferenciação (Coutinho *et al.*, 1999).

Dessa maneira, a partir do momento que a Miogenina começa a se expressar, as fibras musculares serão formadas pelas células produzidas durante a fase de determinação. O gene da Miogenina é, então, um importante regulador da miogênese e deve ser estudado na tentativa de se detectar algum polimorfismo vantajoso à indústria de carnes.

Os objetivos desse trabalho foram verificar a existência de polimorfismos no gene da Miogenina em suínos da raça naturalizada Piau e de linhas comerciais e associar os possíveis polimorfismos identificados com características de produção em suínos

MATERIAL E MÉTODOS

População em estudo

Neste trabalho, foram analisadas amostras de DNA de 14 animais (12 matrizes e dois varrões) da espécie suína (*Sus scrofa*). A identificação das matrizes seguiu a numeração da granja de melhoramento de suínos da UFV. As matrizes foram originárias do cruzamento entre animais Landrace com Large White e do cruzamento Landrace/Large White com Pietrain. Os dois varrões utilizados eram animais da raça naturalizada Piau e foram identificados como C e J.

As matrizes foram alojadas na granja de melhoramento genético de suínos da Universidade Federal de Viçosa, localizada no Campus Universitário. Os varrões foram mantidos na fazenda Boa Vista, município de Viçosa-MG.

Extração do DNA

O sangue foi coletado por meio de punção venosa do *sinus orbitalis* em tubos, de 10 mL contendo EDTA 0,5 M pH 8,0. As amostras foram centrifugadas a 2.500 g durante 20 minutos. A camada de células brancas, localizada entre as células vermelhas (camada inferior) e o plasma sanguíneo (camada superior), foi isolada com o auxílio de pipetas e removida para tubos limpos. A extração do DNA genômico, a partir das células brancas, foi realizada por purificação com fenol - clorofórmio, após tratamento com proteinase K, como descrito por Sambrook *et al.* (1989). As amostras para uso em reações de PCR foram diluídas, de acordo com a quantificação, para a concentração de 25 ng/ μ L em solução contendo 10 mM de Tris-HCl pH 8,0 e 1,0 mM de EDTA pH 8,0, sendo mantidas a 4°C.

Construção dos *Primers*

Foram construídos três pares de *primers* que flanqueiam as regiões dos 3 exons e mais 4 pares que cobririam os dois introns e as regiões 5' e 3' do gene da Miogenina (Fig. 1).

Os *primers* foram desenhados a partir do programa *Web Primers* (<http://alces.med.umn.edu/websub.html>), com base no sequenciamento feito por

Soumillion *et al.* (1998), sob o número de acesso no Genbank X89007 (<http://www.ncbi.nlm.nih.gov/Entrez>). A Tab. 1 descreve os pares de *primers* MYOG, que flanqueiam os três exons, e os pares de *primers* MIOGIN, que flanqueiam os dois introns e as regiões 5' e 3'. Os *primers* foram selecionados conforme suas posições no gene, temperaturas de anelamento e variação de energia livre (ΔG).

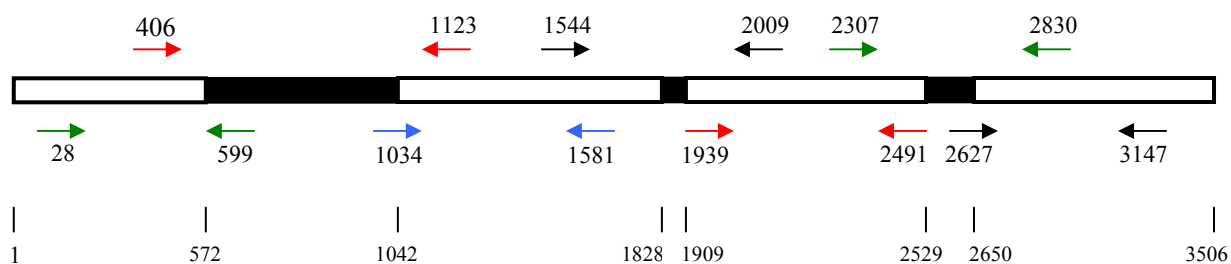


Figura 1 – Orientação e localização dos *primers* MYOG (setas acima do esquema) e dos *primers* MIOGIN (setas abaixo do esquema) em relação à seqüência depositada no Genbank sob número de acesso X89007. Tarjas em branco representam os introns e as preenchidas representam os exons. O posicionamento desses na linha contínua está representada em pares de bases. *Primers forward* (→) e *primers reverse* (←).

Tabela 1: *Primers* utilizados para amplificar o gene da Miogenina em suínos.

Nome	Seqüência do <i>primer</i>	ΔG^1 (Kcal/mol)	TA ² (°C)	Tamanho do fragmento (pb)
MYOG F1	TTGATGTGCAGCAACAGCTTAGA	-4,2	60	738
MYOG R1	TGTCATCCATCTGACCGAGC	-6,3	58	
MYOG F2	AGGAAACCACAGACCCCTTAAG	-4,5	64	466
MYOG R2	AAGAGGACTGGAGCAGGGTGTA	-4,2	64	
MYOG F3	GGGGATTCCGAGAGGACATACT	-3,9	64	524
MYOG R3	CTCTCTGGGTGGTGGGCTAAC	-4,2	62	
MIOGIN F1	ACTGACACAGTCTGGGTAAGGT	-6,0	60	572
MIOGIN R1	CTGTATGAGACATCCCCCTACT	-4,8	60	
MIOGIN F2	CAGCCAGGGGTAAGTGGCTGTC	-4,8	66	553
MIOGIN R2	TAAGCTGAGTGTCTGTGAGGACA	-4,8	66	
MIOGIN F3	GACTCAAGTCCCACAGTCTATC	-4,1	62	576
MIOGIN R3	CATTACCCAGGTAACATCTTTCTA	-4,2	62	
MIOGIN F4	CCAGATGAAACCATGCCCAACTGA	-5,1	66	521
MIOGIN R4	AACAATCCTCTATTCCCAAGTCCCT	-5,1	66	

¹Energia livre de formação de duplex; ²Temperatura de anelamento.

Amplificação dos Fragmentos por PCR

A amplificação dos fragmentos foi feita utilizando um volume final de 100 μ L contendo 4,0 pmoles de cada *primer*; 0,2 mM de cada dNTP; 0,2 mM de MgCl₂; 50 mM de KCl; 20 mM de Tris-HCl; 5 unidades de Taq DNA polimerase e 125 ng de DNA.

A reação foi conduzida em termociclador MJ Research PTC 100-96[®] com o programa de amplificação constituído de um passo inicial de desnaturação a 94°C, por 5 minutos, seguido dos três passos básicos da PCR: desnaturação a 94°C por 1 minuto; anelamento por 1 minuto, com temperaturas alteradas de acordo com o *primer* para diminuir a amplificação inespecífica ou aumentar o rendimento da amplificação; e a extensão de 1 minuto do fragmento pela Taq polimerase a uma temperatura de 72°C. O número de vezes que o ciclo se repetiu, a temperatura de anelamento e a concentração de MgCl₂ na reação de PCR foram alteradas para aumentar a eficiência da reação. Os

resultados das ampliações foram carregados em géis de poliacrilamida a 8% (5 µL de cada reação), corridos em cubas de eletroforese verticais e visualizados por coloração em nitrato de prata.

Purificação dos Fragmentos Amplificados

Com a finalidade de otimizar a técnica de sequenciamento, o DNA amplificado das matrizes foram unidos em 3 grupos, contendo 4 animais em cada, conforme a Tab. 2.

Tabela 2 - Números dos grupos, identificação dos animais e suas respectivas raças.

Número dos <i>Mix</i>	Identificação dos animais	Raças
1	22, 25, 90, 80	Landrace x Large White
2	69, 127, 130, 133	Landrace x Large White
3	147, 159, 160, 175	Landrace/LargeWhite x Pietran

Para a purificação dos fragmentos do gene Miogenina, foram utilizados 100 µL da reação de amplificação de cada macho e 25 µL de cada uma das quatro fêmeas que comporiam aquele mix, num volume final de 100 µl. Foi utilizado o *GFXTM PCR DNA and Gel Band Purification Kit (Amersham Pharmacia Biotech)*, segundo o protocolo que segue.

A coluna GFX foi posicionada sobre o tubo coletor onde foram adicionados 500µl de *Capture Buffer* e 100µl do produto de PCR de cada animal/mix. Este material foi homogeneizado cuidadosamente e posteriormente centrifugado a 13.000 rpm por 30 segundos. O efluente foi então descartado e adicionado 500µl de *Wash Buffer* (10mM Tris-HCl pH8,0; 1mM EDTA) sobre a coluna. Novamente esta foi centrifugada a 13.000 rpm por 30 segundos. Nesta etapa descartou-se o tubo coletor e foi posicionado sobre a coluna um microtubo devidamente identificado. Foi adicionado ao topo da matriz da coluna 20µl de água Mili-Q bidestilada. Nesta fase, incubaram-se as amostras em temperatura ambiente por 5 minutos, seguido de centrifugação a 14.000 rpm por 3 minutos.

Após o procedimento de purificação, obteve-se aproximadamente 20µl de DNA, que foi quantificado em espectrofotômetro GeneQuant II (*Amersham Pharmacia Biotech*) por meio da leitura em absorvância de 260 nm e 280 nm, usando um comprimento de onda de 320 nm, para correção de *background*.

Sequenciamento e Genotipagem

As reações de sequenciamento foram baseadas na técnica de terminação de cadeia por dideoxynucleotídeos (ddNTPs), descrita por Sanger *et al.* (1977). O DNA foi sequenciado utilizando-se o “ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems)”.

Para a reação de sequenciamento, foi utilizado o mesmo protocolo já em uso no Laboratório de Biotecnologia Animal do DZO/UFV, onde cerca de 25 ng do fragmento de DNA obtido via PCR purificado anteriormente foi empregado em uma reação contendo 0,7 µL de *BigDye* (PE Applied Biosystems), 4,0 pmoles do *primer* direto e 2 µL de tampão (200 mM Tris-HCL pH 9,0 e 5 mM MgCl₂), 5,3 µL de água Mili-Q bidestilada, em um volume total de reação de 10 µL. Para o *primer* reverso foi utilizado um “mix” semelhante.

Para as reações de amplificação foi utilizado um termociclador “MJ Research, Inc. modelo PTC-100-Watertown, USA” com um passo inicial de desnaturação a 96°C por 20 segundos, - 46°C (1°C a cada segundo), 50°C por 10 segundos (anelamento), +10°C (1°C a cada segundo) e 60°C por 4 minutos (extensão) em um total de 45 ciclos.

Após a amplificação, o DNA foi precipitado pela adição de 72 µL de isopropanol a 75% e 18 µL de água mili-Q autoclavada e os tubos deixados à temperatura ambiente por 20 minutos. Em seguida, este foi submetido ao Vortex lentamente durante 5 segundos e posteriormente centrifugado a 13.000 rpm por 30 minutos e o sobrenadante foi descartado. O precipitado foi lavado com 200 µL de etanol a 75%, centrifugado a 13.000 rpm por 5 minutos, por três vezes e seco em placa quente ou no ar. O pelete foi ressuspensionado em 25 µL de formamida Hi-Di (PE Applied Biosystems), desnaturado a 95°C por cinco minutos e mantido no gelo até a sua aplicação.

Os produtos da extensão foram então separados por eletroforese capilar, utilizando-se como matriz o polímero POP6 (PE Applied Biosystems), e a fluorescência

emitida foi coletada por câmara CCD, sendo a informação processada pelo “ABI PRISM 310 *Genetic Analyzer (PE Applied Biosystems)*” existente no Laboratório de Biotecnologia Animal do Departamento de Zootecnia da Universidade Federal de Viçosa.

As seqüências geradas foram visualizadas com o programa Chromas 2.3 (<http://www.technelysium.com.au/chromas.html>) e alinhadas e editadas pelo programa ClustalX (Thompson *et al.*, 1997)

RESULTADOS E DISCUSSÃO

Amplificação dos Fragmentos por PCR

As seqüências de nucleotídeos e a localização dos sete pares de *primers* e suas respectivas características estão apresentadas na Tab. 1 e na Fig. 1, respectivamente. Todos os fragmentos do gene da Miogenina amplificaram com especificidade e com bom rendimento nas condições de PCR utilizadas neste estudo, com exceção dos fragmentos gerados pelo par de *primers* MIOGIN 3, onde os reagentes não amplificaram com exatidão o fragmento alvo. O tamanho das bandas foi verificado com um padrão de 1Kb (*Invitrogen*) e eram equivalentes aos tamanhos dos fragmentos que flanqueavam os pares de *primers*

Sequenciamento do gene Miogenina

Todos os 6 fragmentos amplificados foram seqüenciados e produziram um sequenciamento confiável, a não ser pelo fragmento MIOGIN 2, que apresentou um eletroferograma com elevados índices de contaminação apesar de a amplificação ter ocorrido com especificidade.

Na Fig. 2 é apresentado um fragmento do eletroferograma do fragmento MYOG 2, de um dos machos da raça Piau.

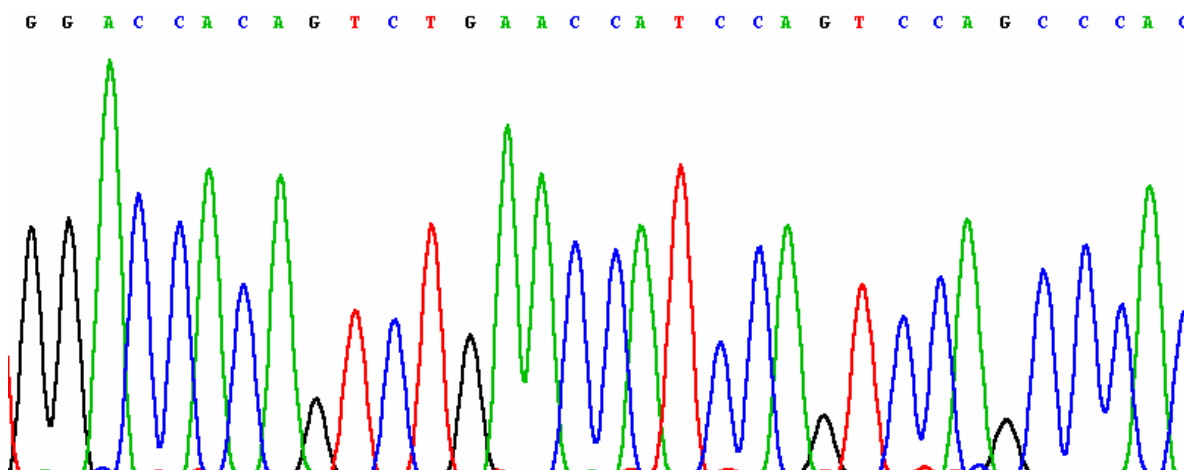


Figura 2 – Exemplificação do sequenciamento do fragmento MYOG 2 visualizado em eletroferograma

Análise do Sequenciamento

As seqüências dos cinco fragmentos do gene Miogenina, tanto dos dois varrões da raça Piau quanto dos 3 grupos de matrizes comerciais, foram alinhadas com o sequenciamento feito por Soumillion *et al.* (1997), sob o número de acesso no Genbank X89007 (<http://www.ncbi.nlm.nih.gov/Entrez>, acessado em 27/02/2004).

Nas seqüências de nucleotídeos produzidas pelo sequenciamento automático dos produtos de PCR, referentes aos fragmentos do gene Miogenina, geradas por esse estudo, não foi observado, tanto na região promotora quanto nos exons e até mesmo nos fragmentos dos introns seqüenciados, nenhum tipo de mutação entre os animais comerciais e os suínos da raça Piau. Na região 3' do gene existem polimorfismos entre os animais utilizados nesse estudo e a seqüência depositada no Genbank, mas considerando apenas as matrizes comerciais e os varões Piau a região continua mostrando-se conservada.

Nas Fig. 3 a 7 são apresentados os alinhamentos de todos os fragmentos seqüenciados. Ressalta-se que, como nem todos os fragmentos foram seqüenciados, por estarem faltando os sequenciamentos dos fragmentos MIOGIN 2 e 3, existem *gaps* entre o fim de um alinhamento e o início do alinhamento do fragmento seguinte.

Fragmento MIOGIN 1 (região promotora)

	1						70
Genbank	GATCTTTTTT	TAAGAGAGTC	TCATCTGACT	GACACAGTCT	GGGTAAGGTG	CTGTGAGGAA	GCAGGGGGAT
C	*****	*****	*****	*****	*****	*****	*****
J	*****	*****	*****
Mix 1	*****	*****	*****	*****	*****	*****	*****
Mix 2	*****	*****	*****	*****	*****	*****	*****
Mix 3	*****	*****	*****	*****	*****	*****	*****
	71						140
Genbank	GCATAAACTG	ACTTCTCCAG	GCCCCTTCCA	GCCTACACCT	ACCCCTTCCT	TCTTCCCCC	C-GCCTCACC
C
JC.....
Mix 1	*****
Mix 2	*****
Mix 3	*****	*
	141						210
Genbank	CCCACCCCA	CTGGGCTTCT	TTGGGACTGG	CGAGTAGGCA	GGCCGCCAG	CTAGGAGTAA	TTGAAAGGAG
C
J
Mix 1
Mix 2
Mix 3
	211						280
Genbank	CAGATGAGAG	GAGAATGTGT	GTCCTCCCC	ACCTCCCCAG	CCCCCATGGG	GGCTGCAGAG	AAATGAAAAC
C
J
Mix 1
Mix 2
Mix 3
	281						350
Genbank	TAATCAAATT	ACACCCTATG	GCCTCCTTAC	CCGTGCACAG	GAGCCTGCTG	GGGGCAGGCT	GGCTGTGGGG
C
J
Mix 1
Mix 2
Mix 3
	351						420
Genbank	AGGGGGGGTG	CAGGGGAGAG	GGGAAGGGGA	ATCACATCTA	ATCCACTGTA	AACGTCTTGA	TGTGCAGCAA
C
J
Mix 1
Mix 2
Mix 3
	421						490
Genbank	CAGCTTAGAG	GGGGGCTCAG	GTTTCTGTGG	CGTTGGCTAT	ATTTATCTCT	GGTTCCATGC	CAGCGGGGAG
C
J
Mix 1
Mix 2
Mix 3
	491						560
Genbank	GGTTTAAATG	GCACCCAGCA	GTTGGCGTGA	GGGGCTGCAG	GAGCTTGGGG	GCTGGTGGCA	GGAACAAGTC
C
J
Mix 1	****	*****	*****	*****	*****
Mix 2	****	*****	*****	*****	*****
Mix 3	****	*****	*****	*****	*****
	561				601		
Genbank	TTTTCTGACC	CCATGGAGCT	GTATGAGACA	TCCCCCTACT	T		
C
J
Mix 1	*****	*****	*****	*****
Mix 2	*****	*****	*****	*****
Mix 3	*****	*****	*****	*****

Figura 3 – Comparação das seqüências de nucleotídeos do fragmento MIOGIN 1, contendo a região promotora do gene Miogenina. C e J são varrões da raça Piau e os mix são fêmeas comerciais. O exon 1 está representado em

negrito. O ponto (.) representa seqüências homólogas e o asterisco (*) corresponde a regiões não seqüenciadas.

Fragmento MYOG 1 (exon 1)

	1						70
Genbank	GTTGGCGTGA	GGGGCTGCAG	GAGCTTGGGG	GCTGGTGGCA	GGAACAAGTC	TTTTCTGACC	CCATGGAGCT
C
J
Mix 1
Mix 2
Mix 3
	71						140
Genbank	GTATGAGACA	TCCCCCTACT	TCTACCAGGA	ACCCCACTTC	TATGACGGGG	AAAACCTACCT	GCCCCGTCCAC
C
J
Mix 1
Mix 2
Mix 3
	141						210
Genbank	CTCCAGGGCT	TTGAGCCACC	AGGCTACGAG	CGGACTGAGC	TGAGTCTGAG	CCCTGAGGCC	CGAGTGCCCC
C
J
Mix 1
Mix 2
Mix 3
	211						280
Genbank	TGGAAGATAA	GGGGCTGGGG	ACCCCCGAGC	ACTGCCCAGG	CCAGTGCCTG	CCGTGGGCAT	GTAAGGTGTG
C
J
Mix 1
Mix 2
Mix 3
	281						350
Genbank	TAAGAGGAAG	TCCGTGTCTG	TGGACCGTCG	GCGGGCCGCC	ACGCTGAGGG	AGAAGCGCAG	GCTCAAGAAG
C
J
Mix 1
Mix 2
Mix 3
	351						420
Genbank	GTGAATGAGG	CCTTTGAGGC	CCTGAAGAGG	AGCACCTGTC	TCAACCCCAA	CCAGCGGCTG	CCCAAGGTGG
C
J
Mix 1
Mix 2
Mix 3
	421						490
Genbank	AGATCCTGCG	CAGCGCCATC	CAGTACATCG	AGCGCCTGCA	GGCCCTGCTC	AGCTCCCTCA	ACCAGGAGGA
C
J
Mix 1
Mix 2
Mix 3
	491					549	
Genbank	GCGAGACCTC	CGCTACCGAG	GCGGGGGCGG	GCCGCAGCCA	GGGGTAAGTG	GCTGTCCCA	
C	
J	
Mix 1	
Mix 2******	
Mix 3	

Figura 4 – Comparação das seqüências de nucleotídeos do fragmento MYOG 1, contendo o exon 1 da Miogenina. C e J são varrões da raça Piau e os mix são fêmeas comerciais. O exon 1 está representado em negrito. O ponto (.)

representa seqüências homólogas e o asterisco (*) corresponde a regiões não seqüenciadas.

Fragmento MYOG 2 (exon 2)

	1						69
Genbank	AGGAAACCAC	AGACCCCTT	AAGCTGAGT	GTCCTGTGAG	GACAGAATTG	CCACCCAGAA	TGGATGCTAG
C
J
Mix 1
Mix 2
Mix 3	*****
	70						139
Genbank	GAAAGGCCTT	GAAGCAGGGG	TTCTGTGGCC	TGACAGAGGC	ATTGAAAGGA	GCTGATTTTG	AAGATTCTGG
C
J
Mix 1
Mix 2
Mix 3
	140						209
Genbank	AATAGGAGGT	GGGCTGGACT	GGATGGTTCA	GACTGTGGTC	CACCTTCCAA	CCCCACCACA	GGTGCTTCTG
C
J
Mix 1
Mix 2
Mix 3
	210						279
Genbank	CCCATTTGAA	CGAGACAAAG	AGACACTGCT	AAGAACAGCG	GAGCCCTGAC	CTTGGGCCTT	CACCCATATCT
C
J
Mix 1
Mix 2
Mix 3
	280						349
Genbank	TGCAGGTGCC	CAGTGAATGC	AGTTCCACACA	GCGCCTCCTG	CAGTCCAGAA	TGGGGCAGTG	CACTGGAGTT
C
J
Mix 1
Mix 2
Mix 3
	350						419
Genbank	CGGCCCCAAC	CCAGGGGGTA	AGTGAGGCCA	GGACCCTGGT	TACCTTGACT	CAAGTCCCAC	AGTCTATCTG
C
J
Mix 1
Mix 2
Mix 3
	420				466		
Genbank	TAGGTCCAGT	TCTCCCCTCT	TACCTTACAC	CCTGCTCCAG	TCCTCTT		
C		
J		
Mix 1		
Mix 2		
Mix 3*		

Figura 5 – Comparação das seqüências de nucleotídeos do fragmento MYOG 2, contendo o exon 2 da Miogenina. C e J são varrões da raça Piau e os mix são fêmeas comerciais. O exon 2 está representado em negrito. O ponto (.) representa seqüências homólogas, o asterisco (*) corresponde a regiões não seqüenciadas.

Fragmento MYOG 3 (exon 3)

	1						70
Genbank	TCTGTCCTGG	AAAACCTGTG	CCACTTCAAA	GAGCTCAGCC	TCTGCCTTGA	GTCTTGAGAC	CCTTTCATTA
C
J
Mix 1
Mix 2
Mix 3
	71						140
Genbank	CCCAGGTAAC	ATCTTTCTAC	CTTCTCCTTA	CAGATCATCT	GCTCACAGCT	GACCCTACAG	ATGCCACAAA
C
J
Mix 1
Mix 2
Mix 3
	141						210
Genbank	TCTGCACTCC	CTCACCTCCA	TCGTGGACAG	CATCACAGTG	GAGGATGTGG	CTGTGGCCTT	CCCAGATGAA
C
J
Mix 1
Mix 2
Mix 3
	211						280
Genbank	ACCATGCCCA	ACTGAGACTG	TCTGCCAGGA	TGGGTGTGCG	TGAGAGCCCC	CCCCAAGGCT	GGCCACAGAC
C
J
Mix 1
Mix 2
Mix 3
	281						350
Genbank	GGCACCACCTT	CTGCAGCAGG	GCTCTCCTAA	GCCAGTTGTC	CTGCTGCCAG	GAAGCCAGCC	CTGGGGTTGC
C
J
Mix 1
Mix 2
Mix 3
	351	370					
Genbank	CAAATGCCAG	ACTAACCCCC					
C					
J					
Mix 1					
Mix 2	****				
Mix 3					

Figura 6 – Comparação das seqüências de nucleotídeos do fragmento MYOG 3, contendo o exon 3 da Miogenina. C e J são varrões da raça Piau e os mix são fêmeas comerciais. O exon 3 está representado em negrito. O ponto (.) representa seqüências homólogas e o asterisco (*) corresponde a regiões não seqüenciadas.

Fragmento MIOGIN 4 (3')

	1						70
Genbank	CTCCATCGTG	GACAGCATCA	CAGTGGAGGA	TGTGGCTGTG	GCCTTCCCAG	ATGAAACCAT	GCCCAACTGA
C	*****	*****	*****	*****	*****	*****	*****
J	*****	*****	*****	*****	*****	*****	*****
Mix 1	*****	*****	*****	*****	*****	*****	*****
Mix 2	*****	*****	*****	*****	*****	*****	*****
Mix 3	*****	*****	*****	*****	*****	*****	*****
	71						140
Genbank	GACTGTCTGC	CAGGATGGGT	GTGCGTGAGA	GCCCCCCCA	AGGCTGGCCA	CAGACGGCAC	CACTTCTGCA
C	*****
J	*****
Mix 1	*****
Mix 2	*****
Mix 3	*****
	141						210
Genbank	GCAGGGCTCT	CCTAAGCCAG	TTGTCCTGCT	GCCAGGAAGC	CAGCCCTGGG	GTTGCCAAAT	GCCAGACTAA
C
J
Mix 1
Mix 2
Mix 3
	211						280
Genbank	CCCCCTCTC	ATCCATATAA	GGTTAGCCCA	CCACCCAGAG	AGGGAATGGA	TGCTCTCATT	TATCTGACTC
C
J
Mix 1
Mix 2
Mix 3
	281						350
Genbank	CTTAGAGCAG	AAAGCAGTTC	TGTTTCCAAG	GGGATAAAAC	AGGGGACCAG	AGTGCCCCCT	TGCGTAASCC
CG..
JG..
Mix_1G..
Mix_2G..
Mix_3G..
	351						420
Genbank	CC-TGGCTCA	GGG-ACAAAC	TCAGGAGCTT	CCCTTTGATC	ATAATGCAGC	CTTCAATTC	ACCCCTTAAA
C	..C.....	..G.....
J	..C.....	..G.....
Mix_1	..C.....	..G.....
Mix_2	..C.....	..G.....
Mix_3	..C.....	..G.....
	421						490
Genbank	AAGAAACAGT	TTGAGAGACG	AAGAGTGTCT	TGACCTGGAC	AAGCTATGCA	CATCTCCTGT	TCGTGTCTCT
C
J
Mix_1
Mix_2
Mix_3
	491						560
Genbank	TCCTAAGCCA	GTGGCTAGGC	TGGGCTG--C	CTGAATTGAG	AGAAGAAAAG	GAGAGGAACA	ATCCTCTATT
CCTGC
J	..*****	*****	*****	*****	*****	*****	*****
Mix_1CTGC
Mix_2CTGC
Mix_3CTGC
	561				604		
Genbank	CCCAAGTCCC	TGGGGGCCA	AGCTTTTGCA	GTGAATATTG	GGAA		
C	*****	*****	*****	*****	****		
J	*****	*****	*****	*****	****		
Mix 1	..*****	*****	*****	*****	****		
Mix 2	..*****	*****	*****	*****	****		
Mix 3	..*****	*****	*****	*****	****		

Figura 7 – Comparação das seqüências de nucleotídeos do fragmento MIOGIN 4, contendo a região 3' do gene Miogenina. C e J são varrões da raça Piau e os mix são fêmeas comerciais. O ponto (.) representa seqüências homólogas e o asterisco (*) corresponde a regiões não seqüenciadas.

Como pode ser observado, o gene da Miogenina apresenta um elevado grau de conservação, inclusive nos fragmentos dos introns que foram seqüenciados. O gene, por ser tão conservado, indica que provavelmente desempenha função fundamental no desenvolvimento da musculatura durante a miogênese e que não existe nenhum outro gene que possa substituir sua função, devido à deficiência muscular e inviabilidade dos animais que tiveram o gene da Miogenina mutado, evidenciando a importância do gene para a formação da estrutura muscular (Hasty *et al.*, 1993).

Soumillion *et al.* (1998) descreveram a localização de quatro haplótipos no loco da Miogenina em suínos que podem ser detectados por meio de PCR-RFLP usando a enzima de restrição *MspI*, mas nenhum dos sítios polimórficos se encontrava na região codificadora. Um dos sítios está localizado na região 3' do gene, outro situado dentro do segundo intron e um terceiro, identificado apenas na raça chinesa Meishan, dentro da região promotora. Porém, devido à distribuição limitada dos sítios polimórficos na região promotora e no segundo intron não foi possível realizar nenhum tipo de estudo de associação desses polimorfismos com características produtivas. No estudo aqui discutido, estas alterações não foram detectadas, mostrando o grau de conservação deste gene na amostra analisada.

Te Pas *et al.* (1999) genotiparam suínos da raça Yorkshire para o sítio polimórfico localizado na região 3' do gene com a enzima *MspI* e realizou associações de características produtivas com os genótipos. Os autores identificaram efeito aditivo do gene para as características de peso ao nascimento, peso de carcaça, taxa de crescimento e conteúdo de carne magra, já que os genótipos homozigotos diferiram significativamente entre si, e concluíram que genótipos do gene da Miogenina influenciam a taxa de crescimento e a massa muscular em suínos. No presente trabalho, este sítio polimórfico, citado por Te Pas *et al.* (1999) não foi encontrado, o que impossibilitou a análise de seu efeito na presente população.

CONCLUSÃO

Com base nos resultados obtidos não foi possível associar polimorfismos no gene da Miogenina com características de interesse econômico, já que esses não foram encontrados nas raças estudadas. Pode-se concluir que o gene da Miogenina foi altamente conservado dentro da amostra de suínos avaliada, não sendo, portanto, recomendável para uso em programas de melhoramento.

REFERÊNCIAS BIBLIOGRÁFICAS

- BENEVENUTO JÚNIOR, A. A., Avaliação de rendimento de carcaça e de qualidade da carne de suínos comerciais, de raça nativa e cruzados. 2001. 93p. Dissertação (Mestrado em Tecnologia de Alimentos) – Universidade Federal de Viçosa UFV, Viçosa.
- COUTINHO, L.L.; GABRIEL, J.E.; ALVARES, L.E. Controle molecular do desenvolvimento da musculatura esquelética. In: SIMPÓSIO INTERNACIONAL DE GENÉTICA E MELHORAMENTO ANIMAL, 1999, Viçosa. Anais... Viçosa: UFV, 1999, p.355-376.
- HASTY, P.; BRADLEY, A.; MORRIS, J.H. Muscle deficiency and neonatal death in mice with a targeted mutation in the *myogenin* gene. *Nature*, v. 364, p.501-506, 1993.
- SAMBROOK, J.; FRITSCH, E. F.; MANIATIS, T. *Molecular cloning: A Laboratory Manual*. 2 ed. New York: Cold Spring Harbor Laboratory Press, 1989. 564p.
- SOUMILLION A.; ERKENS, J.H.; LENSTRA, J.A. et al. Genetic variation in the porcine myogenin gene locus. *Mammalian Genome*. V.8(8), p.564-8, 1997.
- SOUMILLION, A.; ERKENS, J.H.; LENSTRA, J.A. *et al.* Genetic variation in the porcine myogenin gene locus. *Mammalian Genome* 8 (8), p. 564-568, 1998. Disponível em: <http://www.ncbi.nlm.nih.gov/Entrez>. Acesso em: 27 fev. 2004.
- Te PAS, M.F.W.; SOUMILLION, A.; HARDERS, F.L. *et al.* Influences of Myogenin genotypes on birth weight, growth rate, carcass weight, backfat thickness, and lean weight of pigs. *J. Anim. Sci.* v.77, p.2352-2356, 1999.
- Te PAS, M.F.W.; VERBURG, F.J.; GERRITSEN, C.L.M. *et al.* Messenger ribonucleic acid expression of the MyoD gene family in muscle tissue at slaughter in relation to selection for porcine growth rate. *J. Anim. Sci.* v.78, p.69-77, 2000.
- THOMPSON, J.D.; GIBSON, T.J.; PLEWNIAK, F.; JEANMOUGIN, F.; HIGGINS, D.G. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, v.24, p.4876-4882, 1997.

CAPÍTULO 2

ANÁLISE FILOGENÉTICA DO GENE DA MIOGENINA EM ANIMAIS COM SEQUENCIAMENTO DISPONÍVEL

RESUMO

O gene da Miogenina é membro da família MyoD, reguladora da miogênese que ocorre durante o desenvolvimento embrionário. O objetivo desse estudo foi analisar filogeneticamente o gene da Miogenina, estudando sua história evolutiva nas espécies domésticas com a seqüência do gene depositado no Genbank. Um método usado comumente para detectar evolução em genes é a comparação da taxa de substituição de nucleotídeos não sinônimas (Ka) pela taxa de substituições sinônimas (Ks). Genes com valores dessa taxa maiores que um (1) indicam que o gene sofreu mudanças que tornaram o organismo mais adaptado ao ambiente. As espécies analisadas foram: *Sus scrofa* (suínos), *Bos taurus* (bovinos), *Ovis áries* (ovinos), *Homo sapiens* (humanos), *Mus musculus* (camundongos), *Rattus norvegicus* (ratos), *Gallus gallus* (galinhas) e *Meleagris gallopavo* (perus). O alinhamento das seqüências foi feito utilizando o programa Clustal X e as árvores gênicas obtidas por Máxima Verossimilhança e as taxas de substituição sinônimas e não sinônimas por meio do método de Parcimônia. Os resultados indicaram que provavelmente o gene tenha sofrido uma evolução adaptativa nas espécies do grupo Ruminantia (*B. taurus* e *O. aries*) após as duas espécies terem divergido do seu ancestral comum. Nas outras espécies o gene pareceu estar evoluindo de maneira conservativa.

Palavras-chaves – Miogenina, análise filogenética, pressão de seleção.

CHAPTER 2

PHYLOGENY ANALYSIS OF THE MYOGENIN GENE IN ANIMALS WITH AVAILABLE SEQUENCING

ABSTRACT

The Myogenin gene is member of the MyoD gene family, regulator of the myogenesis that takes place during embryonic development. The Objective of this study was to perform a phylogeny analysis of the Myogenin gene to study its evolutionary history in the domestic species that have the sequencing data in the Genbank. One common method to detect gene evolution is made by comparing the ratio of nonsynonymous nucleotide substitution (Ka) by the ratio of synonymous substitutions (Ks). Genes with values higher than one (1) means that the gene had suffered changes that made the organism more adaptive to the environment. The species used in the analysis were *Sus scrofa* (swine), *Bos Taurus* (cows), *Ovis aries* (sheeps), *Homo sapiens* (humans), *Mus musculus* (mice), *Rattus norvegicus* (rats), *Gallus gallus* (chickens) and *Meleagris gallopavo* (turkeys). The sequences alignments were made using the program Clustal X and the phylogenetic trees were obtained by Maximum Likelihood and the synonymous and nonsynonymous substitution rates were by the Parsimony method. The results point out that probably the gene suffered an adaptive evolution in the Ruminantia group (*B. Taurus* and *O. aries*) after these species diverged from their common ancestral. In the other species, the gene seems to be evolving in a conservative way

Keywords – Myogenin, phylogenetic analysis, selection pressure.

INTRODUÇÃO

Ecologicamente, um organismo está mais adaptado ao meio quanto mais eficiente for o uso de sua energia para sua sobrevivência, de maneira que o saldo de energia que possa ser usado com funções reprodutivas seja maior.

A pressão de seleção positiva, na qual mudanças particulares em sítios do genoma dão aos organismos uma vantagem sobre outros da mesma espécie, podem rapidamente alterar a frequência desses sítios mais vantajosos em uma população.

Os genes que desempenham algum tipo de papel importante sobre características que influenciam na sobrevivência do animal, ou genes que maximizam a eficiência de uso de energia, de maneira que o saldo de energia disponível para a reprodução seja maior, são mais prováveis de terem sofrido uma pressão de seleção positiva durante a sua história evolutiva. Por exemplo, os genes que participam do complexo de histocompatibilidade, que atuam no papel de reconhecimento de anticorpos e são importantes determinantes do estado de sanidade do animal, ou genes que tenham papel na regulação do metabolismo de lipídeos, como o gene da Leptina, são exemplos prováveis de possuírem mutações não sinônimas que desencadearam melhor adaptabilidade do animal em seu nicho ecológico (Yang e Nielsen, 2002).

A massa muscular animal é também uma característica importante para a sobrevivência do animal, e por isso modificações que ocorram nos genes envolvidos na sua formação podem vir a oferecer aos animais uma melhor eficiência na utilização desse tecido (Tellgren, *et al.*, 2004).

Durante o desenvolvimento embrionário, uma das famílias de genes que regula a formação muscular é a família MyoD. Ela é composta por quatro genes, o MyoD1, o fator miogênico 5 (Myf - 5), a Miogenina e o fator miogênico 6 (Myf - 6). Cada gene é expresso em um tempo específico do desenvolvimento muscular, também chamado de miogênese. O processo de miogênese pode ser dividido em duas fases: determinação e diferenciação. A determinação é o evento em que as células pluripotentes que estão se multiplicando são mobilizadas para o processo miogênico, se transformando em mioblastos. A diferenciação ocorre quando os genes dos miotúbulos iniciam sua expressão músculo-específica, então os mioblastos param de se multiplicar e as fibras

musculares passam a crescer quase que somente em volume (Coutinho *et al.*, 1999). A Miogenina tem papel importante durante a miogênese pois a sua expressão marca o início da diferenciação. Conseqüentemente, quando a Miogenina é expressa, as fibras musculares se desenvolvem a partir dos mioblastos previamente formados.

O objetivo desse trabalho foi estudar a evolução do gene da Miogenina dentro das espécies de animais domésticos que possuem o gene seqüenciado e verificar se o gene sofreu pressão de seleção positiva em algum período durante a evolução das espécies.

MATERIAL E MÉTODOS

Seqüências utilizadas

Para as análises filogenéticas, as seqüências do gene da Miogenina, das espécies usadas para comparação com as seqüências obtidas no sequenciamento genético dos suínos, foram obtidas do Genbank (<http://www.ncbi.nlm.nih.gov/>, acessado em 15\07\2004), sendo os números de acesso escolhidos baseados na quantidade de informações depositadas sobre o mesmo para cada espécie. No caso de *Ovis aries*, apenas um sequenciamento havia sido depositado e este não estava completo, não contendo o início e o final do gene. Os números de acesso e as espécies estão listados na Tab. 1. Para as seqüências de *Sus scrofa*, foi usado uma seqüência consenso obtida a partir do sequenciamento de todo o gene da Miogenina tanto em raças comerciais como na raça naturalizada brasileira Piau.

Tabela 1 - Espécies com seqüências depositadas no Genbank do gene da Miogenina e que foram usadas para comparação com as seqüências consenso obtidas da raça naturalizada de suíno Piau e de raças comerciais de *Sus scrofa*.

ESPÉCIE	NOME VULGAR	ORIGEM	ACESSO	AUTOR*
<i>Homo sapiens</i>	Humanos	Genômico	AF050501	TSENG <i>et al.</i> (1999)
<i>Mus musculus</i>	Camundongos	Genômico	M95800	EDMONDSON <i>et al.</i> (1989)
<i>Rattus norvegicus</i>	Ratos	mRNA	M24393	WRIGHT <i>et al.</i> (1989)
<i>Ovis aries</i>	Ovinos	mRNA	AF433651	JEANPLONG <i>et al.</i> (2001)
<i>Bos taurus</i>	Bovinos	mRNA	AB110600	MUROYA <i>et al.</i> (2003)
<i>Gallus gallus</i>	Galinhas	mRNA	D90157	FUJISAWA-SEHARA <i>et al.</i> (2002)
<i>Meleagris gallopavo</i>	Peras	mRNA	AY560111	LIU <i>et al.</i> (2004)

* referências obtidas no site <http://www.ncbi.nlm.nih.gov/>

As análises foram feitas considerando as seqüências (caracteres) do exon 1, do exon 2, do exon 3 e do fragmento que se refere à região do mRNA maduro a ser traduzido, ou seja, a união do exon 1, do exon 2 e do exon 3. Foram comparados, separadamente, os fragmentos que compõem o domínio Básico e o domínio HLH (*helix-loop-helix*), que são os dois domínios funcionais encontrados no exon 1. Como

comentado anteriormente, a OTU (*operational taxonomic unit*, unidade taxonômica operacional) *Ovis aries* só possuía a seqüência completa para o domínio HLH.

Alinhamento das Seqüências

As regiões estudadas obtidas do Genbank foram editadas a partir do programa BLAST (Basic local alignment search tool), (Altschul *et al.*, 1997), usando como referência as seqüências obtidas no sequenciamento do capítulo 1. O tamanho dos domínios Básico e HLH foram também obtidos no Genbank. As seqüências dos fragmentos de todas as OTUs e a da raça Piau foram alinhadas com o programa Clustal X (Thompson *et al.*, 1997).

Escolha do Modelo

O modelo usado para a construção da árvore filogenética deve ser capaz de incorporar heterogeneidade nas taxas de substituição nucleotídica ao longo dos sítios, pois segundo Uzzel e Corbin (1971), nem todos os sítios evoluem na mesma taxa de substituição, como, por exemplo, aqueles sob forte restrição seletiva. A distribuição mais usada para modelagem é chamada de função gama.

Os alinhamentos, para cada fragmento separado e todos reunidos, foram analisados utilizando-se o programa Model Test 3.5, desenvolvido por Posada e Crandall (1998), que usa razão de máxima verossimilhança para determinar, com base nas características dos sequenciamentos alinhados, quais parâmetros serão usados na construção das árvores filogenéticas e as estimativas de tais parâmetros.

Construção das Árvores Filogenéticas

Para a construção das árvores filogenéticas foi utilizado o programa PAUP 4.0 beta 10. (Swofford, 1999) que utiliza métodos de parcimônia, máxima verossimilhança e de distâncias. Ele usa os arquivos criados pelo alinhamento com o Clustal X.

As topologias das árvores foram construídas usando o método de máxima verossimilhança, com os modelos evolutivos selecionados e os parâmetros estimados com a utilização do programa Model Test 3.5.

Para cada um dos fragmentos estudados foi escolhido um modelo evolutivo de máxima verossimilhança, baseado nas características das seqüências de nucleotídeos

desses fragmentos. As árvores foram enraizadas usando o grupo externo de *Gallus gallus* e *Meleagris gallopavo*.

Cálculo da Razão Ka/Ks

Muitos métodos estão disponíveis para calcular a razão Ka/Ks, desde o simples método de Nei e Gojobori (1986) até o método computacionalmente mais intensivo e altamente parametrizado de máxima verossimilhança de Yang e Nielsen (2000). O método PBL de Pamilo, Bianchi e Li (Li *et al.*, 1985; Pamilo e Bianchi, 1993, citados por Liberles, 2001), um dos mais populares e de bom desempenho, foi modificado por Benner *et al.* (1998), e chamado de PBLSB, usando uma implementação de reconstruções de seqüências ancestrais. Essa modificação permitiu a comparação de taxas de Ka/Ks ao longo dos ramos de uma árvore filogenética ao invés de uma comparação simples das seqüências par a par (Liberles, 2001).

Liberles (2001) usou o método PBLSB para a construção do banco de dados TAED (The Adaptive Evolution Database, ou Banco de Dados de Evolução Adaptativa). O TAED é um banco de dados virtual que contém exemplos potenciais de adaptação positiva em genes obtidos no *Master Catalog*, uma coleção de dados relacionados às famílias evolutivas, que inclui alinhamento de seqüências múltiplas, árvores filogenéticas e seqüências de ancestrais reconstruídas para os genes encontrados no GenBank.

Essa ferramenta se encontra na Internet <http://www.bioinfo.no/tools/kaks> e foi usada nesse estudo para realizar o cálculo da taxa Ka/Ks. O cálculo dos valores é executado segundo as fórmulas:

$$K_a = (L_2 * B_2 + L_0 * K_0) / (2 * L_2 / 3 + L_0)$$

$$K_s = (L_2 * A_2 + L_4 * A_4) / (L_2 + L_4) + B_4$$

em que:

Ka = taxa de transição não sinônima

Ks = taxa de transição sinônima

L0 = sítio não degenerado

L2 = sítio duas vezes degenerado

L4 = sítio quatro vezes degenerado, respectivamente.

$$K_i = A_i + B_i$$

A_i e B_i = taxa de transição e transversão para o i -ésimo tipo de sítio, respectivamente ($i = 0$ - não degenerado, 2 - duas vezes degenerado, 4 - quatro vezes degenerado)

As seqüências dos fragmentos analisados, alinhadas pelo Clustal X, foram usadas para a reconstrução dos estados ancestrais. Como o algoritmo trabalha com códon, foi realizado um ajuste de que todo códon no qual houvesse um *gap* (espaço) fosse preenchido por *gaps* nas seqüências alinhadas.

Esse procedimento deve ser tomado porque o algoritmo que estima K_a/K_s , da maneira que é implementado na ferramenta do cálculo, não lida com *frameshifts* (mudanças na fase de leitura do gene), já que com muita freqüência tais *frameshifts* são apenas erros no arquivo de entrada. O programa simplesmente não leva em consideração tais casos diretamente pois as seqüências com *frameshift* terão uma taxa muito alta de substituições não sinônimas (K_a), e irão, provavelmente, gerar um alto valor de K_a/K_s se a porção que sofreu *frameshift* for de comprimento significativo. A ferramenta de cálculo do K_a/K_s somente aceita árvores enraizadas e com todos os nós binários.

RESULTADOS E DISCUSSÃO

Foi realizado um alinhamento das seqüências das OTUs por meio do programa BLAST usando como referência o sequenciamento da raça Piau para serem estabelecidos os limites dos fragmentos a serem analisados.

Em relação às espécies *Bos taurus* e *Ovis aries*, a comparação gerou dúvida em relação ao estabelecimento dos limites do exon 2, já que, entre o final do exon 1 e o início do exon 2, estabelecidos por meio do BLAST, existem 6 nucleotídeos que não alinharam nem com o fim do exon 1 e nem com o começo do exon 2. Também não houve similaridade significativa entre o exon 2 da seqüência usada como referência com as espécies *Gallus gallus* e *Meleagris gallopavo*, exatamente as OTUs que formariam o grupo externo na análise.

Em relação ao exon 3, não foi encontrada similaridade significativa com a espécie *Meleagris gallopavo*. Em relação à *Gallus gallus* foi encontrada similaridade em apenas 85 nucleotídeos. A OTU *Ovis aries*, devido a sua seqüência estar depositada incompletamente, não apresenta a seqüência inteira dos exons 1 e 3.

Na Tab. 2 pode-se observar a delimitação dos exons segundo o programa BLAST, tomando como referência o sequenciamento feito no Capítulo 1. As medidas limites dos fragmentos referem-se à seqüência depositada no Genbank para cada espécie.

Três conjuntos de dados, um baseado na seqüência completa que codifica o mRNA do gene (Fig. 1) e os outros dois baseados nas seqüências que codificam os domínios funcionais HLH (Fig. 2) e Básico (Fig. 3) foram usados para gerar três árvores filogenéticas por meio de máxima verossimilhança para estas seqüências. O fragmento de mRNA representa a seqüência codificadora completa do gene, ou seja, a união dos três exons que formam o gene, incluindo os domínios Básico e HLH que pertencem ao exon 1.

Tabela 2 – Fragmentos que tiveram similaridade com a seqüência referência e seus devidos tamanhos, localizações e números de acesso ao Genbank, segundo o programa BLAST.

Espécie	Acesso Genbank	Exon 1*	Exon 2*	Exon 3*
<i>H. sapiens</i>	(AF050501)	1177 → 1647	2450 → 2530	3155 → 3247
<i>M. musculus</i>	(M95800)	1619 → 2089	2604 → 2684	3212 → 3334
<i>R. norvegicus</i>	(M24393)	35 → 505	506 → 584	587 → 708
<i>B. taurus</i>	(AB110600)	3 → 467	474 → 554	555 → 677
<i>O. aries</i>	(AF433651)	11 → 403	410 → 490	491 → incompleto
<i>M. gallopavo</i>	(AY560111)		Sem similaridade significativa	Sem similaridade significativa
<i>G. gallus</i>	(D90157)		Sem similaridade significativa	

* Os nucleotídeos se referem ao início e ao final da similaridade em cada uma das espécies analisadas.

Procedeu-se a análise dos domínios Básico e HLH separadamente em razão do motivo Básico ser o resíduo da proteína responsável pela interação com o DNA e o motivo HLH pela ligação entre os monômeros de bHLH para formarem dímeros funcionais. Há grande chance de que alterações nesses fragmentos possam levar a modificação na eficiência do funcionamento da proteína.

Essa análise diferenciada é aconselhada quando há informações prévias que permitam hierarquizar os sítios em classes, em que, por sua vez, são esperadas existirem diferentes pressões seletivas e dessa maneira diferentes taxas Ka/Ks (Yang e Nielsen, 2002).

O alinhamento das seqüências usadas na construção das árvores para os fragmentos mRNA, domínio HLH e domínio Básico são mostrados nas Fig. 1, 2 e 3, respectivamente.

	1							70
Piau	ATGGAGCTGT	ATGAGACATC	CCCCTACTTC	TACCAGGAAC	CCCCTTCTA	TGACGGGGAA	AACTACCTGC	
<i>M_musculus</i>T..G.T.....T.	
<i>R_norvegicus</i>A.....T..G.T.	
<i>O_aries</i>	-----	-----	-----	-----	-----	-----	--...AA..	
<i>B_taurus</i>C..	T.....	..T.....	..T.....G	
<i>H_sapiens</i>G.T.....	
<i>G_gallus</i>C.	T.....CAA	..T.....T	..T..C...G	AGAGG..T.	C..T.....T...G	
<i>M_gallopavo</i>C.	T.....CAA	..T.....T	..T..C...G	AGAGG..T.	C..T.....T...G	
	71							140
Piau	CCGTCCACCT	CCAGGGCTTT	GAGCC-ACCA	GGCTAC--GA	GCGGACTGAG	CTGAGTCTGA	GCCCTGAGGC	
<i>M_musculus</i>	.T.....	T.....C-C..GT--..C..CT.A.G..A..	
<i>R_norvegicus</i>	.T.....	T.....C-C..GT--..C...	..C..CT.A.G..A..	
<i>O_aries</i>	TT.....-G..--..G...	..C..C...	
<i>B_taurus</i>	.T.....-G..--A.G...	..C..C...	
<i>H_sapiens</i>	.T.....C	..A.-...--..G...	..C..CC...C.....	
<i>G_gallus</i>	G.TC..G.T.	G.....AC	..G.GG..G	C.T.T.CT..	..TC.C...	G...CC...TAAG	
<i>M_gallopavo</i>	G.TC..G.T.	G...A..A.	..G.AG..G	CAT.T.CT..	..C.C..A	G..GC...TC..AAG	
	141							210
Piau	CCGAGTGCCC	CTGGAAGATA	AGGGGCTGGG	GACCCCGAG	CACTGCCAG	GCCAGTGCCT	GCCGTGGGCA	
<i>M_musculus</i>G...A.	...A.....T..	..T..T...	
<i>R_norvegicus</i>G...	..A.....G.	...A.....T..	..T..T...C.....G	
<i>O_aries</i>	T.....	..T....C.	...T.....	..C.TG.G..G.G	
<i>B_taurus</i>	T.....	..T....C.	...T.....	..C.TG.G..G.G	
<i>H_sapiens</i>	..C..G...	..T..G..C.T..G	
<i>G_gallus</i>	.A...G.G.T	T...G..G.	..ACTC.AC	.CTG.....C.	.G..A...T.	..A...T	
<i>M_gallopavo</i>	.A...G.G.T	T...G..G.	..ACTCAAT	.CTA.....T.	.G..A...T.	A.....T	
	211							280
Piau	TGTAAGGTGT	GTAAGAGGAA	GTCCGTGTCT	GTGGACCGTC	GCGGGCCGC	CACGCTGAGG	GAGAAGCGCA	
<i>M_musculus</i>T.....GGA	..A...A..	..A.....	
<i>R_norvegicus</i>T.....G.	..A...A..	..A.....	
<i>O_aries</i>G.....G.C.....A	
<i>B_taurus</i>G.....G.C.....A	
<i>H_sapiens</i>G.....CG.G.	..A.....	
<i>G_gallus</i>	..C..AA.C.	.C...C.C..	AA.....C	A.C...G.	T...G..C.G.	
<i>M_gallopavo</i>	..C..AA.C.	.C...C.C..	AA.....C	A.C...G.	T...G..C.G.	
	281							350
Piau	GGCTCAAGAA	GGTGAATGAG	GCCTTGTAGG	CCCTGAAGAG	GAGCACCTG	CTCAACCCCA	ACCAGCGGCT	
<i>M_musculus</i>	A.....C.....	
<i>R_norvegicus</i>	A.....C.....	..T.....	A.....	
<i>O_aries</i>	.A.....AC.....	..C.....A.....	
<i>B_taurus</i>	.A.....A	.G...CC..	..C.....A.....	
<i>H_sapiens</i>C.....	A.....	
<i>G_gallus</i>	...G.....C..AC.....	..T...AC.	C...T..	
<i>M_gallopavo</i>	...G..A..C..AC.....	..G.....C.	C...T..A.....	
	351							420
Piau	GCCCAAGGTG	GAGATCCTGC	GCAGGCCAT	CCAGTACATC	GAGCGCCTGC	AGGCCCTGCT	CAGCTCCCTC	
<i>M_musculus</i>	...T..A...TA.T.....	
<i>R_norvegicus</i>	...T.....T.....TA.T.....	
<i>O_aries</i>A...TA	
<i>B_taurus</i>A...A	
<i>H_sapiens</i>T.....C.	
<i>G_gallus</i>AG.AG.	
<i>M_gallopavo</i>AG.AG.	
	421							490
Piau	AACCAGGAGG	AGCGAGACC-	-----TC	CGCTACCGAG	GCGGGGCGG	GCCGCAGCCA	GGGGTGCCCA	
<i>M_musculus</i>C..T.-	-----A...C...C	AT.....	
<i>R_norvegicus</i>C..T.-	-----C...G	..T..A...	
<i>O_aries</i>C.....	-----G	A..C...G.G	..C.....	
<i>B_taurus</i>C.....	-----G	A..C...G.G	..C.....	
<i>H_sapiens</i>T.....	-----GC.....AG	
<i>G_gallus</i>C...	...T..G.A	GAGGGAGC.GG--	..CCC.CT.C	A..A..A..T	..CT.CA...	
<i>M_gallopavo</i>C...	...C..G.A	GAGGGAGT.GG--	..CCT.CT.C	A..A..A..T	..CT.CA...	

```

491
Piau      GTGAATGCAG TTCCCACAGC GCCTCCTGCA GTCCAGAATG GGGCAGTGCA CTGGAGTTCG GCCCCAACCC
M_musculus .....A C..... .G..G.. ..A..... .T.....
R_norvegicus .....A C..... .G..G.. ..A..... .T.....
O_aries ..... C....T... ..C.G.. ..C... ..T.....
B_taurus ..... C....T... ..C.G.. ..C... ..T.....
H_sapiens CGA..... C..T..... ..G..... ..A..G.....
G_gallus .C..G...G. C..TGG... T.A..... .C..T..G.. .A...CCCAG .....T. .A.....
M_gallopavo ....G...G. C..TGG... T.G..... .C..C..G.. .A...CACAG .....A ..A.....

561
Piau      AGGGGATCAT CTGCTCACAG CTGACCCTAC AGATGCCAC AATCTGCACT CCCTCACCTC CATCGTGGAC
M_musculus ...A..... T....G.G. ....C..... ..T..G.. .....
R_norvegicus ...A..... T....G... ..G..... .C..... ..T..G.. .....
O_aries .....C... .G..... .C..... -----
B_taurus .....C... .G..... .C..... -----
H_sapiens .....G..... .C..... -----
G_gallus C.CA.....C ..C..G.GC. A....AGG. ...G.A..G. .C..C.... .G...T.... .....G
M_gallopavo T.C.....C ..C..G.GT. A...TG.AG. ...G.A..G. .C..C.... .T...T.... T....T..G

631
Piau      AGCATCACAG TGGAGGATGT GGCTGTGGCC TTCCCAGATG AAACCATGCC CAACTGA
M_musculus .....G. ....A. .T...T... ..C..... .....
R_norvegicus .....G. ....A. .T...CA.. ..C..... .....
O_aries -----
B_taurus ..... C...C.... ..A... A.....
H_sapiens .....A..... .T..... -----
G_gallus .....G.C. ....C... .C...A.G .....G. .GCGGG.C.A A.....
M_gallopavo .....G.C. ....C... .C...A.G .....G..G. .GCGGG.C.A A.....

687

```

Figura 1 – Alinhamento dos fragmentos das OTUs referente ao mRNA transcrito do gene Miogenina.

```

1
Piau      GTGAATGAGG CCTTTGAGGC CCTGAAGAGG AGCACCTGC TCAACCCCAA CCAGCGGCTG CCCAAGGTGG
B_taurus .....A. G...CC.... .C..... ..A..... ..A....
O_aries .....A. ....C.... .C..... ..A..... ..A....
R_norvegicus .....C..... T.....A ..... ..T.....
M_musculus .....C..... ..T..A....
H_sapiens .....C..... .A..... ..T.....
G_gallus .....C.A. ....C.... T....AC.C .....T.....
M_gallopavo .....C.A. ....C.... G.....C.C .....T..... ..A.....

71
Piau      AGATCCTGCG CAGCGCCATC CAGTACATCG AGCGCCTGCA GGCCCTGCTC AGCTCC
B_taurus .....A..... ..A..... ..T.....
O_aries .....T..... ..A..... ..T.....
R_norvegicus .....T..... ..T..... .A... .T.....
M_musculus .....C..... ..T..... .A... .T.....
H_sapiens .....T..... ..C..... ..T.....
G_gallus .....C..... ..AG..... ..AG.
M_gallopavo .....C..... ..AG..... ..AG.

126

```

Figura 2 - Alinhamento dos fragmentos das OTUs referente ao domínio HLH do gene Miogenina.

```

1
Piau ATGGAGCTGT ATGAGACATC CCCCTACTTC TACCAGGAAC CCCACTTCTA TGACGGGGAA AACTACCTGC 70
M_musculus .....T.....G.....T.....T.....T.....T.....
R_norvegicus .....A.....T.....G.....T.....T.....T.....T.....
O_aries -----
B_taurus .....C..T.....T.....T.....G.....G.....
H_sapiens .....G.....T.....T.....T.....T.....T.....T.....
G_gallus .....C.T.....CAA...T.....T.T..C...G. AGAGG..T..C..T.....T...G
M_gallopavo .....C.T.....CAA...T.....T.T..C...G. AGAGG..T..C..T.....T...G

71
Piau CCGTCCACCT CCAGGGCTTT GAGCC-ACCA GGCTAC--GA GCGGACTGAG CTGAGTCTGA GCCCTGAGGC 140
M_musculus .T.....T.....C.....-C..G.....T--.. .....C..CT.A.....G..A..
R_norvegicus .T.....T.....C.....-C..G.....T--.. .....C...C..CT.A.....G..A..
O_aries TT.....-G.....-G.....-A.....G.....C..C.....
B_taurus .T.....-G.....-A.....G.....C..C.....
H_sapiens .T.....C..A.-.....-.....G.....C..CC.....C.....
G_gallus G.TC..G.T.G.....AC...G.GG..G.C.T.T.CT...TC.C...G...CC...T.....AAG
M_gallopavo G.TC..G.T.G...A..A...G.AG..G.CAT.T.CT...C.C..A.G..GC...T.....C..AAG

141
Piau CCGAGTGCCC CTGGAAGATA AGGGGCTGGG GACCCCGAG CACTGCCCAG GCCAGTGCCT GCCGTGGGCA 210
M_musculus .....G.....A.....A.....T...T..T.....
R_norvegicus .....G.....A.....A.....T...T..T.....
O_aries -----
B_taurus T.....T.....C.....T.....C.TG.G.....G.....G.....
H_sapiens ..C..G.....T..G..C.....T.....T.....T.....T.....G.....
G_gallus .A...G.G.T.T...G..G...ACTC.AC.CTG.....C..G..A...T...A...T
M_gallopavo .A...G.G.T.T...G..G...ACTCAAT.CTA.....T..G..A...T.A.....T

211
Piau TGTAAGGTGT GTAAGAGGAA GTCCGTGTCT GTGGAC 246
M_musculus .....T.....G.....
R_norvegicus .....T.....
O_aries -----
B_taurus .....G.....
H_sapiens .....G.....C.....
G_gallus ..C..AA.C..C...C.C..AA.....C.A.C...
M_gallopavo ..C..AA.C..C...C.C..AA.....C.A.C...

```

Figura 3 - Alinhamento dos fragmentos das OTUs referente ao domínio Básico do gene Miogenina.

Com o alinhamento das seqüências pode-se avaliar quais parâmetros devem ser considerados na construção da árvore filogenética para cada fragmento. Os parâmetros para cada fragmento estudado foram estimados utilizando-se o programa Model Test 3.5, permitindo selecionar os seguintes modelos:

a-) Fragmento mRNA

Modelo HKY85 + G - Considera que substituições do tipo transição ocorrem mais freqüentemente que do tipo transversão e que as freqüências dos quatro nucleotídeos nem sempre é similar. Também considera que sítios de substituição possuem distribuição gama.

Frequência das bases: $f(A) = 0,2080$
 $F(C) = 0,3227$
 $F(G) = 0,3034$
 $F(T) = 0,1659$

Modelo de substituição:

$$Ti/Tv \text{ (transições/transversões)} = 2,0734$$

Taxa de variação entre sítios

$$\text{Proporção de sítios invariáveis} = 0$$

Sítios variáveis (G)

$$\text{Parâmetro } a \text{ de variação gama} = 0,3548$$

b-) Fragmento Domínio Básico

Modelo: K80 + G - Considera que substituições do tipo transição ocorrem mais frequentemente que do tipo transversão e os sítios de substituição possuem distribuição gama.

Frequências das bases: Frequências equivalentes

Modelo de substituição:

$$Ti/Tv \text{ (transições/transversões)} = 2,1344$$

Taxa de variação entre sítios

$$\text{Proporção de sítios invariáveis} = 0$$

Sítios variáveis (G)

$$\text{Parâmetro } a \text{ de variação gama} = 0,5381$$

c-) Fragmento Domínio HLH

Modelo: K80 + G - Considera que substituições do tipo transição ocorrem mais frequentemente que do tipo transversão e os sítios de substituição possuem distribuição gama.

Frequência das bases: Frequências equivalentes

Modelo de substituição:

$$Ti/Tv \text{ (transições/transversões)} = 1,8380$$

Taxa de variação entre sítios

$$\text{Proporção de sítios invariáveis} = 0$$

Sítios variáveis (G)

Parâmetro a de variação gama = 0,1076

Normalmente, os evolucionistas estão interessados na árvore filogenética que representa a história evolutiva de um grupo de espécies. Entretanto, quando a árvore é construída apenas com um gene, ela pode não representar estritamente a história evolutiva da espécie, mas sim a do gene, chamada de árvore gênica.

Nesse caso, todas as três árvores mostraram exatamente a mesma topologia, como pode ser observado na Fig. 4, com o grupo externo formado pelas OTUs *G. gallus* e *M. gallopavo*, como foi pré-estabelecido.

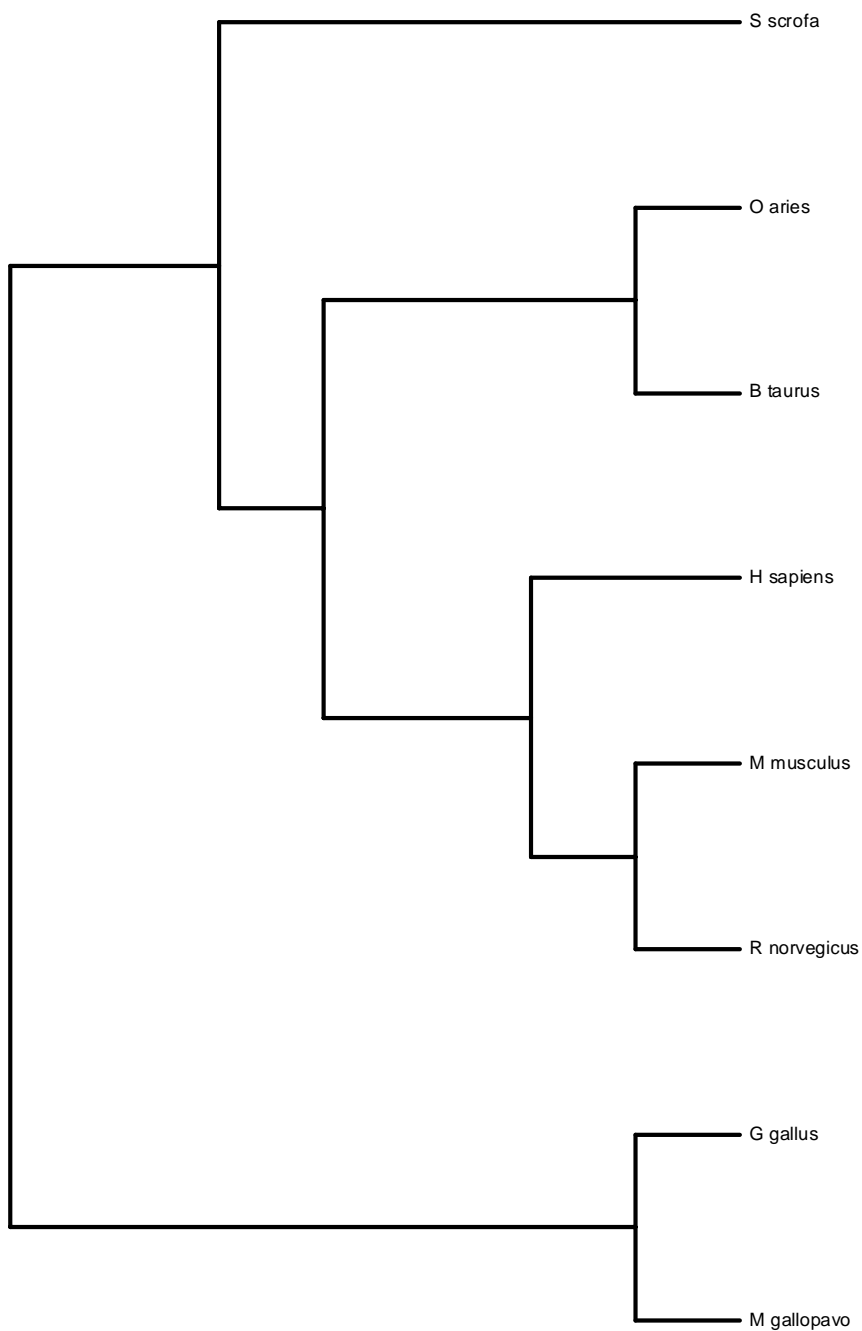


Figura 4 – Topologia coincidente das árvores gênicas para os três fragmentos estudados do gene da Miogenina. As árvores foram obtidas pelo método de máxima verossimilhança com o programa PAUP (Swofford, 1999).

Como pode ser observado na árvore gênica, as OTUs mais proximamente relacionadas são *B. taurus* e *O. aries*, formando o grupo monofilético Ruminantia, e *M. musculus* e *R. norvegicus* formando o grupo monofilético Rodentia. Essa árvore gênica não representa a história evolutiva das espécies já que, apresenta a OTU *S. scrofa* formando um grupo monofilético com Rodentia. A árvore filogenética contendo essas OTUs apresentaria o *S. scrofa* e Ruminantia formando um grupo monofilético chamado Artiodactila (<http://tolweb.org/tree/phylogeny.html>, acessado em 07/12/2004).

Com a árvore gênica pronta, a evolução das seqüências foi analisada para o gene da Miogenina, a partir do método PBLSB (Benner *et al.*, 1998), para a reconstrução dos estados dos caracteres ancestrais juntamente com um método para contagem de sítios de substituições sinônimas e não sinônimas.

A taxa Ka/Ks é uma ferramenta valiosa na detecção de pressões de seleção entre espécies proximamente relacionadas. Uma taxa mais elevada de Ka em relação à Ks indica um regime de seleção positiva onde existem espécies ou populações que apresentam modificações evolutivas significativas (adaptativas) em relação aos grupos que não contém essa taxa positiva.

Tradicionalmente, a seleção positiva é definida por valores da razão Ka/Ks significativamente maiores que um, sendo que valores entre 0,6 e 1 podem ocorrer devido ao relaxamento de restrição funcional, já que banco de dados como TAED consideram proteínas com esses valores como sendo candidatos à evolução adaptativa, pois os valores são relativamente altos quando comparados com a média das outras proteínas. Na prática existem valores contínuos que vão entre uma pressão de seleção fortemente negativa (onde apenas um único aminoácido seria para que a proteína seja funcional) passando por uma pressão negativa (por exemplo, apenas 5 aminoácidos hidrofílicos seriam viáveis) até a neutra (onde os 20 aminoácidos se enquadrariam equivalentemente).

Nas Tab. 3, 4 e 5 são mostradas os valores de Ka, Ks e da razão Ka/Ks obtidos nos ramos das árvores filogenéticas para cada fragmento estudado. Os nós e os ramos pertencentes a cada valor das tabelas podem ser visualizados respectivamente nas Fig 5, 6 e 7.

Primeiramente, pode-se notar o nível de conservação que o gene apresenta e, em todos os fragmentos (Tab. 3, 4 e 5 e Fig 5, 6, e 7), parece estar sob pressão de

seleção positiva ($Ka/Ks = 2,4225$) o ramo que leva a espécie *O. aries*, exatamente a OTU que não apresentava a seqüência completa.

No fragmento referente ao mRNA, todos os ramos, a não ser os do grupo Ruminantia (ramos do nó 4, na Tab. 3), apresentaram um valor de Ka/Ks inferior a 0,3 e a grande maioria com valores menores que 0,1. Isso indica que a proteína gerada pelo gene analisado (Miogenina), selecionada por milhões de anos, atingiu sua função otimizada previamente a divergência do grupo interno e do grupo externo. Isso implica em uma pressão de seleção negativa ou evolução conservativa, onde a maioria das mutações não sinônimas era prejudicial à adaptação das espécies. Dentro do grupo Ruminantia, o ramo que leva a OTU *Bos taurus* apresentou um valor de 0,37 (ramo 1 do nó 4, Tab. 3) e o que leva a *Ovis aries* de 2,42 (ramo 2 do nó 4, Tab. 3).

A mesma conclusão pode ser feita para os outros dois fragmentos. Com relação ao domínio Básico, novamente o ramo que leva a OTU *O. aries* se mostrou sob pressão de seleção positiva, com um valor da razão Ka/Ks de 1,68. Já o domínio HLH, a seleção apresentou-se muito mais conservativa, provavelmente porque esse foi o único fragmento que a OTU *O. aries* apresentou seqüência completa. Para esse domínio, quase todos os ramos apresentaram uma razão Ka/Ks igual a zero, com exceção dos ramos que levam as OTUs *O. aries* e *B.taurus* que apresentaram razões elevadas de 13,15 e 19,57, respectivamente. Tais taxas indicam que as espécies analisadas do grupo Ruminantia estiveram sob grande pressão de seleção positiva.

Tabela 3 - Valores de Ka, Ks e da razão Ka/Ks obtidos nos ramos das árvores filogenéticas para o mRNA relativo ao gene Miogenina usando o método PBLSB.

Nó	Ka/Ks	Ramo 1		Ramo 2		
		Ka	Ks	Ka/Ks	Ka	Ks
1	0,1052	0,0101	0,0965	0,0384	0,0031	0,0800
2	0,0885	0,0048	0,0547	0,0400	0,0024	0,0600
3	0,2392	0,0207	0,0864	0,0928	0,0157	0,1692
4	0,3760	0,0070	0,0187	2,4225	0,0092	0,0038
5	0,0879	0,0062	0,0701	0,1227	0,0191	0,1558
6	0,0278	0,0015	0,0548	0,0472	0,0031	0,0653
7	0,2903	0,0816	0,2811	0,2632	0,0813	0,3090

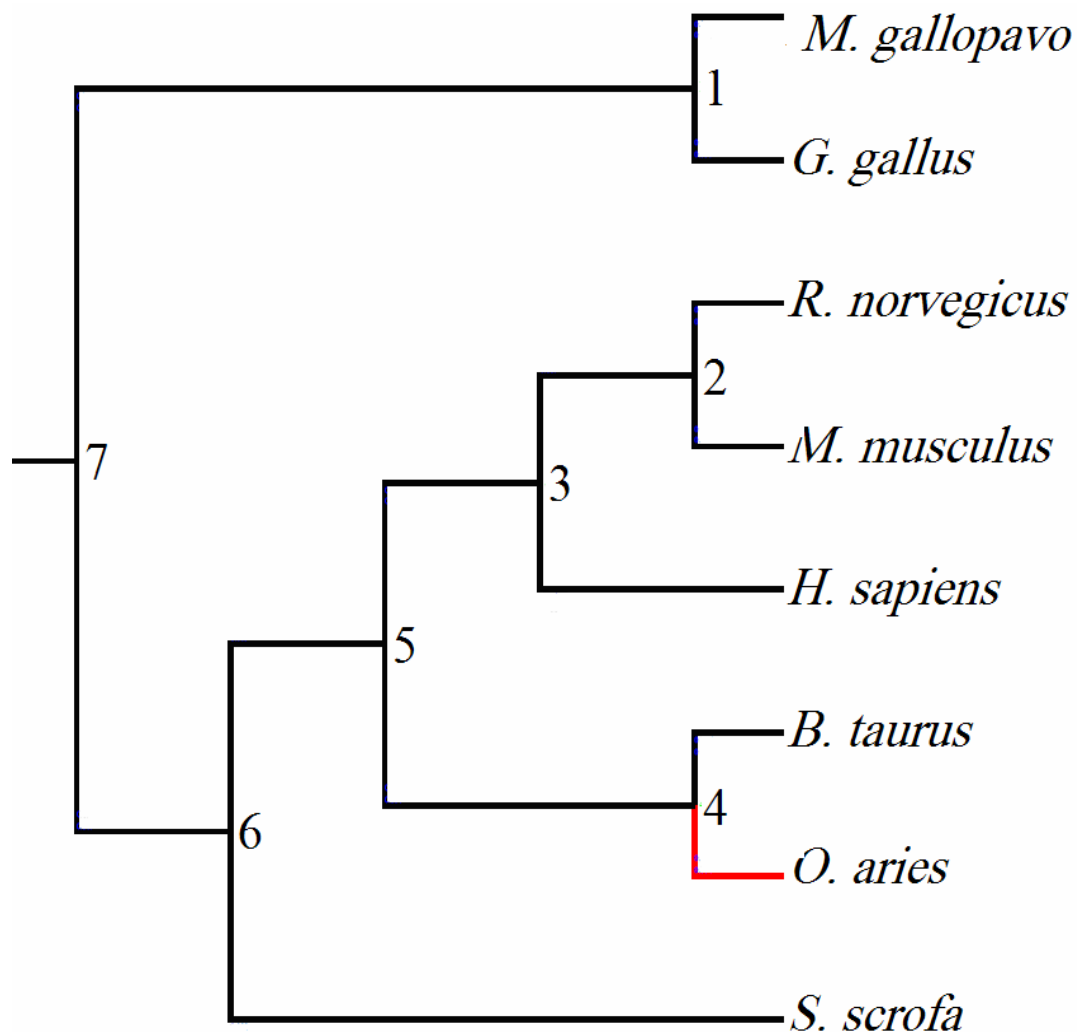


Figura 5 – Árvore gênica do fragmento relativo ao mRNA do gene da Miogenina e seus valores Ka e Ks obtidos usando o método PBLSB (Benner *et al.*, 1998). As árvores foram geradas pelo método de Máxima Verossimilhança.

Tabela 4 - Valores de Ka, Ks e da razão Ka/Ks obtidos nos ramos das árvores filogenéticas para o domínio Básico do gene Miogenina usando o método PBLSB.

Nó	Ka/Ks	Ramo 1		Ramo 2		
		Ka	Ks	Ka/Ks	Ka	Ks
1	0,1774	0,0165	0,0931	0,0406	0,0033	0,0802
2	0,1580	0,0133	0,0844	1,6893	0,0482	0,0285
3	0	0	0,0709	0	0	0,0588
4	0,2294	0,0205	0,0895	0,0642	0,0133	0,2069
5	0,1252	0,0236	0,1882	0,0966	0,0073	0,0754
6	0,0821	0,0023	0,0277	0,0445	0,0060	0,1340
7	0,3710	0,1212	0,3266	0,2964	0,1150	0,3881

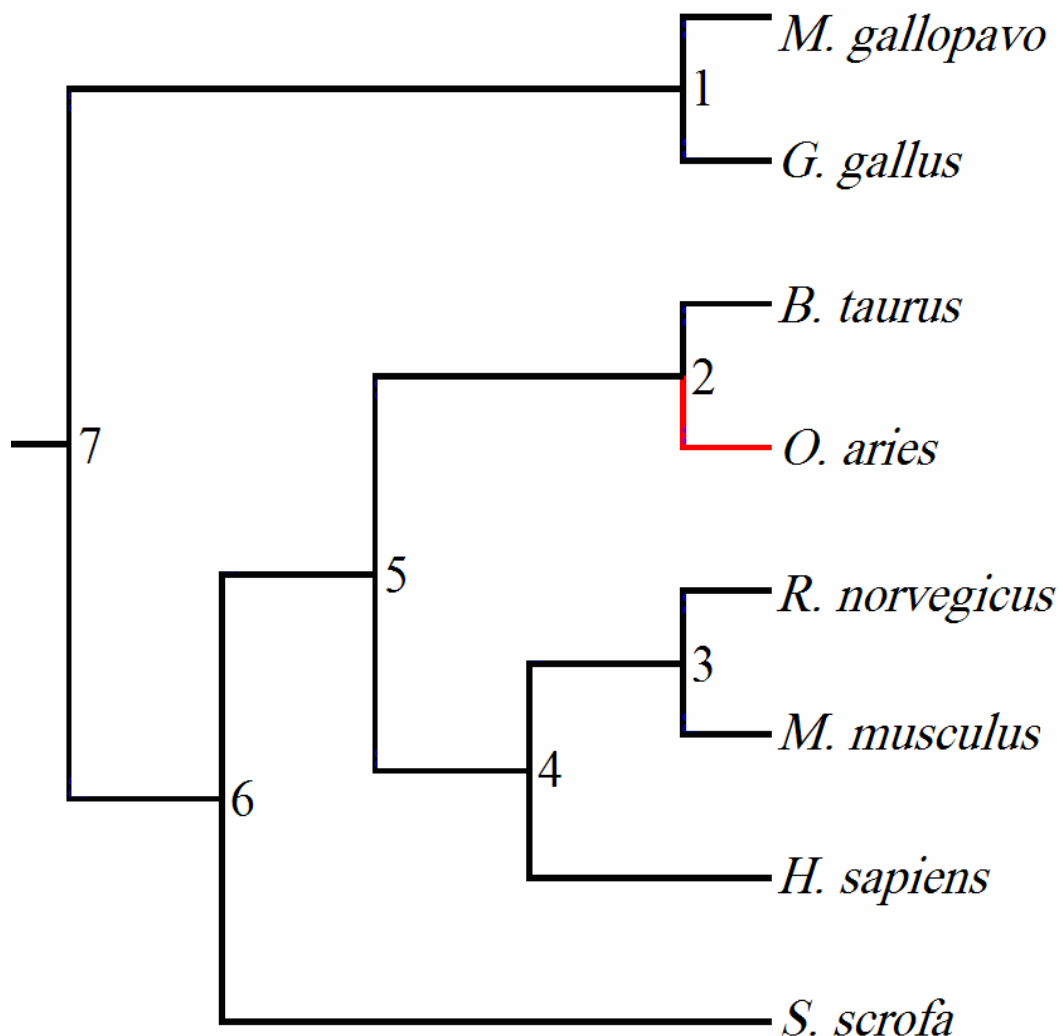


Figura 6 – Árvore gênica do fragmento relativo ao domínio Básico do gene Miogenina e seus valores Ka e Ks obtidos usando o método PBLSB (Benner et al., 1998). As árvores foram geradas pelo método de Máxima Verossimilhança.

Tabela 5 - Valores de Ka, Ks e da razão Ka/Ks obtidos nos ramos das árvores filogenéticas para o domínio HLH do gene Miogenina usando o método PBLSB.

Nó	Ramo 1			Ramo 2		
	Ka/Ks	Ka	Ks	Ka/Ks	Ka	Ks
1	0	0	0,0496	0	0	0,0461
2	13,1587	0,0131	1e-10	19,5704	0,0196	1e-10
3	0	0	0,0432	0	0	0,0417
4	0	0	0,0522	0	0	0,1122
5	0	0	0,1290	0	0	0,0891
6	0	0	0,0102	0	0	0,0244
7	0,2073	0,0218	0,1051	0,2577	0,0219	0,0852

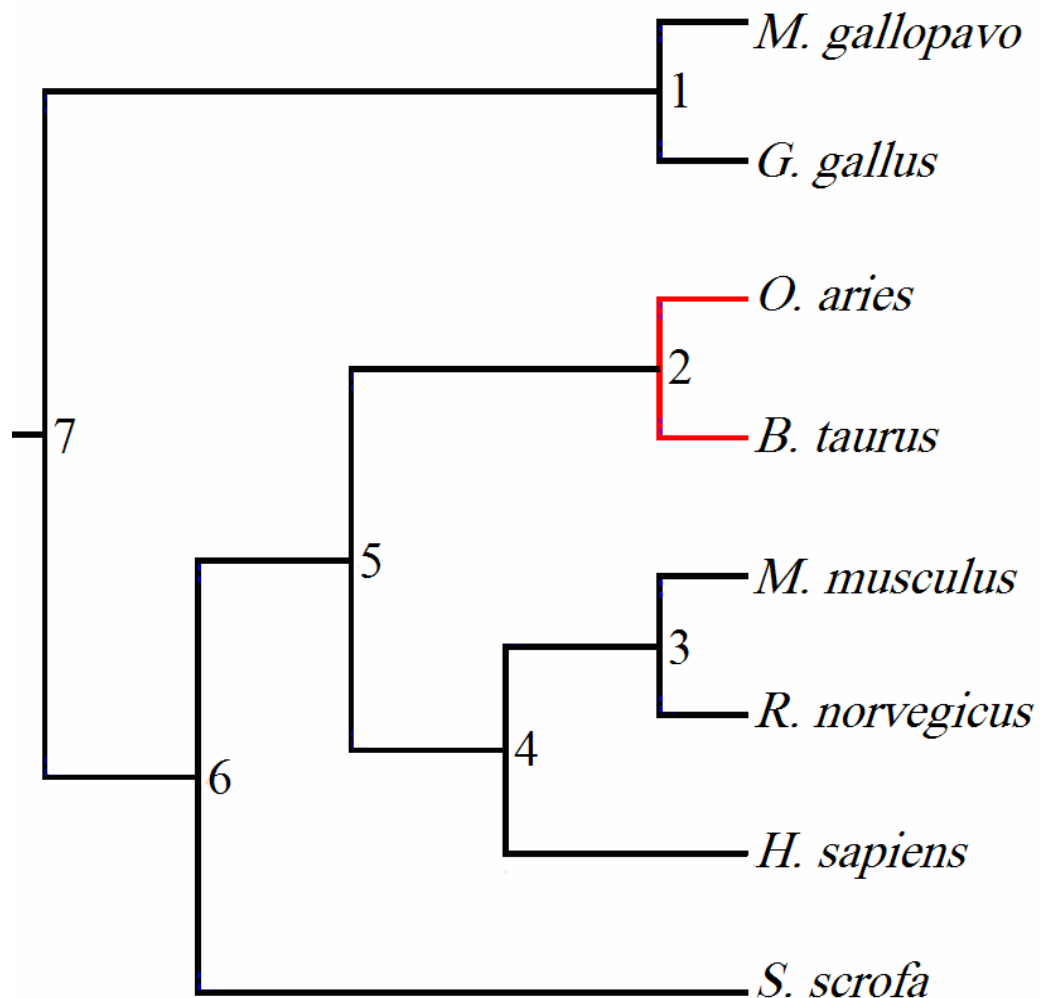


Figura 7 – Árvore gênica do fragmento relativo ao domínio HLH do gene Miogenina e seus valores Ka e Ks obtidos usando o método PBLSB (Benner *et al.*, 1998). As árvores foram geradas pelo método de Máxima Verossimilhança.

A quantidade de informações relativas ao gene da Miogenina ainda é muito escassa em banco de dados como o Genbank, com um número muito pequeno de espécies com o gene seqüenciado. Isso prejudica a realização de um delineamento metodológico mais desejável, que incluísse outras raças de suínos selvagens ou parentes próximos como o javali. Espécies com informações muito distantes umas das outras podem gerar uma perda na qualidade da reconstrução dos estados ancestrais, especialmente quando o tamanho do ramo sob estudo aumenta, já que isso leva a uma diluição de um episódio de adaptação positiva (com $Ka/Ks > 1$) com episódios de evolução conservativa.

Apesar dessa distância entre as OTUs, é importante ressaltar que a evolução do gene da Miogenina ocorre em uma dinâmica diferente da evolução das espécies. A maneira conservativa que o gene evolui faz com que as espécies acabem sendo proximamente relacionadas quando considera-se a seqüência de nucleotídeos do gene da Miogenina, ou seja, os caracteres que forneceram as informações para a construção das árvores gênicas. Além disso, métodos baseados em parcimônia podem superar os métodos baseados em máxima verossimilhança na análise de seqüências proximamente relacionadas, mas são conhecidos por subestimar a quantidade de substituições em comparações mais distantemente relacionadas (Yang e Nielsen, 2002).

Esses resultados devem ser testados e confirmados com outros testes, já que a reconstrução das seqüências de ancestrais não é sempre acurada (até mesmo quando são usados critérios de máxima parcimônia), usando um delineamento mais apropriado para esse tipo de estudo, permitindo analisar mais precisamente a história evolutiva do gene da Miogenina.

CONCLUSÕES

Com exceção do ramo levando ao grupo Ruminantia, o gene da Miogenina evolui sobre uma seleção conservativa pelo menos desde a divergência dos mamíferos com as aves. Sobre a evolução do gene no ramo que leva as espécies *Ovis aries* e *Bos taurus* o gene sofreu polimorfismos no domínio HLH que possivelmente promoveram uma maior adaptabilidade desses organismos a seus ambientes.

Os resultados constatam que é pequena a possibilidade de encontrar variantes no gene da Miogenina para usá-lo em programas de melhoramento de suínos

LITERATURA CITADA

- ALTSCHUL, S.F.; MADDEN, T.L.; SCHAFFER, A.A. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, v.25(17), p.3389-3402, 1997. Disponível em: <http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>. Acesso em: 13 set. 2004.
- BENNER, S.A.; TRABESINGER, N.; SCHREIBER, D. Post genomic science: converting primary structure into physiological function. *Adv. Enzyme Regul.*, v.38, p. 155-180, 1998.
- COUTINHO, L.L.; GABRIEL, J.E.; ALVARES, L.E. Controle molecular do desenvolvimento da musculatura esquelética. In: SIMPÓSIO INTERNACIONAL DE GENÉTICA E MELHORAMENTO ANIMAL, 1999, Viçosa. Anais...Viçosa: UFV, 1999, p.355-376.
- EDMONDSON, D.G.; OLSON, E.N. A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. *Genes Dev.* v.3(5), p.628-640, 1989.
- FUJISAWA-SEHARA, A.; NABESHIMA, Y.; HOSODA, Y. *et al.* Myogenin contains two domains conserved among myogenic factors. *J. Biol. Chem.* 265 (25), p.15219-15223, 2002.
- JEANPLONG, F.; BASS, J.J.; SMITH, H.K. *et al.* Chronic underfeeding represses the activity of satellite cells and alters myostatin and IGFs. Não publicado. Disponível em: <http://www.ncbi.nlm.nih.gov/Entrez>. Acesso em: 27 jul. 2004.
- LI, W.H.; WU, C.I.; LUO, C.C. A new method for estimating synonymous and non synonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.*, v.2, p.150-174, 1985.
- LIBERLES, D.A.; SCHREIBER, D.R.; GOVINDARAJAN, S. *et al.* The Adaptive Evolution Database (TAED), *Genome Biology*, v.2(8), research 0028.1 – 0028.6, 2001.
- LIU, C.; VELLEMAN, S.G. 2004. Submetido em 26/02/2004 *Animal Sciences*. Disponível em: <http://www.ncbi.nlm.nih.gov/Entrez>. Acesso em: 27 jul. 2004.
- MUROYA, S.; NAKAJIMA, I.; CHIKUNI, K. Effect of myogenic regulatory factor siRNAs on differentiation of bovine skeletal muscle cells. Não

publicado. Disponível em: <http://www.ncbi.nlm.nih.gov/Entrez>. Acesso em: 27 jul. 2004.

- NEI, M.; GOJOBORI, T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.*, v.3, p.418-426, 1986.
- PAMILO, P.; BIANCHI, N.O. Evolution of the Zfx and Zfy genes: rates and interdependence between the genes. *Mol. Biol. Evol.*, v.10, p. 271-281, 1993.
- POSADA, D.; CRANDALL, K.A. Model test: testing the model of DNA substitution. *Bioinformatics*, v.14, p.817-818, 1998.
- SWOFFORD, D.L. PAUP*: Phylogenetic Analysis Using Parsimony (* and Others Methods), versão 4.0. Sunderland: Sinauer Associates Inc. USA. 1999.
- TELLGREN, Å.; BERGLUND, A.C.; SAVOLAINEN, P. *et al.* Myostatin rapid sequence evolution in ruminants predates domestication. *Molecular Phylogenetics and Evolution*, v.33, p.782-790, 2004.
- THOMPSON, J.D.; GIBSON, T.J.; PLEWNIAK, F.; JEANMOUGIN, F.; HIGGINS, D.G. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, v.24, p.4876-4882, 1997.
- TSENG, B.S.; CAVIN, S.T.; HOFFMAN, E.P. *et al.* Human bHLH transcription factor gene myogenin (MYOG): genomic sequence and negative mutation analysis in patients with severe congenital myopathies. *Genomics*, v.57(3), p.419-423, 1999.
- UZZEL, T.; CORBIN, K.W. Fitting discrete probability distribution to evolutionary events. *Science*, v.172, p.1089-1096, 1971.
- YANG, Z.; NIELSEN, R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.*, v.17, p.32-43, 2000.
- YANG, Z.; NIELSEN, R. Codon-Substitution Models for Detecting Molecular Adaptation at Individual Sites Along Specific Lineages. *Molecular Biology Evolution*, v.19(6), p.908-917, 2002.
- WRIGHT, W.E.; SASSOON, D.A.; LIN, V.K. Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. *Cell*, v. 56 (4), p.607-617, 1989.