**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Identificação de polimorfismos em região do cromossomo 3 da
galinha associado ao desempenho de deposição de gordura**

Gabriel Costa Monteiro Moreira

Dissertação apresentada para obtenção do título de
Mestre em Ciências. Área de concentração: Ciência
Animal e Pastagens

**Piracicaba
2014**

Gabriel Costa Monteiro Moreira
Engenheiro Agrônomo

**Identificação de polimorfismos em região do cromossomo 3 da galinha
associado ao desempenho de deposição de gordura**
versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:
Prof. Dr. **LUIZ LEHMANN COUTINHO**

Dissertação apresentada para obtenção do título de
Mestre em Ciências. Área de concentração: Ciência
Animal e Pastagens

Piracicaba
2014

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - ESALQ/USP**

Moreira, Gabriel Costa Monteiro

Identificação de polimorfismos em região do cromossomo 3 da galinha associado ao desempenho de deposição de gordura / Gabriel Costa Monteiro Moreira. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2014.

83 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2014.
Bibliografia.

1. Gordura abdominal 2. INDEL 3. QTL 4. Sequenciamento de nova geração 5. SNP
I. Título

CDD 636.513
M872i

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

**À Marina Monteiro (*in memoriam*) e Wilson D'Amarante (*in memoriam*),
à Maria Lúcia Costa Moreira e Carlos Augusto Moreira (*in memoriam*), dedico!**

AGRADECIMENTO

A vida em uma pequena caixa de fósforos!

É tempo de agradecer a todas as lembranças que hoje tenho...

É noite, risco o primeiro fósforo e agradeço à Deus pela presença, força e permissão; pela fé e renovação. Aproveito a luz e agradeço à Nossa senhora, mãe de amor, carinho e proteção. Agradeço ao meu incansável guia: Salve Jorge! A chama? Ainda acesa! Mas olho para a caixa e vejo que ainda tenho fósforos, lembranças, agradecimentos.

Risco o segundo, e uma grande chama se estabelece... é ela, a família. Vejo rostos, e me pego agradecendo... lembrando de Rita, Luis e Ariane, lembrando de Dona Lúcia, Marina, Joana e Valdelice... Carlos e Wilson. Agradeço pelo sorriso despertado, pela doce lembrança.

Risco o terceiro...e lembro-me dos irmãos de alma que fiz em outras terras, em outras bandas... Eduardo de Carli, Isabella Pimentel e Camila Dourado! Parceria, companhia, decisão e cumplicidade.

O frio ainda insiste e aos poucos vou perdendo meus fósforos... tiro mais um e sem pestanejar, risco outra vez! Olho para o lado e vejo companhia... companhia de DE, de torcida, de aprendizado, vejo Flávia Rocha, Marília Ribeiro, Fabiane Silva, Daiane Fausto, Thaline Pacheli, Patrícia Maloso, Greicieli Moraes, Aliedson, Liliane (Kraxá).

Lembro-me das risadas, estudos e discussões com Minestra, Gregori, Jú, Joanita, Simone e Laiza e assim, invisto mais um palito no prazer de lembrar e agradecer aos membros do GEMA.

Pego mais um! Descubro um palito maior na minha caixa... indicativo de mais tempo de chama, mais tempo de luz, é hora de fortes e intensas lembranças. Risco este, pensando em todos do laboratório de Biotecnologia animal, uma das minhas muitas moradas nesse meu caminho. Agradeço emocionado ao professor Luiz Coutinho e à minha co-orientadora e companheira Clarissa Boschiero. Um turbilhão de agradecimentos me invade...muito obrigado pequenina (Tássia), menina Thaís, menina Ariana, menina Felício, Marcela, Lilian, Nirlei, Priscila, Carla, Sônia, Larissa, Verônica, Fabi, Berna, Gabi. Muito obrigado menino do baum (Fábio), Vinícius, Ricardo, Jorge, Gustavo, Ribamar, Millor. E no meio dessa chama, um agradecimento especial à Dênia Attílio pela amizade, acolhimento, ensinamentos, risadas e alegrias... aprendi muito com você! A chama reduz mas ainda me pego agradecendo à CAPES e à FAPESP.

É meu último palito... nele, lembranças de todos e todas que me fizeram sorrir, que me fizeram continuar, que ajudaram a acumular fósforos! Essa chama logo se apaga até que eu descubro mais um palito ao meu lado, mais uma chama, mais uma chance... e esse, eu dedico e agradeço a você leitor. Obrigado pela leitura, pela consideração!

“[...]”
Sigo meu caminho a musicar
Sinto que saí do canto
Do esconderijo do encanto
Do óbvio do apenas sonhar”

Gabriel Costa

SUMÁRIO

RESUMO.....	11
ABSTRACT.....	13
1 INTRODUÇÃO.....	15
2 REVISÃO BIBLIOGRÁFICA.....	17
2.1 Importância da avicultura.....	17
2.2 Deposição de gordura na carcaça: um parâmetro genético.....	18
2.3 Importância e caracterização do genoma da galinha.....	20
2.4 QTLs mapeados para deposição de gordura em galinhas.....	21
2.5 Genes associados à deposição e metabolismo de gordura.....	25
2.6 Variações genéticas e expressão das características fenotípicas.....	26
2.7 Mutações causais.....	28
2.8 Novas tecnologias de sequenciamento para a identificação de variações genéticas.....	30
2.9 Ferramentas para identificação de variações genéticas.....	33
3 OBJETIVOS.....	35
3.1 Objetivo Geral.....	35
3.2 Objetivos Específicos.....	35
4 MATERIAL E MÉTODOS.....	37
4.1 População experimental.....	37
4.2 Preparação de bibliotecas e sequenciamento.....	37
4.3 Região de QTL estudada.....	38
4.4 Descoberta de variações genéticas.....	39
4.5 Anotação funcional.....	39
5 RESULTADOS E DISCUSSÃO.....	41
5.1 Sequenciamento e alinhamento.....	41
5.2 Descoberta de SNPs e INDELs.....	41
5.3 Variantes genéticas homozigotas e heterozigotas.....	44
5.4 Anotação funcional.....	45
5.5 Análise de genes a partir dos SNPs deletérios.....	53
5.6 Vias metabólicas relacionadas às variantes localizadas em regiões exônicas.....	57
6 CONCLUSÕES.....	63

REFERÊNCIAS.....	65
APÊNDICE.....	79

RESUMO

Identificação de polimorfismos em região do cromossomo 3 da galinha associado ao desempenho de deposição de gordura

Dezoito galinhas de uma população experimental utilizada em um cruzamento recíproco entre as linhagens de frangos de corte (TT) e de postura (CC) foram sequenciadas pela tecnologia de nova geração na plataforma Illumina com uma cobertura média de 10X. A descoberta de variantes genéticas foi realizada em uma região de locos de característica quantitativa (*Quantitative Trait Locus*, QTL), associado anteriormente com peso e percentagem de gordura abdominal no cromossomo 3 da galinha (GGA3), entre os marcadores microssatélites *LEI0161* e *ADL0371* (33,595,706-42,632,651 pb). O programa SAMtools foi utilizado na identificação de 136.054 SNPs únicos e 15.496 INDELs únicas nos 18 animais sequenciados e após a filtragem das mutações, 92.518 SNPs únicos e 9.298 INDELs únicas foram mantidas. Uma lista de 77 genes foi analisada buscando genes relacionados ao metabolismo de lipídios. Variantes localizadas na região codificante (386 SNPs e 15 INDELs) foram identificadas e associadas com vias metabólicas importantes. Variantes nos genes *LOC771163*, *EGLN1*, *GNPAT*, *FAM120B*, *THBS2* e *GGPS1* foram identificadas e podem ser responsáveis pela associação do QTL com a deposição de gordura na carcaça em galinhas.

Palavras - chave: Gordura abdominal; INDEL; QTL; Sequenciamento de nova geração; SNP

ABSTRACT

Identification of polymorphisms in a region of chicken chromosome 3 associated with the performance of the fat deposition

Eighteen chickens from a parental generation used in a reciprocal cross with broiler and layer lines were sequenced by new generation technology with an average of 10-fold coverage. The DNA sequencing was performed by Illumina next generation platform. The genetic variants discovery was performed in a quantitative trait loci (QTL) region which was previously associated with abdominal fat weight and percentage in chicken chromosome 3 (GGA3) between the microsatellite markers *LEI0161* and *ADL0371* (33,595,706-42,632,651 bp). SAMtools software was used to detect 136,054 unique SNPs and 15,496 unique INDELs for the 18 chickens, and after quality filtration 92,518 unique SNPs and 9,298 unique INDELs were retained. One list of 77 genes was analysed and genes related to lipid metabolism were searched. Variants located in coding region (386 SNPs and 15 INDELs) were identified and associated with important metabolic pathways. Loss of functional variants in the genes *LOC771163*, *EGLN1*, *GNPAT*, *FAM120B*, *THBS2* and *GGPS1* may be responsible for the QTL associated with fat deposition in chicken.

Keywords: Abdominal fat; INDEL; Next generation sequencing; QTL; SNP

1 INTRODUÇÃO

Programas de melhoramento genético de aves são desenvolvidos para melhorar o desempenho e características de carcaça. Mais especificamente, o melhoramento genético de frangos de corte tem como objetivo aumentar a taxa de crescimento e melhorar a condição corporal (características de desempenho) (BERRI et al., 2001). As características de desempenho estão correlacionadas com a qualidade da carne e podem ter impactos negativos sobre a produção.

Linhagens de frangos de corte tendem a apresentar aumento na deposição de gordura (BAEZA e BIHAN-DUVAL, 2013). A presença de gordura na carne de frango diminui seu valor nutricional e consequentemente, seu valor comercial (JENNEN et al., 2004). A deposição de gordura em frangos é uma característica poligênica (RESNYK et al., 2013) e a descoberta de mutações potencialmente associadas com essa característica pode contribuir para reduzir os níveis de gordura na carne de frango.

O sequenciamento de nova geração (NGS) é uma ferramenta útil para estudos genômicos, pois permite a geração de uma quantidade maior de dados e a identificação de muitas mutações, como SNPs (*Single nucleotide polymorphism*) e pequenas inserções/deleções (INDELs), com um custo mais baixo e menor tempo em comparação aos métodos convencionais (CARVALHO e SILVA, 2010). O NGS também permite a descoberta de genes que apresentam mutações funcionais, e, em seguida, a identificação das vias metabólicas que podem ser relacionados com as características econômicas importantes, tais como a deposição de gordura.

O locus de característica quantitativa (*Quantitative Trait Locus*, QTL) é responsável por uma percentagem da variância fenotípica de uma determinada característica (WONG et al., 2004), e a identificação de polimorfismos nesta região permite o conhecimento da informação contida no código genético para posterior associação às características fenotípicas do animal; permite a busca de mutações causais.

De acordo com dados do ChickenQTL database (CHICKEN QTLDB, 2013) foram mapeados 174 QTLs para peso de gordura abdominal e 118 para o percentagem de gordura abdominal em frangos. Desse total, nove QTLs para percentagem de gordura abdominal e 11 QTLs para peso de gordura abdominal foram mapeados no cromossomo 3 da galinha (GGA3).

Em uma população experimental F₂ de galinhas, desenvolvida para estudos genômicos, Campos et al. (2009) mapearam QTLs associados a deposição de gordura

em uma região específica (entre os marcadores *LEI0161* e *ADL0371*) no cromossomo 3 da galinha (GGA3). QTLs para a deposição de gordura também foram mapeados em outras populações nesta mesma região por Park et al. (2006); Nadaf et al. (2009); Wang et al. (2012).

O objetivo deste estudo foi ressequenciar com alta cobertura, por meio da tecnologia de nova geração, 18 indivíduos de duas linhagens distintas (corte e postura) da população experimental F_2 da EMBRAPA para identificar SNPs e INDELs em uma região de QTL no cromossomo 3 da galinha que foi associada anteriormente com a deposição de gordura (CAMPOS et al., 2009).

2 REVISÃO BIBLIOGRÁFICA

2.1 Importância da avicultura

O modelo industrial de produção avícola se estabeleceu no cenário brasileiro a partir dos anos 70, quando a carne de frango produzida passou a ser exportada. A modernização da agricultura, impulsionada pela instalação de grandes indústrias avícolas na região sul do país, possibilitou a importação de insumos e aporte tecnológico à produção de frango de corte.

Em concordância com a forte demanda, países como a China, Índia e Brasil impulsionaram a sua produção avícola, adequando-se às exigências do mercado consumidor na produção de proteína animal para rápida disponibilização. Segundo dados do USDA – FSA (*United States Department of Agriculture - Foreign Agricultural Service*), esses países são responsáveis por mais de três quartos do crescimento global na produção de aves, estimado em 1% para o ano de 2013.

No que tange a produção avícola brasileira, o USDA – FDA (*Food and Drug Administration*) prevê um aumento de aproximadamente 2,7% na produção de carne de frango até o final do ano de 2014; indicando limitações na produção em comparação aos índices de anos anteriores. Entre os anos de 2010 e 2011, o Brasil apresentou um crescimento de 6,8%, chegando a uma produção estimada de 13,1 milhões de toneladas em 2011 (UBABEF, 2012). Estes índices conferem ao país a terceira posição no *ranking* dos países produtores de carne de frango, depois dos Estados Unidos e China, de acordo com pesquisas do USDA (2013).

A carne do frango é o principal produto das exportações avícolas brasileiras e apresentou no ano de 2011, uma receita de aproximadamente 8,2 bilhões de dólares, com um preço médio de venda em torno de 2.093 dólares por tonelada (UBABEF, 2012). Para o ano de 2014, estudos realizados pela USDA (2013) revelam um aumento de 3,46% na taxa de exportação e consequentemente no valor da receita, estimada em 22,9%. O valor da receita brasileira ainda é influenciado pelo alto custo de insumos e alimentos para formulação de ração.

Segundo MORAES e CAPANEMA (2012), o consumo de carne de frango no Brasil é superior ao das carnes bovina e suína. Isto é devido em parte ao valor final do produto, sendo a carne de frango mais barata, proporcionando consequentemente um maior poder de compra. No Brasil, o consumo *per capita* de carne de frango apresentou um crescimento acentuado de cerca de 60% entre os anos de 2000 e 2011,

passando de 29,91 kg/hab para 47,38 kg/hab (UBABEF, 2012). Segundo previsões da USDA (2013), o consumo de carne de frango apresentou um incremento de 1,58% no ano de 2013 e 2,5% para o ano de 2014.

Para atender às exigências do mercado consumidor de carne de frango, os animais abatidos devem apresentar maior deposição de músculo e menor deposição de gordura na carcaça além de cortes com boa aparência (sem avarias) (FURTADO et al., 2011). Os animais para criação devem ser selecionados de acordo com sua performance de crescimento e composição de carcaça. O conhecimento dos mecanismos que regulam a deposição de gordura na carcaça é fundamental aos programas de melhoramento, permitindo a seleção de animais de acordo com a sua aptidão.

2.2 Deposição de gordura na carcaça: uma abordagem genética

A deposição de gordura tem sido alvo de pesquisas nas ciências zootécnicas, devido ao seu impacto negativo sobre a eficiência alimentar e rendimento de carne magra (IKEOBI et al., 2002). Embora a gordura esteja associada com a melhoria do sabor da carne, em excesso é um atributo indesejável pelos consumidores (LAGARRIGUE et al., 2006).

A maior parte da gordura na carcaça em galinhas está depositada sob pele, nos órgãos e na porção abdominal (NASSAR et al., 2012). Aves com deposição excessiva de gordura podem causar problemas no processamento da carne e também uma rejeição por parte dos consumidores, preocupados com a qualidade nutricional dos cortes (ABASHT et al. 2006; ZHOU et al. 2007). Sendo assim, gordura abdominal é um problema na produção de frangos de corte, afetando a rentabilidade do sistema produtivo.

Choct e colaboradores (2000) estudaram linhagens experimentais de frangos de corte e relataram que estas linhagens tinham cerca de 150-200 g de gordura por kg de peso corporal e 85% deste conteúdo de gordura não possuía importância fisiológica conhecida. A seleção de animais com maior aptidão para deposição de músculo (frangos de corte) levou ao aumento do peso, da taxa de crescimento e da gordura depositada/armazenada nos cortes. O perfil da gordura depositada, bem como a percentagem de gordura nos diferentes cortes, influencia a qualidade nutricional da carne de frango (MURAKAMI et al., 2010).

Segundo Michelin Filho (1986) a gordura abdominal é o principal depósito de gordura nas aves e está diretamente relacionada ao teor de gordura na carcaça (KESSLER et al., 2000). A seleção de animais com menor deposição de gordura abdominal favorece então a redução no teor total de gordura na carcaça. A seleção contra a gordura abdominal é de extrema importância para a eficiência do sistema produtivo de aves de corte.

Oguz et al. (1996) relataram um incremento na gordura abdominal de galinhas selecionadas para aumento do peso corporal quando comparada ao teor de gordura abdominal em animais não selecionados. Madeira e colaboradores (2010), discutiram ainda a resposta diferenciada nas características de carcaça de acordo com a alimentação fornecida e com o manejo empregado, entretanto, o resultado final observado na deposição de gordura depende da seleção genética empregada na linhagem, levando-se em consideração o objetivo final pretendido.

Segundo Lagarrigue (2006) a seleção comercial contra a deposição de gordura é de difícil estabelecimento, levando-se em consideração a dificuldade e custo de abate e dissecação. As aves geneticamente melhoradas apresentam incremento no peso corporal (peso da carcaça), mobilizando aporte energético para o desenvolvimento muscular. Estas aves têm o apetite aumentado, em resposta à demanda fisiológica. O excesso de energia consumida é mobilizado então para o depósito de gordura na carcaça. Assim, o progresso no desempenho favorável da galinha está associado ao aumento na deposição de gordura (ZEREHDARAN et al., 2004) possivelmente devido a correlações genéticas positivas entre o peso corporal e esta característica.

Foram obtidas estimativas de correlações genéticas na galinha de 0,35 (GAYA, 2006) e $0,58 \pm 0,46$ (LEENSTRA E PIT, 1988) entre a conversão alimentar e o peso da gordura abdominal, mensurados em linhagens de corte com a mesma idade de abate. Na linhagem experimental de corte (TT), da EMBRAPA Suínos e Aves, a correlação genética estimada foi de $0,36 \pm 0,13$ entre o peso aos 42 dias e o peso da gordura abdominal (CRUZ et al., 2011). Assim, aves selecionadas com maior peso vivo e consequentemente maior ingestão de alimentos, tendem à maior deposição de gordura.

Segundo Abasht e colaboradores (2006) a deposição de gordura é uma característica altamente hereditária, com herdabilidade variando entre 0,5 e 0,8 (LECLERCQ et al., 1988; CHAMBERS, 1990; LE BIHAN-DUVAL et al., 1999). Em estudos desenvolvidos com animais comerciais de corte, Cahaner e Nitsan (1985)

relataram uma herdabilidade de $0,82 \pm 0,28$ para peso da gordura abdominal. Os valores de herdabilidade encontrados indicam, segundo Gaya et al. (2003), uma boa resposta aos programas de seleção contra a deposição de gordura abdominal.

2.3 Importância e caracterização do genoma da galinha

Quando comparada aos bovinos e aos suínos, a galinha é considerada um importante modelo para estudos na área da genômica animal por apresentar estrutura social que permite grandes grupos sociais de machos com fêmeas, facilidade de estabelecer cruzamentos entre as gerações, precocidade sexual, agilidade limitada e adaptação a uma variedade de ambientes (CROOIJMANS et al., 1996; SIEGEL et al., 2006), além da facilidade de obtenção do DNA das células sanguíneas vermelhas nucleadas.

A galinha foi o primeiro animal doméstico a ter o genoma sequenciado, publicado por Hiller e colaboradores (2004). O genoma sequenciado foi de uma única fêmea do ancestral selvagem da galinha doméstica, a *Red Jungle Fowl* (ABPLANALP, 1992; DODGSON, 2011), que apresentou um tamanho de 1,050 Mb e foi realizado com uma cobertura de sequenciamento de 6,6X. O genoma apresentou regiões com baixa qualidade, devido às limitações tecnológicas no sequenciamento e por isso, essas regiões não foram adicionadas ao tamanho final do genoma sequenciado.

Posteriormente, no ano de 2006, uma nova sequência foi disponibilizada e foram adicionados 198 Kbp em regiões de baixa qualidade (DODGSON, 2011). De acordo com a última versão do genoma (*Gallus_gallus* 4.0), publicada no banco de dados ENSEMBL em 2012, o genoma da galinha possui 15.508 genes codificantes, 1.558 não codificantes, 42 pseudogenes e 17.954 genes transcritos. Além disso, foram descritos cerca de 4 milhões de pequenas variantes SNPs, INDELs e mutações somáticas. A última versão do genoma referência da galinha apresenta 1,072 Mb.

A galinha doméstica é uma exceção ao modelo biológico de cariótipo dos machos e fêmeas, pois a fêmea é o sexo heterogamético, e o macho, o sexo homogamético. Desta forma, ela se enquadra no sistema ZW onde o macho é ZZ e a fêmea ZW. Na montagem do genoma por Hiller e colaboradores (2004), estes cromossomos foram mal representados (BURT, 2005), devido ao excesso de repetições e/ou duplicações do cromossomo W, apresentado pela fêmea. Na montagem do genoma publicada em

2012, disponível no ENSEMBL, estes cromossomos ainda apresentaram regiões com baixa qualidade de sequenciamento.

Além disso, o grande tamanho amostral da população (inúmeras linhagens de galinhas domésticas no mundo) e a rica diversidade genética influenciada pela presença de mutações, estimada em uma variante única a cada 374 pb ao longo do genoma (WONG et al., 2004), são atributos facilitadores para estudos genômicos em galinhas (CARLBORG et al., 2006). Brandstrom e Ellegren (2007) em um estudo de inserções e deleções (INDELs), relataram uma densidade média de cerca de um polimorfismo (SNP) a cada 350 pb ao longo do genoma da galinha. A densidade de INDELs foi altamente correlacionada com a densidade de SNPs, apresentando $1,9 \times 10^{-4}$ eventos de INDEL/pb.

Segundo Siegel e colaboradores (2006) a alta taxa de recombinação encontrada na galinha doméstica é uma característica que possibilita o desenvolvimento de linhagens experimentais para estudos de mapeamento. Estudos realizados com linhagens experimentais permitem a identificação dos aspectos genéticos envolvidos na determinação do fenótipo, com foco na presença ou identificação de mutações.

A identificação de mutações em regiões anteriormente descritas e relacionadas à regulação das características de produção (fenótipo) aumenta a probabilidade de identificação de mutações causais, explicando parte da variação fenotípica encontrada. As regiões podem ser identificadas ao longo do genoma pelo mapeamento de QTL.

2.4 QTLs mapeados para deposição de gordura em galinhas

O mapeamento de QTLs, a identificação de genes candidatos e a descoberta de novos genes são importantes para o conhecimento do genoma do animal (BURT, 2002), bem como a identificação das regiões e componentes responsáveis pelas expressões fenotípicas. O loco determinado pelo QTL representa parte da variância fenotípica (WONG et al., 2004), e neste sentido, a identificação dos sítios polimórficos é fundamental para a interpretação das informações contidas no código genético, expressas no fenótipo do animal.

Por isso, a determinação de regiões de QTLs auxilia a localização de regiões no genoma que estão associadas com determinadas características e podem conter polimorfismos que merecem ser melhor investigados, como os polimorfismos SNPs ou as INDELs.

Diversos estudos foram desenvolvidos a fim de mapear regiões do genoma associadas com o depósito de gordura em galinhas. Modelos experimentais foram criados e foram mapeados diversos QTLs associados com peso e percentagem da gordura abdominal (TATSUDA e FUJINAKA et al., 200; IKEOBI et al., 2002; JENNEN et al., 2004; LAGARRIGUE et al., 2006; CAMPOS et al., 2009; PINTO et al., 2010; NONES et al., 2012), porém, abrangendo regiões compreendidas entre dezenas de centimorgans (cM) (ABASHT et al., 2006). Segundo dados do *Chicken QTLdatabase* (CHICKEN QTLDB, 2013), já foram mapeados 174 QTLs para peso da gordura abdominal e 118 para percentual de gordura abdominal. Deste total, cerca de nove QTLs para percentagem de gordura abdominal e 11 QTLs para peso da gordura abdominal foram mapeados no GGA3.

Existem diversas linhagens estabelecidas para fins experimentais ou para estudos genômicos, que podem ser usadas para estudos da relação entre genótipo e fenótipo (SIEGEL et al., 2006). O cruzamento entre linhagens divergentes geneticamente (corte e postura, por exemplo), permite ainda o estudo dos diferentes fenótipos transmitidos às próximas gerações. Estas linhagens foram também desenvolvidas para a criação de mapas de ligação e mapeamento de QTLs. Estas populações são geralmente menores em número de animais, podendo ser utilizadas também na identificação de genes candidatos e mutações.

Foram mapeados QTLs significativos associados com peso e percentagem da gordura abdominal em linhagens divergentes de corte e postura no cromossomo 1 da galinha (IKEOBI et al., 2002; JENNEN et al., 2004; LAGARRIGUE et al., 2006; XU et al., 2013); GGA3 (IKEOBI et al., 2002; CAMPOS et al., 2009); GGA5 (IKEOBI et al., 2002; LAGARRIGUE et al., 2006; MCELROY et al., 2006; SUN et al., 2013); GGA7 (IKEOBI et al., 2002; SCHUTZ et al., 2002; DEMEURE et al., 2013); GGA13 (IKEOBI et al., 2002; JENNEN et al., 2004; PINTO et al., 2010), GGA27 (SCHUTZ et al., 2002; DEMEURE et al., 2013) e GGA28 (IKEOBI et al., 2002).

No Brasil, foi estabelecida em 1999 uma colaboração entre a EMBRAPA Suínos e Aves e a ESALQ/USP para a condução de pesquisas em genômica de aves, que possibilitou o mapeamento de diversos QTLs e a identificação de genes associados com características de interesse econômico, principalmente para desempenho e carcaça. Os estudos foram conduzidos em uma população obtida de cruzamentos recíprocos entre linhagens não endogâmicas de corte (TT) e uma linhagem de postura

(CC). Essas populações, em esquema F_2 , foram denominadas CTCT e TCTC (detalhes em ROSÁRIO et al., 2009).

Em TCTC, após a conclusão de uma varredura do genoma, foram mapeados diversos QTLs associados com características relacionadas com a deposição da gordura: peso de gordura abdominal, percentagem de gordura abdominal, triglicerídeos e colesterol. Campos et al. (2009) mapearam QTLs sugestivos e significativos associados com níveis séricos de triglicerídeos no GGA5, 23 e 27; com percentagem de gordura abdominal no GGA3, 10, 12 e 27 e associados com peso da gordura abdominal no GGA2, 3 e 27. Ruy (2004) mapeou os mesmos QTLs significativos para as características de peso da gordura abdominal e porcentagem de gordura no GGA3 entre os marcadores *LEI0029* e *ADL0371* (Tabela 2). Na população CTCT foram identificados QTLs associados com peso de gordura abdominal e colesterol no GGA5 (SILVA et al., 2011).

Em resumo, diversos QTLs associados com deposição de gordura foram mapeados nos GGA2, 3, 5, 10, 12, 23 e 27 na população da EMBRAPA. Os QTLs significativos que foram mapeados entre os marcadores *LEI0161-LEI0029* e *LEI0029-ADL0371* no GGA3 foram associados com o peso e a percentagem de gordura abdominal. O QTL mapeado entre os marcadores *LEI0029 – ADL0371* foi descrito em dois estudos em galinhas (Tabela 2), o que pode indicar uma região de interesse para a investigação de polimorfismos que podem estar associados com deposição de gordura. Além disso, foi detectado um efeito significativo de *imprinting* para o QTL associado com porcentagem de gordura no GGA3 por Campos e colaboradores (2009). Segundo Nolan e Pettite (2001), o *imprinting* genômico ocorre quando a expressão do gene e/ou QTL é dependente dos parentais. Sendo assim, os animais parentais apresentam importância significativa na segregação do QTL apresentado.

Tabela 2 – QTLs mapeados pela análise de irmãos completos (*full-sib*) entre os marcadores *LEI0029* e *ADL0371* no cromossomo 3 da galinha para as características de deposição da gordura na população F_2 da EMBRAPA.

Característica ¹	Posição (cM)	Marcadores Flanqueadores	IC ² (cM)	Valor de F	PV ³
Campos et al., 2009					
GORD (g)	115	<i>LEI0029–ADL0371</i>	80-195	7,5*	4,08
% GORD	112	<i>LEI0029–ADL0371</i>	75-125	10,8**	6,05
Ruy, 2004					
GORD (g)	121	<i>LEI029–ADL0371</i>	68-211	7,9**	4,15
% GORD	121	<i>LEI029–ADL0371</i>	102-130	13,6**	6,37

¹GORD = peso da gordura abdominal; % GORD = porcentagem de gordura abdominal.

²Intervalo de confiança.

³Porcentagem de variação explicada pelo QTL.

**Significativo a 1% no cromossomo. *Significativo a 5% no cromossomo.

Os QTLs descritos por Campos e colaboradores (2009) (Tabela 2), levaram em consideração a análise global do efeito de todos os 544 indivíduos analisados no percentual de variação fenotípica, de acordo com metodologia da análise de irmãos completos (*full-sib*). No mesmo estudo a análise de meio irmãos (*half-sib*) também foi realizada e os QTLs mapeados diferiram quanto à sua localização (intervalo de marcadores flanqueadores) dos mapeados com a análise de irmãos completos. Por exemplo, os QTLs mapeados para peso de gordura abdominal e percentual de gordura abdominal que foram mapeados entre *LEI0029* e *ADL0371* na análise de F_2 , foram mapeados entre os marcadores *LEI0161* e *LEI0029* pela análise de meio irmãos. Embora o intervalo de marcadores seja diferente, eles estão localizados muito próximos.

Segundo outros estudos, diversos QTLs foram mapeados para percentual de gordura abdominal e peso da gordura abdominal no GGA3 (Tabela 3), em intervalos próximos aos mapeados pelas pesquisas de Campos e colaboradores (2009) (Tabela 2).

Tabela 3 – QTLs mapeados para percentual de gordura abdominal e peso da gordura abdominal no GGA3, em regiões próximas ao intervalo de marcadores *LEI0161-ADL0371*

Característica ¹	Marcadores Flanqueadores	IC ² (cM)	Valor de F	% PV ³
Park et al., 2006				
GORD (g)	<i>MCW0222-MCW0004</i>	85-155	--	--
Campos et al., 2009				
GORD (g)	<i>LEI0161-LEI0029</i>	113-126	18,5 ^{**}	1,12
% GORD	<i>LEI0161-LEI0029</i>	113-126	13,1 ^{**}	0,05
Nadaf et al., 2009				
% GORD	--	32-168	7,9 [‡]	1,03
GORD (g)	--	32-144	7,4 [‡]	0,90
Wang et al., 2012				
GORD (g)	<i>HUJ0006-LEI0161</i>	84-113	4,7 [‡]	2,61
% GORD	<i>HUJ0006-LEI0161</i>	84-113	6,5 [*]	3,54

¹GORD = peso da gordura abdominal; % GORD = porcentagem de gordura abdominal.

²Intervalo de confiança.

³Porcentagem de variância fenotípica explicada pelo QTL.

^{**}Significativo a 1% no cromossomo. ^{*}Significativo a 5% no cromossomo. [‡]Sugestivo.

Com base nestes resultados apresentados, uma região-alvo de estudo foi definida no GGA3 entre os marcadores *LEI0161* e *ADL0371*. A escolha foi feita principalmente devido aos QTLs mapeados no estudo de Campos e colaboradores (2009), que foram associados com peso de gordura abdominal e porcentagem de gordura abdominal, alguns deles com alta proporção da variância fenotípica (6,05%) explicada pelo QTL (Tabelas 2 e 3).

2.5 Genes associados à deposição e metabolismo de gordura

A característica de deposição de gordura é poligênica/complexa, ou seja, um número desconhecido de genes pode estar envolvido no controle desta característica. Diferentes genes já foram associados à deposição de gordura em galinhas, por exemplo, *lipoprotein lipase* (*LPL*, localizado no GGAZ, COOPER et al., 1992); *patatin-like phospholipase domain containing 2* (*PNPLA2*, Localizado no GGA5, NIE et al., 2010); *ghrelin/obestatin prepropeptide* (*GHRL*, localizado no GGA12, NIE et al., 2009); *insulin-like growth factor 2 (somatomedin A)* (*IGF2*, localizado no GGA5, WANG et al.,

2005) e *fatty acid binding protein 7, brain (FABP7*, localizado no GGA3, GODBOUT, 1993).

Resnyk et al. (2013) identificaram ainda, genes candidatos expressos na gordura abdominal em um estudo de expressão diferencial de genes em linhagens de galinha selecionadas para alta e baixa deposição de gordura na carcaça. Alguns dos genes foram: *Fibrinogen alpha (FGA*, localizado no GGA4) e *Thrombospondin 2 (THBS2*, localizado no GGA3) relacionados à homeostase; *Adiponectin (ADIPOQ*, localizado no GGA9) e *Angiopoietin-like 4 (ANGPTL4*, localizado no GGA28) relacionados às síntese de adipocinas; *7-Dehydrocholesterol reductase (DHCR7*, localizado no GGA5) e *Fatty acid desaturase 2 (FADS2*, localizado no GGA5) relacionados à lipogênese e *Acetyl-CoA acetyltransferase 1 (ACAT1*, localizado no GGA1) e *Alcohol dehydrogenase 1C (class I) gamma polypeptide (ADH1C*, localizado no GGA4) relacionados à lipólise.

Wu e colaboradores (2006) identificaram em populações experimentais selecionadas para corte, postura e ainda em linhagens indígenas, polimorfismos no gene *proliferadores de peroxissoma ativado receptor- γ coactivator-11 (PPARGC1A)* (localizado no GGA4), que foi associado com o aumento da gordura abdominal nestas aves. Oguro e colaboradores (2012) identificaram um polimorfismo SNP no gene *adenosine deaminase (DSRAD)* associado aos níveis séricos de triglicerídeos e circunferência abdominal em galinhas. Em estudos na população experimental da EMBRAPA Suínos e Aves, Souza (2004) estudou mutações no gene *miogenina* (GGA26), encontrando associação com a gordura abdominal.

Diversos estudos de identificação de mutações e associação com características fenotípicas foram desenvolvidos para diversas características de carcaça em galinhas. A identificação de mutações ainda não descritas e possivelmente associadas à deposição de gordura na carcaça é fundamental à seleção genômica contra esta característica.

2.6 Variações genéticas e expressão das características fenotípicas

A descoberta de mutações ao longo do genoma permite a identificação, a partir da sua posição, das regiões onde elas estão inseridas e o seu potencial papel na regulação da expressão gênica ou ainda, sua influência na determinação do fenótipo. Os polimorfismos são encontrados tanto nas regiões codificadoras quanto nas regiões não codificantes, entretanto, as mutações que ocorrem nas regiões de éxon ou ainda nas junções dos éxons, podem modificar o RNA mensageiro (mRNA) processado,

produzindo um mRNA maduro que codifica para uma proteína diferente da original (PATTHY, 1994).

Polimorfismos identificados em regiões não codificantes podem estar relacionados à regulação da expressão gênica. A Figura 1 apresenta as regiões ao longo do genoma e do RNA mensageiro onde as mutações podem ser identificadas.

Figura 1 – Esquema representativo das regiões distribuídas ao longo do genoma e do RNA mensageiro

Mutações identificadas em regiões que antecedem as regiões codificadoras dos genes são classificadas como *upstream*; as mutações à jusante das regiões codificadoras dos genes são classificadas como *downstream* (Figura 1). As regiões não traduzidas do RNA mensageiro (UTR 5' e UTR 3') podem relacionar-se à regulação da expressão gênica após transcrição (DENG et al., 2013).

As regiões de UTR 3' estão implicadas na formação da cauda Poli-A e influenciam a expressão dos genes na etapa pós transcricional (KUERSTEN e GOODWIN et al., 2003). Deng et al. (2013) descreveram a associação da região UTR 3' com a regulação da expressão do gene *LIMK1*, associado ao desenvolvimento neural em estudo com seres humanos, sugerindo esta região como alvo de pesquisas de identificação de variações genéticas. Ainda em estudos com humanos, mutações na região não

transcrita (UTR 5') foram associadas à manifestação de distúrbios cardíacos (*Cardiomiopatia arritmogênica do ventrículo direito*) sugerindo uma função reguladora desta região (BEFFAGNA et al., 2005).

Waters et al. (2010) identificaram SNPs na região UTR 5' do gene responsável pela síntese do hormônio de crescimento em bovinos. Os animais nos quais essas mutações foram identificadas apresentaram variações nas características de desempenho.

Em galinhas, Rao et al. (2013) descreveram variações na expressão gênica em animais que apresentaram alterações da sequência na região de UTR 5'. Ainda neste estudo, foram identificadas variações no tamanho dessas regiões e suas respectivas implicações na regulação da expressão sugerindo assim, a importância do estudo de identificação de INDELs ao longo do genoma.

Além das regiões de UTR, das anteriores à região codificadora dos genes (*upstream*), entre genes (intergênicas) e à justante da região codificadora dos genes (*downstream*), segundo Wang (2010) as mutações podem ainda ser identificadas em regiões codificantes (regiões de éxon), sítios de *splicing* (intervalo de 2 pb) e ncRNA (sobreposição de transcrito não anotado; sem definição de gene).

As mutações identificadas em regiões codificantes (éxons), são classificadas ainda em sinônimas, não sinônimas, *stopgain* e *stoploss*. As mutações sinônimas provocam alterações nos códons, mas não nos aminoácidos, enquanto que, as mutações não sinônimas provocam alteração nos aminoácidos que compõem a proteína (SAUNA e KIMCHI-SARFATY, 2011). As mutações *stopgain* provocam a inserção de um *stop códon* enquanto que as mutações *stoploss* provocam a deleção ou perda de um *stop códon* (WANG, 2010).

A alteração na composição dos códons e consequentemente na sequência de aminoácidos pode ter efeitos significativos na variação da expressão gênica (PLOTKIN e KUDLA, 2010). Assim, o aprimoramento das técnicas para identificação de variações genéticas é fundamental ao estudo dos mecanismos envolvidos na regulação das características fenotípicas.

2.7 Mutações causais

Como discutido anteriormente, diversos QTLs já foram mapeados para as características fenotípicas de interesse econômico em aves, entretanto, existe uma disparidade entre o número de QTLs mapeados e o número de mutações causais

identificadas (BRAUNSCHWEIG, 2010). Segundo o mesmo autor, a identificação de mutações causais na região de QTL associadas à característica fenotípica estudada, permite um maior entendimento da forma com que a expressão do fenótipo é regulada.

Assim, o detalhamento da região de QTL pode fornecer informações de um grande número de pares de base e prováveis mutações causais (AHSAN et al., 2013). Características poligênicas, a exemplo da deposição de gordura na carcaça, necessitam de mapeamento amplo do genoma para determinação das variantes possivelmente relacionadas à regulação da característica, levando em consideração a interação entre os elementos do genótipo e a influência deles no fenótipo (BRAUNSCHWEIG, 2010).

Mutações causais podem ser identificadas em todo o genoma seja em regiões codificadoras ou ainda em regiões não codificadoras, como descrito no tópico anterior. Os SNPs e/ou INDELs identificados em regiões não codificadoras podem alterar a afinidade da enzima RNA-polimerase no início da transcrição ou ainda, estar em desequilíbrio de ligação com mutações causais (ANDERSSON, 2013). Sendo assim, mutações em regiões não codificadoras podem responder por parte da variação fenotípica observada.

Clop et al. (2005), discutiram a interferência de SNP na expressão do gene da *miostatina*, em ovinos. O SNP foi descrito em uma região não traduzida (UTR 3'), e a presença deste polimorfismo resulta na síntese de um microRNA responsável pela regulação negativa na expressão deste gene. Em um estudo com suínos, Markljung et al. (2009) descreveram uma mutação em íntron, relacionada à regulação da expressão do gene *IGF2*. Estes estudos evidenciam a importância de buscar mutações causais ao longo de todo o genoma.

Em galinhas, mutações causais foram descritas em diversos estudos relacionados ao formato e coloração da crista. Wright et al. (2009), demonstraram que uma CNV (*copy number variation*) em uma região de íntron no gene *SOX5* é responsável pelo fenótipo de crista de tamanho médio (*Pea-comb*) em galinhas. Gunnarsson et al. (2011) identificaram uma deleção no gene *SOX10*, responsável pelo fenótipo de cor marrom escura em galinhas (*Dark brown colour*), e Imsland et al. (2012) identificaram uma inversão no gene *MNR2*, responsável pelo fenótipo de crista em formato cilíndrico (*Rose-comb*). Assim, a identificação de mutações em genes e o conhecimento acerca do efeito dos genes ou suas respectivas vias metabólicas é fundamental à identificação de mutações causais (BRAUNSCHWEIG, 2010).

Em trabalhos desenvolvidos na linhagem experimental da EMBRAPA Suínos e Aves, Ledur et al. (2012) investigaram genes associados ao metabolismo de lipídios a exemplo dos *ADIPOR2*, *LEPR*, *HNF4 α* dentre outros. No gene *LEPR* foram identificados SNPs segregando na linhagem de corte (TT) (NINOV et al., 2006), associados às características de desempenho. No gene *HNF4 α* , Silva et al. (2012) identificaram um SNP associado à redução da gordura abdominal em galinhas.

Em estudo de identificação de polimorfismo no gene *OBR*, realizado em diferentes gerações de uma população de frangos de corte, Wang et al. (2004) identificaram um polimorfismo associado à deposição de gordura abdominal. Desta forma, o mapeamento fino de regiões (QTL) contendo genes associados às características de produção, favorece o conhecimento de um número grande de mutações, para estudo de associação (AHSAN et al., 2013). É necessário então, o emprego de tecnologias capazes de analisar um grande volume de dados.

2.8 Identificação de variações genéticas

O conhecimento do genoma é fundamental para o uso das tecnologias baseadas em DNA na promoção do melhoramento genético. Embora o mapeamento do genoma tenha evoluído ao longo dos anos, dispondo de tecnologias de alto rendimento e acurácia, o entendimento da forma com que as informações contidas no genótipo são expressas ainda é algo complexo (SIEGEL et al., 2006). Diversas pesquisas são desenvolvidas com o objetivo de associar SNPs com características fenotípicas de interesse, tendo como foco então, as variantes de menor tamanho (polimorfismos de base única - SNPs). Os SNPs têm permitido o desenvolvimento de novas metodologias e estratégias para o mapeamento de genes de interesse, pois são marcadores extremamente úteis para promover o mapeamento fino, cujo objetivo é delimitar a menor região genômica em que um QTL foi mapeado (CARLSON et al., 2004).

Estudos de associação de SNPs com características de produção foram largamente desenvolvidos, possibilitados por diversas tecnologias aplicadas à identificação de variações genéticas, em detrimento das inserções e deleções (INDELs). As variações estruturais, especificamente as INDELs, apresentam limitações das técnicas para identificação acurada e posterior associação, entretanto, podem responder por parte da variação fenotípica (MILLS et al., 2011). A taxa de INDELs, o padrão de distribuição e as implicações evolutivas das INDELs são pouco caracterizados se comparadas às dos SNPs, em estudos desenvolvidos em galinhas.

INDELs são alterações genéticas de inserção e/ou deleção de um ou mais nucleotídeos na sequência do genoma (MATTE, 2011). As inserções e deleções podem afetar a estrutura do gene bem como a regulação das vias metabólicas a ele associadas. Assim, as INDELs contribuem para a diversidade genética e consequente variação da expressão gênica (BRITTEN, 2002; TIAN et al., 2008; RAO et al., 2010).

As INDELs podem ser multialélicas ou bialélicas (WEBER et al., 2002). As INDELs multialélicas, em sua maioria, ocorrem em pequenas sequências repetidas no DNA, também conhecidas com microssatélites. Estes marcadores são largamente utilizados na pesquisa genética, em estudos de diversidade e variabilidade (PENA, 2005).

Segundo Hill (2005), mutações em genes podem aumentar a variabilidade nas populações, bem como, na variação genética de seleção induzida (EITAN e SOLLER, 2004; CARLBORG et al., 2006). Na galinha, mais de 2,8 milhões de SNPs já foram identificados a partir da comparação da sequência do genoma do ancestral (*Red Jungle Fowl*) com sequências obtidas em três linhagens domesticadas: um macho de corte (*White Cornish*), uma fêmea de postura (*White Leghorn*) e outra fêmea de uma espécie ornamental (*Silkie chinesa*) (WONG et al., 2004). Groenen e colaboradores (2009) publicaram um mapa de alta densidade de SNPs que possibilitou a inclusão de 8.559 marcadores no mapa consenso de ligação a partir da genotipagem de três populações desenhadas para o mapeamento: a *Wageningen*, *East Lansing* e *Uppsala*.

Novos SNPs identificados, em diferentes linhagens estão disponíveis no banco de dados dbSNP NCBI. Ainda segundo dados do dbSNP (NCBI, 2013), existem 8,97 milhões SNPs depositados referentes ao genoma da galinha. Destes, aproximadamente 1 milhão estão localizados no cromossomo 3. Para a região-alvo, entre os marcadores *LEI0161–ADL0371*, aproximadamente 84 mil SNPs já estão depositados. Ainda segundo dados disponíveis no dbSNP (NCBI, 2013), cerca de 450 mil pequenas INDELs já estão depositadas para o genoma da galinha, 53 mil no GGA3 e 4,1 mil presentes na região-alvo definida neste estudo.

O estudo de identificação de marcadores (SNPs) e mapas de haplótipos contendo QTLs mapeados em populações experimentais é válido se considerada a correspondência com populações comerciais, a partir do catálogo de mutações descritas (COUTINHO et al., 2010). Ainda segundo estes autores, o ressequenciamento de regiões do genoma associadas às características quantitativas (QTLs) permitirá a identificação de mutações causais descritas e não descritas.

Além da detecção de milhares de SNPs com uma alta acurácia, taxas de variação de nucleotídeos e aminoácidos, e as pequenas inserções/deleções podem ser investigadas de forma mais aprofundada a partir dos dados de sequenciamento de nova geração (LERNER e FLEISCHER, 2010). Marklund e Carlborg (2010) investigaram a sensibilidade de detecção de SNPs, a partir da utilização de marcadores flanqueadores como uma nova medida para prever a variabilidade, utilizando a tecnologia de sequenciamento de nova geração. Foram avaliados *pools* de DNA de duas linhagens divergentes de galinhas selecionadas para peso corporal aos 56 dias de idade. Foram identificados cerca 31 mil SNPs dentro ou entre as linhagens e uma distribuição de cerca de 3,7 SNPs/kb.

Rubin e colaboradores (2010) usaram a tecnologia de nova geração de sequenciamento para a identificação de deleções e SNPs em genes candidatos e assinaturas de seleção, em oito diferentes populações de galinhas domésticas e na *Red Jungle Fowl*. Neste estudo foram identificados cerca de 7 milhões de SNPs e aproximadamente 1.300 deleções, com uma cobertura de sequenciamento de 44,5X. Também por meio da mesma tecnologia de sequenciamento, Kranis et al. (2013) utilizaram 243 aves de 24 linhagens divergentes em aptidão (corte, postura, comerciais e experimentais) para a identificação de mutações e para o desenvolvimento de um painel de alta densidade de SNPs. Foram identificados cerca de 78 milhões de SNPs segregando em uma ou mais linhagens. Após filtragens, foram selecionados cerca de 1,8 milhões de SNPs para a montagem de um *chip* denso de SNPs (600K).

Brandström e Ellegren (2007) identificaram 140.484 INDELs no genoma da galinha, analisando polimorfismos identificados por meio dos métodos convencionais de sequenciamento (SANGER) de três linhagens de galinhas domésticas. Na análise foram removidas as regiões de microssatélites. A densidade média foi de $1,9 \times 10^{-4}$ INDELs/pb, e uma menor densidade nos microcromossomos do que nos macrocromossomos foi constatada.

Portanto, as novas plataformas de sequenciamento apresentam à genômica perspectivas futuras positivas no conhecimento sobre o genoma das aves, principalmente no estudo de variantes alélicas, SNPs, INDELs, e ao desenvolvimento de marcadores para seleção assistida, representando uma importante ferramenta no melhoramento animal (KATO, 2009). No que tange à genômica avícola, os trabalhos com as novas plataformas são promissores no sentido de completar a sequência do genoma, cobrindo regiões com baixa leitura ou baixa confiabilidade, permitindo assim o

estudo comparativo aplicado à expressão das características de interesse zootécnico (BURT, 2005).

2.9 Ferramentas para identificação das variações genéticas

O avanço na tecnologia de sequenciamento de genomas criou uma grande demanda por análises de bioinformática. A bioinformática é definida pela integração de ferramentas da informática para diagnóstico e interpretação de dados possibilitando o uso de modelos estatísticos para a análise de grandes quantidades de dados biológicos (ORDINE et al., 2011). A análise de um grande conjunto de dados permite a associação do genótipo de um maior número de animais com a característica estudada.

As predições da sequência de modo acurado bem como a análise das informações nela contidas tornaram-se fundamentais às pesquisas genéticas. A tecnologia de sequenciamento de nova geração gera milhões de *reads* (pequenas sequências) de DNA, em uma única corrida e o mapeamento deste grande volume de sequências, comparando-o com um genoma referência, é um desafio aos programas de alinhamento (MU et al., 2012; PABINGER et al., 2013). A qualidade do alinhamento interfere diretamente na montagem do genoma ou ainda, na identificação de mutações. Quanto maior o número de *reads* alinhadas na mesma posição, maior será a qualidade da base sequenciada.

Neste sentido, diversos programas de alinhamento foram desenvolvidos a exemplo do Bowtie2 (LANGMEAD e SALZBERG, 2012), BWA (LI e DURBIN, 2009), SeqAlto (MU et al., 2012), RMAP (SMITH et al., 2008), MAQ (LI et al., 2008), ZOOM (LIN et al., 2008), SeqMap (JIANG e WONG, 2008), CloudBurst (SCHATZ, 2009), Novoalign (<http://novocraft.com/>) e SHRIMP (<http://compbio.cs.toronto.edu/shrimp>). Algoritmos foram desenvolvidos para o alinhamento de sequências aos pares ou até mesmo para alinhamento de sequências múltiplas (ORDINE et al., 2013).

O tipo de alinhamento está relacionado ao perfil dos dados gerados pela plataforma de sequenciamento. O alinhamento das *reads* pelo Bowtie2 demanda menor tempo em comparação aos outros programas anteriormente citados. O Bowtie2 realiza o alinhamento combinando a sequência indexada do genoma e ainda, algoritmos de programação dinâmica configurados para otimização do uso do *hardware* possibilitando uma maior velocidade, sensibilidade e precisão (LANGMEAD e SALZBERG, 2012).

As variações (SNPs e INDELS) podem ser identificadas por diversos programas, a exemplo do SAMtools (LI et al., 2009), Dindel (ALBERS et al., 2011), GATK (MCKENNA et al., 2010), ou Corona Lite pipeline (Life Technologies). O SAMtools permite, além da identificação de mutações, outras operações, como a indexação da sequência, a visualização das *reads* alinhadas em relação a sequência do genoma referência, bem como a obtenção da qualidade do alinhamento e porcentagem de *reads* alinhadas/mapeadas.

A anotação funcional das mutações identificadas de modo acurado pode ser feita por diversos programas tais como ANNOVAR (WANG et al., 2010), SNPdat (DORAN e CREEVEY, 2013), VEP (*Variant Effect Predictor*, MCLAREN et al., 2010) dentre outros. O fluxo de trabalho para identificação de variações genéticas incluindo os programas empregados em cada etapa está descrito na Figura 2.

Figura 2 – Fluxo de trabalho para identificação de variações genéticas pela tecnologia de sequenciamento de nova geração

Sendo assim, a integração das ferramentas de bioinformática permitem a identificação de variações genéticas (SNPs/INDELS) a partir dos dados de sequenciamento de nova geração.

3. OBJETIVOS

3.1 Objetivo geral

Identificar polimorfismos (SNPs) e inserções/deleções (INDELs) por meio de sequenciamento de nova geração (Illumina) em uma região-alvo no GGA3 anteriormente associada à deposição de gordura na carcaça em uma população experimental de galinhas, visando à identificação de potenciais mutações causais relacionadas com deposição de gordura. A identificação permitirá reunir as informações para a elaboração de um catálogo completo de variações genômicas de alta qualidade/acurácia nas linhagens parentais de corte e postura que deram origem à população F₂ brasileira.

3.2 Objetivos específicos

- 1) Ressequenciar o genoma completo de 18 animais das linhagens parentais de corte (TT) e postura (CC) da EMBRAPA Suínos e Aves, utilizando métodos de sequenciamento de nova geração (Illumina).
- 2) Identificar polimorfismos (SNPs e INDELs) com o uso de ferramentas de Bioinformática em uma região-alvo entre os marcadores *LEI0029* e *ADL0371* no GGA3 que foi associada anteriormente com deposição da gordura.
- 3) Realizar uma filtragem dos polimorfismos inicialmente identificados por meio de índices de qualidade.
- 4) Realizar a anotação funcional dos polimorfismos selecionados.
- 5) Reunir as informações dos polimorfismos selecionados para a construção de um catálogo completo de variações genômicas da região-alvo que possam estar causando a diferença em deposição de gordura em linhagens brasileiras.

4 MATERIAL E MÉTODOS

4.1 População experimental

Dezoito galinhas de duas linhagens parentais experimentais (uma linhagem de corte chamada TT e uma linhagem de postura chamada CC), desenvolvidas pela EMBRAPA Suínos e Aves foram selecionados para este estudo (nove galinhas de cada linhagem). Nesta população, foram desenvolvidos estudos de mapeamento de QTLs na geração F_2 , que é um resultado de cruzamentos recíprocos entre a linhagem de postura CC e a linhagem de corte TT (detalhes em ROSÁRIO et al., 2009).

A linhagem de frangos de corte foi selecionada para peso corporal, conversão alimentar, rendimento de carcaça e peito, viabilidade, fertilidade, eclodibilidade, redução da gordura abdominal e síndromes metabólicas, durante seis gerações (ROSÁRIO et al., 2009). Ainda de acordo com o mesmo autor, a linhagem de postura foi selecionada para a produção e peso de ovos, conversão alimentar, viabilidade, maturidade sexual, fertilidade, eclodibilidade, qualidade do ovo e redução do peso corporal, durante oito gerações.

As 18 aves estudadas foram: cinco machos de corte (TT5661, TT5921, TT6037, TT6232 e TT6270) e sete fêmeas de postura (CC5, CC37, CC88, CC241, CC332, CC570 e CC886) usadas em cruzamento para obtenção da geração F_1 (TC); dois machos de postura (CC1 e CC372) e quatro fêmeas de corte (TT5461, TT5586, TT5642 e TT5649) usados no cruzamento para obter a geração F_1 (CT).

4.2 Preparação de bibliotecas e sequenciamento

O DNA genômico foi extraído de amostras de sangue das 18 aves pelo protocolo da proteinase K. Após a extração de DNA, as amostras foram quantificadas pelo fluorômetro Qubit® (Life Technologies), e em seguida, diluídas e normalizadas para uma concentração final de 2,5 ng/mL.

A preparação das bibliotecas para o sequenciamento, incluindo a fragmentação do DNA, além da adição de sequências do adaptador e do index, foi realizada de acordo com o protocolo *Nextera DNA Sample Preparation kit* (Illumina). Depois da preparação das bibliotecas de DNA, as bibliotecas foram quantificadas por PCR em tempo real, utilizando o kit de quantificação *KAPA Library Quantification Kit* (KAPA Biosystems). Para uma subsequente diluição com concentração de 20 nM. As amostras diluídas foram preparadas para clusterização pelo *TruSeq kit PE Kit Cluster v3* (Illumina - San

Diego, EUA), e em seguida, clusterizadas (agrupadas) pelo equipamento cBOT (Illumina - San Diego, EUA).

O sequenciamento foi realizado pelas plataformas HiSeq1000 e HiscanSQ (Illumina - San Diego, EUA) e, a cobertura inicial estimada foi de 18X por amostra. A cobertura foi determinada de acordo com a capacidade de geração de dados de cada equipamento de sequenciamento utilizado.

4.3 Região-alvo estudada

Uma região-alvo de 9.036.945 pb foi definida no GGA3 entre os marcadores microssatélites *LEI0161* e *ADL0371* (33,595,706-42,632,651 pb) de acordo com a última sequência de referência de galinha disponível no banco de dados do NCBI (*Gallus_gallus*-4.0). Campos et al. (2009) mapearam quatro QTLs nesta região com base em duas análises diferentes (*line-cross* e *half-sib*) em 544 indivíduos F_2 , sendo três deles com 1% de significância do genoma e um com 5% de significância do genoma.

Os resultados das análises de *line-cross* indicaram que os QTLs para peso e percentagem da gordura abdominal foram mapeados entre os marcadores microssatélites *LEI0029-ADL0371*: um QTL para peso de gordura abdominal (F -ratio = 7,5 e 4,08% de variância explicada pelo QTL), e um para percentual de gordura abdominal (F -ratio = 10,8 e 6,05% da variação fenotípica explicada pelo QTL).

Os resultados da análise de *half-sib* indicaram que ambos os QTLs foram mapeados entre os marcadores *LEI0161-LEI0029* e segregam nos descendentes de três famílias (7822, 7769 e 7716). Essas três famílias tiveram os seguintes parentais: 7822 (macho TT6232 × fêmea CC332); 7769 (macho TT5661 × fêmea CC88) e 7716 (macho TT5649 × fêmea CC886). O QTL para peso de gordura abdominal (F -ratio = 18,5, intervalo de confiança = 80-120 cM) teve maior efeito do QTL na família 7769 e o QTL para percentagem de gordura abdominal (F -ratio = 13,1, intervalo de confiança = 76-120 cM) apresentou maior efeito do QTL também na família 7769. Em conclusão, os dois parentais TT5661 e CC88 foram considerados muito importantes para este estudo considerando as análises do QTL nesta região, indicando a necessidade de uma investigação mais aprofundada das possíveis variantes funcionais que podem estar relacionados com a deposição de gordura.

Estas regiões de QTL (*LEI0161-LEI0029* e *LEI0029-ADL0371*) têm no total 9.036.945 pb de acordo com última sequência referência da galinha (*Gallus_gallus*-

4.0), e nesta região já foram descritos 120 genes (BIOMART, 2013), sendo assim, é importante analisar detalhadamente todas as variantes genômicas apresentadas nesta região e também os seus efeitos genômicos.

4.4 Descoberta de variações genéticas e filtragem

A ferramenta CASAVA (Illumina) foi utilizada para o processamento inicial dos dados. A limpeza das *reads* (qualidade ≥ 24 e tamanho mínimo da *read* ≥ 65 pb) foi realizada para remover *reads* ou fragmentos com baixa qualidade pela ferramenta SeqClean (v.1.3.12, ZHBANNIKOV et al., 2013).

Após limpeza, as *reads* foram alinhadas contra o genoma de referência da galinha (Gallus_gallus-4.0, NCBI) pelo *software* Bowtie 2 (v.2.1.0, LANGMEAD e SALZBERG, 2012). Os SNPs e INDELs foram identificados em uma região-alvo no GGA3 (33,595,706-42,632,651 pb) pelo programa SAMtools (v.0.1.19, LI et al., 2009), com a opção *mpileup*, adotando qualidade de mapeamento e qualidade das bases ≥ 20 . O comando usado foi: `samtools mpileup -q20 -Q20 -AB -r Chr3:33,595,706-42,632,651 -ugf [file.fa] [alignfile.bam] | bcftools view -bvcg - > [output.raw.bcf]; bcftools view [output.raw.bcf] | vcfutils.pl varFilter -D99999 -S > [output.vcf]`.

Após a identificação dos SNPs e INDELs, foi realizada uma rigorosa filtragem das variantes genéticas inicialmente detectadas. A filtragem foi baseada em quatro critérios: qualidade da variante (≥ 30); cobertura mínima de SNPs/INDELs (≥ 5); seleção das variantes presentes em ambas as fitas (direta e reversa), e remoção das variantes com o máximo de cobertura maior do que a cobertura média somada ao desvio padrão, multiplicado por três (AMARAL et al., 2009; KUMAR et al., 2012). A cobertura total de sequenciamento foi calculada com base na soma dos alelos presentes na referência no sentido direto, nos presentes na referência no sentido reverso, alelos alternativos no sentido direto e alelos alternativos no sentido reverso (parâmetro do SAMtools chamado DP4).

4.5 Anotação funcional

Após a filtragem das variantes genômicas, um conjunto de SNPs e INDELs únicos (sem duplicatas) para as 18 aves foi obtido e posteriormente anotado pelo programa ANNOVAR (v. 2013aug23, WANG et al., 2010) com os parâmetros predefinidos (*default*).

Após anotação, os SNPs responsáveis por alterações nos aminoácidos foram selecionados para predição dos seus efeitos funcionais pela ferramenta VEP (*Variant Efeito Predictor*, MCLAREN et al., 2010). A ferramenta VEP prediz o SIFT (*Sorting Intolerant From Tolerant*) score para os SNPs, e classifica-os em deletérios e tolerados. A ferramenta VEP também fornece as variantes que já estão presentes em bancos de dados públicos, como dbSNP (NCBI), e fornece seus respectivos *RefSNP* (rs).

A lista de genes obtida a partir da anotação no ANNOVAR foi analisada gene a gene, a partir de consulta em bancos de dados disponíveis (OMIM-NCBI, Ensembl, Biomart etc.) e ainda foi realizada uma consulta na literatura disponível para seleção dos genes relacionados ao metabolismo de lipídios.

Variantes genéticas em regiões exônicas como os SNPs não sinônimos, *stopgain* e *stoploss* e as INDELs *frameshift*, *stopgain* e *stoploss*, foram ainda selecionadas para uma análise mais aprofundada pela ferramenta DAVID (HUANG et al, 2009.) para uma classificação funcional dos genes, com base nos parâmetros predefinidos (*default* do programa), para encontrar possíveis vias metabólicas relacionadas às variantes funcionais selecionadas, e para tentar relacioná-las com o QTL associado com o peso de gordura abdominal no GGA3.

5 RESULTADOS E DISCUSSÃO

5.1 Sequenciamento e alinhamento

No total, foram obtidas 2.785.354.494 *reads* (tamanho de 100 pb) a partir do sequenciamento de 18 galinhas e após limpeza, foram mantidas 2,132,638,003 *reads* (~77%). As *reads* mantidas foram alinhadas contra o genoma de referência da galinha (*Gallus_gallus*-4.0, NCBI) e ~99% delas foram mapeadas.

Após remoção das duplicatas de PCR pelo SAMtools, a cobertura do sequenciamento foi calculada com a opção *depth* do SAMtools, considerando a relação entre a cobertura total das bases e o tamanho do genoma em pares de bases (pb). A cobertura de sequenciamento foi ~10X, em média, para as 18 galinhas. Obtivemos uma cobertura mínima de 5,7X para o animal CC372 e, cobertura máxima de 16X para o animal TT6232.

A cobertura de sequenciamento da região-alvo entre os marcadores microsatélites *LEI0161* e *ADL0371* (33.595.706-42.632.651 pb na sequência referência *Gallus_gallus*-4.0, NCBI) foi obtida com a opção *mpileup* do SAMtools, permitindo a seleção do intervalo e cálculo da cobertura total das bases para então divisão pelo tamanho da região em pares de base (pb). Foi obtida uma cobertura média de 11X na região-alvo.

5.2 Descoberta de SNPs e INDELs e filtragem

A descoberta de SNPs e INDELs foi feita em uma região-alvo no GGA3, usando os parâmetros iniciais de mapeamento e qualidade Phred acima de 20, foram identificados 136.054 SNPs únicos e 15.496 INDELs únicas para as 18 galinhas. As mutações identificadas foram classificadas de acordo com a sua qualidade (Figura 3).

Figura 3 - Percentual de variantes classificadas de acordo com sua qualidade

A Figura 3 mostra que 5% dos SNPs e aproximadamente 19% das INDELs apresentou um índice de qualidade inferior a 30. Essas variantes foram removidas pois suas qualidades foram insatisfatórias (AMARAL et al., 2009; ECK et al., 2009). O percentual de INDELs com baixa qualidade (≤ 30) foi maior quando comparado ao dos SNPs (14% mais INDELs), mas, isso era esperado uma vez que a detecção de INDELs é mais complicada do que a de SNPs (LI et al., 2008). Segundo Neuman et al. (2012), *reads* com as inserções/deleções são mais difíceis de alinhar contra o genôma referência, e *gaps* podem ser interpretados como INDELs no mapeamento, ou ainda, erros de alinhamento podem reduzir a precisão da detecção, o que tende a aumentar a taxa de INDELs identificadas com baixa qualidade (ALBERS et al., 2011).

Em média foram identificados 55.000 SNPs e 5.000 INDELs por animal (Tabela 4). Na linhagem de corte (TT) foram identificados em média 116.616 SNPs únicos e 12.577 INDELs únicas, enquanto na linhagem de postura (CC) foram identificados 95.030 SNPs únicas e 10.452 INDELs únicas.

A filtragem de SNPs e INDELs foi feita com base em quatro critérios (como descrito anteriormente), com a remoção de ~22% dos SNPs únicos (12.646) e ~40% das INDELs únicas (2.056) para as 18 galinhas, e em média, 45.837 SNPs e 3.113 INDELs (únicos) foram mantidos na linhagem de corte, e 43.769 SNPs e 3.073 INDELs foram mantidos na linha de postura (Tabela 4). A filtragem pela qualidade da variante (≥ 30) removeu 19% dos SNPs e 5% das INDELs. A remoção de um maior número de INDELs em relação ao número de SNPs está de acordo com o maior número de

INDELs com baixa qualidade (como mostrado na Figura 3). A filtragem de variantes presentes nas duas fitas (direta e reversa) removeu 7% dos SNPs e 6% das INDELs para os 18 animais.

Tabela 4 - Número de SNPs e INDELs inicialmente identificadas na região-alvo do GGA3 e o número após a filtragem para as 18 galinhas

Animais	Nº SNPs inicialmente detectados	Nº SNPs (depois da filtragem)	Nº INDELs inicialmente detectados	Nº INDELs (depois da filtragem)
CC1	55.952	47.470	5.072	3.359
CC5	57.208	51.360	5.908	4.247
CC37	57.330	44.503	5.152	3.021
CC088	59.505	49.356	5.613	3.582
CC241	59.441	51.337	5.833	4.034
CC332	57.222	43.709	5.022	2.827
CC372	43.940	22.561	3.607	1.261
CC570	52.848	39.465	4.408	2.481
CC886	55.560	44.158	4.597	2.842
TT5461	52.302	40.223	4.539	2.581
TT5586	58.959	44.249	4.763	2.696
TT5642	59.270	39.392	4.434	2.142
TT5649	62.106	40.086	4.872	2.257
TT5661	65.303	54.766	5.928	3.765
TT5921	62.464	53.058	6.118	4.115
TT6037	49.731	29.144	4.445	1.725
TT6232	60.658	53.961	5.921	4.085
TT6270	64.285	57.656	6.450	4.651
Média	57.449	44.803	5.149	3.093

Do total das mutações identificadas após a filtragem em cada animal, alguns SNPs e INDELs foram identificados exclusivamente em uma linhagem. Na linhagem de corte 37.118 SNPs e 3.616 INDELs foram identificados exclusivamente nesta linhagem. Na linhagem de postura foram identificados 18.904 SNPs e 2.291 INDELs exclusivas nesta linhagem. O maior número de mutações presentes em uma linhagem particular, tal como a linhagem de corte, pode indicar uma maior variabilidade na região do QTL estudado, em condições equânimes de sequenciamento (número de *reads* e cobertura por animal).

O animal com maior número de SNPs exclusivos na linhagem de postura foi o CC88 (1.739), seguido pelo CC37 (1.039), e na linhagem de corte, o animal TT5661 (2.683), seguido pelo TT5586 (2.319). O animal com maior número de INDELs

esclusivas na linhagem de postura foi o CC5 (242), seguido pelo CC88 (212) e na linhagem de corte o TT5921 (360), seguido pelo TT6270 (333). Os dois animais parentais TT5661 e CC88 (com maior efeito do QTL na família 7769, descrito anteriormente) foram considerados muito importantes para este estudo e mostraram um maior número de SNPs e INDELs únicas.

5.3 Variantes genéticas homozigotas e heterozigotas

Foram classificados todos os SNPs e INDELs em homozigotos e heterozigotos para as 18 galinhas (Figura 4). Considerando a média entre os 18 animais, do total de SNPs filtrados, 59% são homozigotos e 41% são heterozigotos e do total de INDELs filtradas para os 18 animais, 69% são homozigotos e 31% são heterozigotos.

De acordo com a Figura 4, foram observados mais SNPs e INDELs homozigotos do que heterozigotos, nas duas linhagens estudadas. Na linhagem de corte foi observado um alto percentual de SNPs e INDELs homozigotos (esta linhagem apresentou menor porcentagem de SNPs/INDELs heterozigotos em comparação com a linhagem de postura). Uma elevada porcentagem de SNPs heterozigotos (57%) também foi observada em uma linhagem de corte, no estudo desenvolvido por Fan et al. (2013).

Figura 4 (a, b) - Percentual de variantes filtradas classificadas em homozigotas e heterozigotas para os 18 animais. (a) Percentagem de SNPs únicas filtradas classificadas em homozigotas e heterozigotas, para os 18 animais. (b) Percentagem de INDELs únicos filtrados classificadas em homozigotos e heterozigotos, para os 18 animais

5.4 Anotação funcional

A anotação funcional de 123.985 SNPs únicos e 11.298 INDELs únicas foi realizada pelo *software* ANNOVAR e os resultados obtidos a partir dos SNPs e INDELs são mostrados nas Tabelas 5 e 6.

Tabela 5 – Anotação dos SNPs únicos filtrados, para as 18 galinhas (linhagem de corte e postura)

Variantes	Total	Porcentagem (%)
<i>Total de SNPs</i>¹	124.509	100
intergênicos	58.099	46,67
intrônicos	61.129	49,10
exônicos	1.226	0,98
<i>splicing</i>	7	0,01
ncRNA	14	0,01
UTR5'	211	0,17
UTR3'	1.016	0,82
<i>upstream</i> (1 kb)	1.393	1,12
<i>downstream</i> (1 kb)	1.392	1,12
Éxons		
sinônimos	840	0,68
Não sinônimos	376	0,30
<i>stopgain</i>	10	0,008

¹Para o cálculo da porcentagem foi considerado o número total de SNPs anotados.

A anotação de 124.487 SNPs classificou 1.226 em regiões exônicas, 61.129 em regiões intrônicas e 58.099 em regiões intergênicas além de outras regiões (*upstream*, *downstream*, etc.). SNPs localizados na região exônica foram classificados em não sinônimos, sinônimos e *stopgain*. Não foram detectados SNPs do tipo *stoploss*. A anotação de 11.805 INDELs classificou 21 em regiões exônicas, 6.106 em regiões intrônicas e 5.260 em regiões intergênicas além de outras regiões (como *upstream*, *downstream*, etc.) INDELs localizadas em regiões exônicas foram classificados como deleções *frameshift*, inserções *frameshift*, deleções *nonframeshift* e inserções *nonframeshift* (Tabela 6). Inserções/deleções de um número de nucleotídeos não divisível por 3, são classificadas como *frameshift*, alterando a sequência de aminoácidos e consequentemente o códon. Inserções/deleções de um número de nucleotídeos divisível por 3 são classificadas como *nonframeshift* e também podem alterar a sequência de aminoácidos.

Tabela 6 – Anotação dos INDELs únicos filtrados, para as 18 galinhas (linhagem de corte e postura)

Variantes	Total	Porcentagem (%)
<i>Todas as INDELs</i>¹	11.805	100
intergênicos	5.260	44,56
intrônicos	6.106	51,72
exônicos	21	0,18
<i>splicing</i>	2	0,02
ncRNA	2	0,02
UTR5'	19	0,16
UTR3'	134	1,14
<i>upstream</i> (1 kb)	113	0,96
<i>downstream</i> (1 kb)	148	1,25
Éxons		
Deleção <i>frameshift</i>	3	0,02
Inserção <i>frameshift</i>	12	0,10
Deleção <i>nonframeshift</i>	2	0,01
Inserção <i>nonframeshift</i>	4	0,03

¹Para o cálculo da porcentagem foi considerado o número total de INDELs anotadas.

Os SNPs não sinônimos, *stopgain* e *stoploss* são mutações em regiões que alteram o aminoácido que codifica a proteína. Estes SNPs podem alterar ou afetar a função da proteína, e são classificados como deletérios (com função da proteína alterada) e tolerado (não altera a função da proteína) (HUANG et al., 2009). Para este tipo de SNP é importante prever o escore SIFT por meio da ferramenta SIFT, que é um programa que prediz modificações no aminoácido que alteram a função de proteínas (NG e HENIKOFF, 2003).

A ferramenta SIFT usa um algoritmo para calcular o escore SIFT e classificar o SNP em deletério e tolerado. O algoritmo considera a posição na transcrição em que a alteração ocorreu (SNP) e em seguida, calcula a probabilidade do novo códon e consequente sequência de aminoácidos alterar a função da proteína (NG e HENIKOFF, 2003). Para a predição do escore SIFT para os SNPs não sinônimos e *stopgain*, foi utilizada a ferramenta VEP, como descrito anteriormente.

Foram anotados um total de 386 SNPs únicos para as 18 galinhas (não sinônimos e *stopgain*) com a ferramenta VEP. Obteve-se o total de 324 SNPs com escore SIFT

calculado e destes, 49 SNPs foram classificados como deletérios (isto é, tinham escore SIFT inferior ou igual a 0,05) (Tabela 7). Os restantes 274 SNPs foram classificados como tolerados porque tiveram uma pontuação maior do que 0,05 (NG e HENIKOFF, 2003). Mutações deletérias podem causar uma alteração na proteína e na sua função, e estas alterações podem influenciar as características fenotípicas.

Do total de 49 SNPs (Tabela 7) classificados como deletérios 11 foram encontrados exclusivamente na linha de postura e cinco SNPs foram identificados em apenas um animal (cada um deles em um animal diferente). Na linhagem de corte, 14 SNPs deletérios foram identificados exclusivamente nesta linhagem e cinco SNPs foram identificados em apenas um animal (cada um deles em um animal diferente), quatro deles em dois animais, e os restantes cinco SNPs foram encontrados em mais de três animais. Os 25 SNPs restantes foram identificados nas duas linhas, distribuídas em mais de três animais de cada linha, mas nenhum presente em todos os 18 animais (Tabela 7).

Do total de SNPs deletérios (49), 26 já estão presentes na base de dados dbSNP (NCBI). Os restantes 23 ainda não foram descritos, bem como, 115 SNPs tolerados. Sendo assim, um total de 138 SNPs não sinônimos e/ou *stopgain* foram descritos pela primeira vez em galinhas a partir da população F₂ Brasileira estudada.

Os parentais TT5661 e CC88 (com maior efeito do QTL na família, como descrito anteriormente) apresentaram 35 SNPs deletérios (Tabela 7). No animal CC88 da linhagem de postura, foram identificados quatro SNPs únicos, e um deles não foi descrito nas bases de dados públicas. No animal TT5661, foram identificados seis SNPs únicos, e cinco destes (previsto como SNPs deletérios) não foram descritos em bases de dados públicas.

Tabela 7 – Descrição dos 49 SNPs deletérios, identificados nos 18 animais (linhagem de corte e de postura)

(continua)

Posição (pb) ¹	Base trocada (SNP)	Aminoácido alterado	Escore SIFT	RefSNP	Gene	Linhagem	Animais
34.630.512	C/A	gGc/gTc	0,03	-	<i>CEP170</i>	Corte	TT5649*, TT5661*
37.040.403	C/A	caG/caT	0,03	-	<i>HEATR1</i>	Corte	TT5586*, TT6232*, TT6270*, TT5461
37.785.852	G/A	tCg/tTg	0,01	-	<i>TARBP1</i>	Corte	TT6232*, TT6270*
38.171.785	G/A	Cgt/Tgt	0,00	-	<i>LOC771163</i>	Corte	TT5586*
38.171.819	A/C	Tac/Gac	0,00	-	<i>LOC771163</i>	Corte	TT6232*, TT6270*
38.406.429	G/A	gCa/gTa	0,02	rs15328615	<i>NOVEL GENE</i>	Corte	TT5461*, TT5642*
38.966.052	G/A	Cct/Tct	0,03	-	<i>DISC1</i>	Corte	TT5642*, TT5649, TT5661, TT6270*
39.095.765	G/A	Ctc/Ttc	0,01	-	<i>SPRTN</i>	Corte	TT5661
39.434.655	T/C	cAc/cGc	0,03	-	<i>URB2</i>	Corte	TT5642, TT5649, TT5661*, TT6037, TT6232*, TT6270*
39.434.967	G/C	Ctg/Gtg	0,01	-	<i>URB2</i>	Corte	TT5461, TT5586, TT5661*, TT5921
39.815.176	G/A	Cgt/Tgt	0,00	rs14344092	<i>CAPN9</i>	Corte	TT5461
39.839.793	C/T	tGt/tAt	0,00	rs312436429	<i>CAPN9</i>	Corte	TT5661*, TT6232*, TT6270*
40.496.914	A/G	Tat/Cat	0,00	rs316694950	<i>WDR27</i>	Corte	TT5921*
41.356.984	A/G	gTg/gCg	0,04	rs313819539	<i>KIF25</i>	Corte	TT5921*, TT6037
34.015.812	G/A	gCg/gTg	0,00	-	<i>HNRNPU</i>	Postura	CC5*
34.079.656	C/G	Gaa/Caa	0,02	rs312714986	<i>C1ORF101</i>	Postura	CC37*, CC332*, CC570*
35.638.209	C/T	cGa/cAa	0,01	rs312499211	<i>FMN2</i>	Postura	CC1*, CC5*, CC88*, CC332*, CC1*, CC88, CC241, CC332, CC372, CC570
37.283.839	C/G	gGa/gCa	0,05	rs313722206	<i>LYST</i>	Postura	CC886*
37.532.903	G/T	cCc/cAc	0,00	-	<i>RBM34</i>	Postura	CC1*, CC88*, CC241*, CC372, CC570*, CC886*
38.315.047	A/C	tTg/tGg	0,03	rs317131899	<i>PCNXL2</i>	Postura	

Tabela 7 – Descrição dos 49 SNPs deletérios, identificados nos 18 animais (linhagem de corte e de postura)

(continuação)

Posição (pb) ¹	Base trocada (SNP)	Aminoácido alterado	Escore SIFT	RefSNP		Linhagem	Animais
38.924.862	A/G	Tat/Cat	0,02	rs15329391	<i>DISC1</i>	Postura	CC886*
39.321.156	G/A	aCg/aTg	0,05	rs10728012	<i>CCSAP</i>	Postura	CC570
40.020.999	T/A	Agt/Tgt	0,00	-	<i>GNPAT</i>	Postura	CC241*
40.496.909	C/T	Gga/Aga	0,03	-	<i>WDR27</i>	Postura	CC1*, CC88*, CC332*, CC886*
34.015.921	C/T	tGt/tAt	0,03	-	<i>HNRNPU</i>	Corte e Postura	CC5*, TT5461*
34.077.897	C/T	cGt/cAt	0,03	rs317346885	<i>C1ORF101</i>	Corte e Postura	CC1, CC37, CC88, CC332, CC570, TT5461*, TT5661, TT5921*
34.079.804	G/A	Cgt/Tgt	0,00	-	<i>C1ORF101</i>	Corte e Postura	CC1, CC5*, CC88, TT5661*, TT5921*
34.079.815	C/A	aaG/aaT	0,00	rs15322872	<i>C1ORF101</i>	Corte e Postura	CC5*, CC241*, CC372*, TT5649*, TT5921*, TT6232, TT6270
34.081.032	G/A	Cgg/Tgg	0,01	-	<i>C1ORF101</i>	Corte e Postura	CC1, CC88, CC241*, TT5642*, TT5661*, TT5921*
34.081.742	A/C	Tcc/Gcc	0,02	rs312859290	<i>C1ORF101</i>	Corte e Postura	CC1, CC88, CC241*, CC332*, TT5586*, TT5642*, TT5661*, TT5921*
34,084,356	G/C	Ctg/Gtg	0,00	rs13723758	<i>C1ORF101</i>	Corte e Postura	CC1, CC5, CC37*, CC88, CC241, CC332*, CC570*, CC886*, TT5461, TT5586, TT5642, TT5649, TT5661, TT5921, TT6037, TT6232, TT6270, CC37*, CC332*, CC372*, CC570*, CC886*, TT5586*, TT5649*, TT5661*
34.088.017	G/T	tCt/tAt	0,00	rs15322892	<i>C1ORF101</i>	Corte e Postura	CC37*, CC332*, CC372*, CC570*, CC886*, TT5586*, TT5649*, TT5661*
34.088.020	T/A	Act/Tct	0,01	rs15322894	<i>C1ORF101</i>	Corte e Postura	CC37*, CC332*, CC372*, CC570*, CC886*, TT5586*, TT5649*, TT5661*

Tabela 7 – Descrição dos 49 SNPs deletérios, identificados nos 18 animais (linhagem de corte e de postura)

(continuação)

Posição (pb) ¹	Base trocada (SNP)	Aminoácido alterado	Escore SIFT	RefSNP		Linhagem	Animais
34.088.097	C/T	Gat/Aat	0,00	rs312628425	<i>C1ORF101</i>	Corte e Postura	CC1, CC5, CC241, CC332*, CC372*, CC88, TT5461, TT5586*, TT5642, TT5649*, TT5921, TT6037, TT6232, TT6270
34.092.767	G/C	tCc/tGc	0,05	rs15322935	<i>C1ORF101</i>	Corte e Postura	CC1*, CC5*, CC37*, CC332*, CC570*, TT5586*, TT5649*, TT5661*
34.097.824	C/T	aGt/aAt	0,00	rs313199696	<i>C1ORF101</i>	Corte e Postura	CC1, CC5, CC37, CC88, CC241, CC332, CC886, TT5461, TT5921*, TT6232, TT6270
34.100.663	T/C	tAt/tGt	0,01	rs312273066	<i>C1ORF101</i>	Corte e Postura	CC1, CC5, CC332, CC886, TT5461, TT5649, TT5661, TT5921*, TT6037, TT6232, TT6270,
34.100.699	A/C	aTc/aGc	0,00	-	<i>C1ORF101</i>	Corte e Postura	CC5, CC332, CC570, CC886, TT5586, TT5661, TT5921*, TT6037, TT6232, TT6270
34.609.622	T/C	cAt/cGt	0,00	-	<i>CEP170</i>	Corte e Postura	CC37, TT5461
35.212.434	G/C	tgC/tgG	0,01	rs313142642	<i>PIGM</i>	Corte e Postura	CC241*, CC886*, TT5586*, TT5642*, TT5649, TT5661*, TT5921*, TT6037, TT6232, TT6270
36.914.484	C/T	Gag/Aag	0,02	-	<i>MTR</i>	Corte e Postura	CC241*, TT5642*, TT5649*, TT5661*, TT5921*, TT6037, TT6232*, TT6270*
37.042.513	T/C	Aag/Gag	0,01	-	<i>HEATR1</i>	Corte e Postura	CC37*, TT5649*, TT5661, TT6037, TT6232*, TT6270*
37.788.171	C/T	aGc/aAc	0,01	rs16733664	<i>TARBP1</i>	Corte e Postura	CC37*, TT5642*, TT5649*

Tabela 7 – Descrição dos 49 SNPs deletérios, identificados nos 18 animais (linhagem de corte e de postura)

(conclusão)

Posição (pb) ¹	Base trocada (SNP)	Aminoácido alterado	Escore SIFT	RefSNP		Linhagem	Animais
38.890.502	T/C	Aaa/Gaa	0,02	-	<i>DISC1</i>	Corte e Postura	CC88*, CC241*, TT5461*, TT5586, TT5921
40.160.345	C/T	Gaa/Aaa	0,03	rs13675883	<i>FAM120B</i>	Corte e Postura	CC241*, CC372*, CC570*, TT5649*
40.461.109	C/T	cGc/cAc	0,04	rs16258880	<i>ERMARD</i>	Corte e Postura	CC1, CC5, CC88, CC241, CC332, CC372, CC570, CC886, TT5461, TT5586*, TT5642, TT5921*, TT6037, TT6232, TT6270

¹ Posição do SNP no GGA3 de acordo com última sequência referência da galinha (Gallus_gallus-4.0, NCBI).

*SNPs classificados como heterozigotos.

5.5 Análise de genes a partir dos SNPs deletérios

Um total de 77 genes foi obtido a partir dos SNPs não sinônimos e *stopgain* e INDELs *frameshift* e não *frameshit*. Os genes com SNPs deletérios foram discutidos/relacionados a partir da informação disponibilizada no banco de dados OMIM (NCBI). Os genes expressos no tecido adiposo e/ou relacionados ao metabolismo de lipídios foram selecionados, similar à metodologia adotada por Ahsan et al. (2013) de análise das informações de cada gene, disponível nos bancos de dados. Os genes ainda foram pesquisados na literatura disponível.

Quatro genes foram relacionados ao metabolismo de lipídio (Tabela 8) sendo eles: *similar to mixed lineage kinase 4* (LOC771163); *egl-9 family hypoxia-inducible factor 1* (EGLN1); *glyceronephosphate O-aciltransferase* (GNPAT); *family with sequence similarity 120B* (FAM120B).

Segundo Seit-Nebi et al. (2012), o gene LOC771163 atua como regulador negativo do gene *tumor necrosis factor* (TNF). O gene TNF está relacionado à síntese e secreção de citocinas e estas, afetam o metabolismo de lipídios (ROSMOND et al., 2001). Segundo os mesmos autores, o aumento na expressão do gene TNF pode resultar em maior deposição de gordura (obesidade) em seres humanos. Sendo assim, mutações deletérias no gene LOC771163 podem induzir a deposição de gordura na carcaça em galinhas.

O gene EGLN1 está relacionado à homeostase de oxigênio em mamíferos (Epstein et al., 2001) e ainda está relacionado ao mecanismo de regulação da diferenciação celular em adipócitos e osteoblastos em um estudo desenvolvido com seres humanos (Wang et al., 2013). O gene GNPAT também atua na diferenciação de células a partir de oócitos em precursores iniciais de lipídios, fonte de energia do embrião (Bertolesi et al., 2012).

O gene GNPAT foi descrito e relacionado com a síntese e metabolismo de glicerofosfolipídios ou fosfoglicerídios. Fosfoglicerídios são os componentes mais abundantes das membranas das células biológicas (Bertolesi et al., 2012) e participam de várias funções e atividades celulares tais como sinalização, transporte de biomoléculas ou ainda, o armazenamento de energia em seres humanos (Acosta-Rodríguez et al., 2013). Ainda em seres humanos, o ácido

fosfatídico (PA) é um componente intermédio da via do metabolismo de glicerofosfolipídios, responsável pela síntese de triacilglicerol (Bertolesi et al., 2012).

Além disso, o gene *GNPAT* é responsável pela síntese de duas enzimas encontradas nos peroxissomas, associadas com a síntese de lípidos, que são: aciltransferase dihidroxiacetona-fosfato (DHAPAT) (Liu et al., 2005) e glyceronephosphate O-aciltransferase (GNPAT) (Mizuno et al., 2013). Segundo Bertolesi et al. (2012), em humanos, estas duas enzimas regulam o metabolismo de glicerofosfolipídios na fase embrionária.

Em um estudo com camundongos, Rodemer et al. (2003) identificaram uma mutação no gene *GNPAT*, associada com o aumento dos níveis de ácidos graxos e de colesterol. Teigler et al. (2009), também estudaram uma deleção no mesmo gene, associada a uma redução nos níveis de lípidos totais, também em camundongos. Em galinhas ainda não foram realizados estudos de associação neste gene e as mutações identificadas no atual estudo são candidatas para estudos de associação com a deposição de gordura em galinhas.

O gene *FAM120B* está relacionado ao receptor PPAR responsável pelo incremento na síntese dos adipócitos (Weivoda e Hohl., 2011). Assim, a expressão deste gene pode estar associada com uma variação na deposição de gordura na carcaça em frangos. Ainda de acordo Resnyk et al. (2013), os genes associados com a deposição de gordura tem geralmente a sua ação mediada pelo receptor PPAR e este receptor é sintetizado por genes envolvidos no metabolismo energético, a partir de células adiposas, distribuídos por todo o genoma, tais como o *PPARGC1A*, *PPARGC1B* e *PPRC1*, localizados no GGA4, 13 e 6, respectivamente.

Em um estudo com camundongos, Li et al. (2007), relataram o gene *FAM120B* expresso no tecido adiposo. Ainda segundo os mesmos autores, a expressão aumentada deste gene (observada no tecido adiposo) ocasionou a diferenciação dos pré-adipócitos em adipócitos. Assim, mutações deletérias no gene *FAM120B* poderiam causar menor deposição de gordura na carcaça.

No gene *chromosome 1 open reading frame 101 (C1orf101)* foram identificados 14 SNPs deletérios (Tabela 7) e destes, três não foram descritos anteriormente. O número de mutações deletérias identificadas no mesmo gene pressupõe uma maior variabilidade nos transcritos e consequente variação fenotípica. Estudos desenvolvidos em seres humanos, por Zhang et al. (2011), apontaram a atuação deste gene na diferenciação de queratinócitos (células diferenciadas do tecido epitelial) em melanócitos (células produtoras de melanina). Não foram encontrados estudos de correlação entre a função/expressão deste gene e o metabolismo de lipídios.

As mutações identificadas em cada um dos genes estão descritas na Tabela 8.

Tabela 8 – Genes selecionados de acordo com sua função, relacionados ao metabolismo de lipídios

Gene	SNP deletério	Posição (pb)	Associação (G.O) ¹	Linhagem	Animais
LOC771163	c.575G>A	38.171.785	Não descrito	Corte	TT5586 (Heterozigoto)
	c.541C>T	38.171.819	Não descrito	Corte	TT6232 (Heterozigoto) TT6270 (Heterozigoto)
					CC886 (Heterozigoto)
EGLN1	c.29C>A	39.053.960	Oxi-redução	Corte e Postura	TT5586 (Homozigoto) TT6232 (Heterozigoto) TT5461 (Heterozigoto)
GNPAT	c.377C>T	40.020.999	Organização da membrana celular	Postura	CC241 (Heterozigoto)
FAM120B	c.1496T>G	40.160.345	Diferenciação celular e regulação da transcrição.	Corte e Postura	CC241 (Heterozigoto)
					CC372 (Heterozigoto)
					CC570 (Heterozigoto)
					TT5649 (Heterozigoto)

¹Gene Ontology.

Considerando os animais parentais das famílias 7822, 7769 e 7716 com maior efeito do QTL (descrito anteriormente), e os genes selecionados na Tabela 8, foram identificados dois SNPs heterozigotos no animal de corte TT6232 sendo um no gene *LOC771163* (c.541C>T) e um no gene *EGLN1* (c.29C>A); um SNP heterozigoto no animal de corte TT5649 no gene *FAM120B* (c.1496T>G) e um SNP heterozigoto no animal de postura CC886 também no gene *EGLN1* (c.29C>A).

As mutações heterozigotas, classificadas como deletérias, podem relacionar-se com a variabilidade entre os animais e consequentemente entre as linhagens. Os SNPs identificados nos genes relacionados ao metabolismo de lipídios são mutações candidatas a estudos posteriores de associação.

Além da análise da lista de 77 genes, a identificação de vias relacionadas ao metabolismo de lipídios é uma ferramenta interessante na busca e identificação de genes e mutações causais.

5.6 Vias metabólicas relacionadas às variantes localizadas em regiões exônicas

Foram selecionados os genes (Ensembl GeneID's) a partir de todos os SNPs e INDELs localizados em regiões exônicas (classificados como SNPs não sinônimos e *stopgain*, INDELs *frameshift*). Estes *GeneID's* foram obtidos pela anotação funcional no ANNOVAR e foram utilizados para identificar as vias metabólicas relacionadas com os genes por meio da ferramenta de bioinformática DAVID. A ferramenta DAVID, utilizada para analisar a lista de genes e posterior identificação das vias metabólicas correlacionadas, permite compreender o papel biológico de genes e também as mutações encontradas em cada gene (HUANG et al., 2009).

Para os SNPs não sinônimos foram obtidos 72 genes, sendo que 10 deles participam de vias metabólicas, e dos SNPs *stopgain*, seis genes foram analisados e um foi descrito e relacionado a uma via metabólica. Das INDELs *frameshift*, oito genes foram analisados e um deles foi relacionado à uma via metabólica. As 12 vias metabólicas identificadas estão descritas na Tabela 9.

Tabela 9 – Lista de genes relacionados às vias metabólicas pelo *KEGG Pathway Database*

Variante	Gene	Via metabólica (gga ¹)
SNP não sinônimo	<i>TAF5L</i>	Fatores de transcrição basais (gga03022)
	<i>AGT</i>	Sistema renina-angiotensina (gga04614)
	<i>CCR6</i>	Interação receptor citoquina-citocina (gga04060)
	<i>EXO1</i>	Reparo de erros (gga03430)
	<i>FMN2</i>	Formação do eixo dorso ventral (gga04320)
	<i>GNPAT</i>	Metabolismo de glicerofosfolípidios (gga00564)
	<i>HNRNPU</i>	Spliceossomo (gga03040)
	<i>PIGM</i>	Biossíntese da âncora glicosilfosfatidilinositol (GPI) (gga00563)
	<i>RPS6KA2</i>	Via de sinalização da MAPK (gga04010)
	<i>THBS2</i>	Via de sinalização da TGF-beta (gga04350)
SNP <i>stopgain</i>	<i>TAF5L</i>	Fatores de transcrição basais (gga03022)
INDEL <i>frameshift</i>	<i>GGPS1</i>	Biossíntese de espinha dorsal (gga00900)

¹Código identificador da via metabólica, descrito para *Gallus gallus*, pelo *KEGG Pathway Database*.

A partir das 12 vias metabólicas identificadas pela ferramenta DAVID (Tabela 9), foram selecionados apenas os que foram relacionados ao metabolismo lipídico e/ou deposição de gordura, devido a região de QTL estudada no GGA3, que foi anteriormente associada com deposição de gordura (Tabela 10). Os genes e as suas vias relacionadas com a deposição de gordura foram: *glyceronephosphate O-aciltransferase (GNPAT)* associado com o metabolismo dos lípidos na via de metabolismo de glicerofosfolípidios (gga00564); *trombospondina 2 (THBS2)* associado com o equilíbrio homeostático dos níveis de TGF-Beta na via de sinalização (gga04350) e *geranilgeranil-difosfato-sintetase 1 (GGPS1)* associado com o metabolismo lipídico em terpenóides na via de biossíntese de espinha dorsal (gga00900).

Tabela 10 – Genes selecionados de acordo com sua função, relacionados ao metabolismo de lipídios

Gene	Variante	Posição (pb)	Associação (G.O) ¹	Linhagem	Animais
<i>GNPAT</i>	nsSNP deletério	40.020.999	Organização da membrana celular	Postura	CC241 (heterozigoto)
<i>THBS2</i>	nsSNP tolerado	40.680.860	Adesão celular e biológica	Postura	CC886 (heterozigoto)
		40.692.317	Adesão celular e biológica	Corte	TT6232 (homozigoto) TT6270 (homozigoto)
<i>GGPS1</i>	INDEL frameshift	37.421.178	Metabolismo de isoprenóides, biosíntese de isoprenóides e de lipídios	Postura e Corte	CC001 (homozigoto)
					CC241 (homozigoto)
					CC37 (homozigoto)
					CC570 (homozigoto)
					CC5 (homozigoto)
					CC886 (homozigoto)
					CC88 (homozigoto)
					TT5461 (homozigoto)
					TT5642 (homozigoto)
					TT5661 (homozigoto)
					TT5921 (homozigoto)
					TT6037 (homozigoto)
					TT6270 (homozigoto)

¹Gene Ontology.

O gene *GNPAT* foi descrito e relacionado com o metabolismo de lipídios, como discutido no tópico anterior. A participação dele na via do metabolismo de glicerofosfolipídios indica este gene como candidato ao estudo de associação para posterior seleção genômica contra deposição de gordura excessiva em aves.

O gene *THBS2* foi relacionado com a síntese da proteína da matriz celular trombospondina-2 (MATOS et al., 2013). Esta proteína estimula a síntese de osteoblastos (osteoblastogênese), que são as células progenitoras da matriz óssea. A adipogênese e a osteoblastogênese são mutuamente reguladas e, esta regulação recíproca permite inferir que thrombospondin-2 inibe a adipogênese (SHITAYE et al. 2010). Segundo os mesmos autores, a expressão aumentada do gene *THBS2* representa um maior acúmulo de células adiposas durante a adipogênese conduzindo à maior deposição de gordura. Em um estudo com camundongos, Chun (2012) analisou a expressão do gene *THBS2* e a condição corporal dos indivíduos, e os camundongos com expressão nula deste gene apresentaram maior resistência à obesidade, mesmo quando submetidos à dietas hiperlipídicas. Este gene foi diferencialmente expresso na gordura abdominal em um estudo da expressão de genes no tecido adiposo em linhagens divergentes de galinhas quanto à deposição de gordura (RESNYK et al., 2013). As variantes identificadas neste gene podem estar relacionadas à uma maior deposição de gordura abdominal em frangos.

O gene *GGPS1* foi relacionado com o metabolismo dos terpenóides e também está envolvido na síntese da enzima geranylgeranil (GGP) sintetase. Esta enzima é sintetizada no primeiro passo da via metabólica de terpenóides e regula a biossíntese de alguns derivados de esteróides incluindo colesterol (o principal esteroide sintetizado em animais) (CORTES et al., 2013). A enzima geranylgeranil GGP sintetase medeia a diferenciação de células em osteoblastos e adipócitos, portanto, quando em baixas concentrações da enzima, a diferenciação em osteoblastos é estimulada e quando em concentrações mais elevadas, ativa o receptor (PPAR - gama2) responsável pelo incremento na diferenciação em adipócitos (WEIVODA e HOHL., 2011).

Assim, a expressão do gene *GGPS1* pode estar associada com uma variação na deposição de gordura na carcaça em frangos.

Para as variantes identificadas nestes três genes, o nsSNP deletério (c.377C>T) no gene *GNPAT* foi identificado em apenas um animal de postura heterozigoto. A linhagem de postura foi selecionada durante muitas gerações para um menor peso corporal, conforme descrito em Rosário et al. (2009), sugerindo uma menor deposição de gordura na carcaça, tendo em conta a correlação genética e fenotípica positiva entre o peso corporal e a deposição de gordura (WANG et al., 2012). Segundo Campos et al. (2009), a linhagem de corte (TT) mostrou uma deposição de gordura 15 vezes maior do que a linhagem de postura (CC), sendo assim, a mutação c.377C>T no gene *GNPAT*, apesar de ter sido identificada em apenas um dos 18 animais, é candidata para futuros estudos de associação de deposição de gordura na carcaça em galinhas.

Em um animal da linhagem de postura foi identificado um SNP heterozigoto (c.52C>T) no gene do *THBS2*, e este gene, foi relacionado ao metabolismo lipídico por Shitaye et al. (2010) e diferencialmente expresso na gordura abdominal por Resnyk et al. (2013). Este animal parental pertence à família 7716, com efeito do QTL na família de 1,12 para peso de gordura abdominal e 0,05 para percentual de gordura abdominal (CAMPOS et al., 2009). Também no gene *THBS2*, o animal TT6232 apresentou um SNP homozigoto (c.1553T>C). Este animal parental pertence à família 7822, com efeito do QTL na família de 3,48 para peso de gordura abdominal e 0,23 para o percentual de gordura abdominal (CAMPOS et al., 2009).

No gene *GGPS1* uma inserção (c.285dupA) foi identificada em 13 dos 18 animais sequenciados. Os parentais TT5661 e CC88, que compreendem a família 7769 (com maior efeito do QTL na família) apresentaram esta inserção, entretanto, ela foi classificada como homozigota nos dois animais. A descoberta de mutações no gene *GGPS1* relacionadas com gordura pode nos ajudar a compreender a função energética do metabolismo de isoprenóides de acordo com Weivoda e Hohl (2011), e consequentemente, a deposição de gordura na carcaça em frangos.

Os animais TT5661 e CC88 apresentaram ainda, um grande número de SNPs e INDELs únicos na linhagem de corte e postura (descritos anteriormente), no entanto, a maioria destas mutações ainda não foi identificada em vias metabólicas relacionadas com o metabolismo lipídico pela ferramenta de bioinformática DAVID. A análise de mutações presentes apenas nas regiões exônicas com a ferramenta DAVID pode ter desconsiderado uma grande parte das variantes possivelmente relacionados ao metabolismo lipídico e consequente deposição de gordura na carcaça.

6 CONCLUSÕES

A redução da gordura abdominal em frangos tem grande importância para os programas de melhoramento, considerando seu grande impacto sobre os custos de produção e seu impacto negativo sobre a qualidade da carne e eficiência de produção. Este estudo relata a caracterização de variantes de boa qualidade presentes numa região de QTL no GGA3 que podem ter efeito sobre o peso e percentagem da gordura abdominal em galinhas. Foram identificadas mutações deletérias possivelmente relacionadas ao peso e percentagem de gordura abdominal em galinhas, candidatas a estudos posteriores de associação, a partir da validação (genotipagem) em outras populações. Mutações foram identificadas nos genes *LOC771163*, *EGLN1*, *GNPAT*, *FAM120B*, *THBS2* e *GGPS1*, que foram anteriormente relacionados ao metabolismo de lipídios.

REFERÊNCIAS

ABASHT, B.; DEKKERS, J.C.M.; LAMONT, S.J. Review of quantitative trait loci identified in the chicken. **Poultry Science**, Savoy, v. 85, p. 2079-2096, 2006.

ABPLANALP, H. Inbred lines as genetic resources of chickens. **Poultry Science**, Savoy, v. 4, p. 29–39, 1992.

ACOSTA-RODRÍGUEZ, V.A.; MÁRQUEZ, S.; SALVADOR, G.A.; PASQUARÉ, S.J.; GORNÉ, L.D.; GARBARINO-PICO, E.; GIUSTO, N.M.; GUIDO, M.E. Daily rhythms of glycerophospholipid synthesis in fibroblast cultures involve differential enzyme contributions. **Journal of Lipid Research**, Bathesda, p. 1–47, 2013.

AHSAN, M.; LI, X.; LUNDBERG, A.E.; KIERCZAK, M.; SIEGEL, P.B.; CARLBORG, O.; MARKLUND, S. Identification of candidate genes and mutations in QTL regions for chicken growth using bioinformatic analysis of NGS and SNP-chip data. **Frontiers in Genetics**, Lousanne, v. 4, n. 226, p. 1-8, 2013.

ALBERS, C.A.; LUNTER, G.; MACARTHUR, D.G.; MCVEAN, G.; OUWEHAND, W.H.; DURBIN, R. Dindel: accurate indel calls from short-read data. **Genome Research**, New York, v. 21, n. 6, p. 961–973, 2011.

AMARAL, A.J.; MEGENS, H.J.; KERSTENS, H.H.D.; HEUVEN, H.C.M.; DIBBITS, B.; CROOIJMANS, R.P.M.; DEN DUNNEN, J.T.; GROENEN, M.A.M. Application of massive parallel sequencing to whole genome SNP discovery in the porcine genome. **BMC Genomics**, London, v. 10, p. 374, 2009.

ANDERSSON, L. Molecular consequences of animal breeding. **Current Opinion In Genetics & Development**, London, v. 23, n. 3, p. 295–301, 2013.

BAÉZA, E., E LE BIHAN-DUVAL, E. Chicken lines divergent for low or high abdominal fat deposition: a relevant model to study the regulation of energy metabolism. **Animal : An International Journal Of Animal Bioscience**, Cambridge, v.7, n.6, p. 965–973, 2013.

BEFFAGNA, G.; OCCHI, G.; NAVA, A.; VITIELLO, L.; DITADI, A.; BASSO, C.; BAUCE, B.; ET AL. Regulatory mutations in transforming growth factor-beta3 gene cause arrhythmogenic right ventricular cardiomyopathy type 1. **Cardiovascular Research**, London, v. 65, n. 2, p. 366–373, 2005.

BERRI, C.; WACRENIER, N.; MILLET, N.; LE BIHAN-DUVAL, E. Effect of selection for improved body composition on muscle and meat characteristics of broilers from experimental and commercial lines. **Poultry Science**, Savoy, v. 80, n. 7, p. 833–838, 2001.

BERTOLES, G.E.; IANNATONE, S.; JOHNSTON, J.; ZAREMBERG, V.; MCFARLANE, S. Identification and expression analysis of GPAT family genes during early development of *Xenopus laevis*. **Gene Expression Patterns**, Amsterdam, v. 12, n. 7/8, p. 219–227, 2012.

BIHAN-DUVAL E.L.; MILLET N.; REMIGNON H. Broiler meat quality: effect of selection for increased carcass quality and estimates of genetic parameters. **Poultry Science**, Savoy, v. 78, p. 822–826, 1999.

BIOMART. **Consula de genes de acordo com Ensembl 73 Genes (WTSI, UK)**. Disponível em: <<http://central.biomart.org/>>. Acesso em: 24 de nov.2013.

BRANDSTRÖM, M.; ELLEGREN, H. The genomic landscape of short insertion and deletion polymorphisms in the chicken (*Gallus gallus*) Genome: a high frequency of deletions in tandem duplicates. **Genetics**, Bathesda, v.176, n. 3, p. 1691–1701, 2007.

BRAUNSCHWEIG, M.H. Mutations in the bovine ABCG2 and the ovine MSTN gene added to the few quantitative trait nucleotides identified in farm animals: a mini-review. **Journal of Applied Genetics**, Poznań, v. 51, n. 3, p. 289–297, 2010.

BRITTEN, R.J. Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, n. 99, p. 13633–13635, 2002.

BURT, D.W. Applications of biotechnology in the poultry industry. **International Journal of Poultry Science**, Faisalabad, n. 58, p. 5-13, 2002.

_____. Chicken genome: current status and future opportunities. **Genome Research**, New York, v. 15, n. 12, p. 1692–1698, 2005.

CAHANER, A.; NITSAN, Z. Evaluation of simultaneous selection for live body weight and against abdominal fat in broilers. **Poultry Science**, Savoy, v.64, p.1257- 263, 1985.

CAMPOS, R.L.R.; NONES, K.; LEDUR, M.C.; MOURA, A. S. A M. T.; PINTO, L. F. B.; AMBO, M.; BOSCHIERO, C.; RUY, D.C.; BARON, E.E.; NINOV, K.; ALTENHOFEN, C.A.B.; SILVA, R.A.M.S.; ROSÁRIO, M.F.; BURT, D.W.; COUTINHO, L.L. . Quantitative trait loci associated with fatness in a broiler-layer cross. **Animal Genetics**, Malden, v. 40, n. 5, 729–736, 2009.

CARLBORG, O.L.; JACOBSSON, P.; AHGREN, P.; SIEGEL, P.B.; ANDERSSON, L. Epistasis and the release of genetic variation during long-term selection. **Nature Genetics**, New York, v. 38, n. 418–420, 2006.

CARLSON, C.S.; EBERLE, A.M.; RIEDER, J.M.; YI, Q.; KRUGLYAK, L.; NICKERSON, D.A. Selecting a maximally informative set of singlenucleotide polymorphisms for association analyses using linkage disequilibrium. **The American Journal of Human Genetics**, Cambridge, v. 74, p.106–120, 2004. CARVALHO, M.C. da C.G.; SILVA, D.C.G. da. Next generation DNA sequencing and its applications in plant genomics, **Ciência Rural**, Santa Maria, v. 40, n. 3, p. 735–744, 2010.

CHAMBERS J.R. Genetics of growth and meat production in chickens. **Poultry Breeding and Genetics (Elsiever)**, Amsterdam, p. 599–643, 1990.

CHICKEN QTLDB. Consulta de QTLs no genoma da galinha. Disponível em: <<http://www.animalgenome.org/cgi-bin/QTLdb/GG/index>>. Acesso em: 28 Out. 2013.

CHOCT; NAYLOR, A.; HUTTON, O.; NOLAN J. **Increasing efficiency of lean tissue composition in broiler chickens**. 2000. Disponível em: <http://www.rirdc.gov.au/reports/CME/98-123.pdf>. Acesso em 05 Maio 2013.

CHUN, T.-H. Peri-adipocyte ECM remodeling in obesity and adipose tissue fibrosis. **Adipocyte**, Austin, v. 1, n. 2, p. 89–95, 2012.

CLOP. A; MARCQ. F; TAKEDA. H; PIROTTIN. D; TORDOIR. X; BIBE. B; BOUIX. J; CAIMENT. F; ELSEN. J; EYCHENNE. F; LARZUL, C.; LAVILLE, E.; MEISH, F.; MILENKOVIC, D.; TOBIN, J.; CHARLIER, C.; GEORGES, M. A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. **Nature Genetics**, New York, v. 38, p. 813–818, 2006.

COOPER, D.A.; LU, S.C.; VISWANATH, R. The structure and complete nucleotide sequence of the avian lipoprotein lipase gene. **Biochimica et Biophysica Acta**, Amsterdam, n. 1129, p. 166–171, 1992.

CORTES, V.A; BUSSO, D.; MARDONES, P.; MAIZ, A.; ARTEAGA, A.; NERVI, F.; RIGOTTI, A. Advances in the physiological and pathological implications of cholesterol. **Biological reviews of the Cambridge Philosophical Society**, Cambridge, v. 4, n. 88, p. 825–843, 2013.

COUTINHO, L.; ROSARIO, M.F.; JORGE, E.C. Biotecnologia animal. **Estudos Avançados**, São Paulo, v.24, n.70, 2010.

CROOIJMANS, R.P.; GROEN, A.F.; VAN KAMPEN, A.J.; VAN DER BEEK, S.; VAN DER POEL, J.J.; GROENEN, M.A. Microsatellite polymorphism in commercial broiler and layer lines estimated using pooled blood samples. **International Journal of Poultry Science**, Faisalabad, n. 75, p. 904–909, 1996.

CRUZ, V.A.R.; VENTURINI, G.C.; PEIXOTO, J.O.; LEDUR, M.C.; SCHMIDT, G.S.; MUNARI, D.P. **Estimativa de parâmetros genéticos para peso corporal, gordura abdominal e peso de peles em linhagem pura de frangos de corte**. 2011. Disponível em: <

<http://ainfo.cnptia.embrapa.br/digital/bitstream/item/53466/1/estimativas-de-parametros-geneticos-para-peso-corporal.pdf>>. Acesso em 27 Nov. 2013.

DEMEURE, O.; DUCLOS, M.J.; BACCIU, N.; LE MIGNON, G.; FILANGI, O.; PITEL, F.; BOLAND, A.; LAGARRIGUE, S.; COGBURN, L.A.; SIMON, J.; LE ROY, P.; LE BIHAN-DUVAL, E. Genome-wide interval mapping using SNPs identifies new QTL for growth, body composition and several physiological variables in an F2 intercross between fat and lean chicken lines. **Genetics Selection Evolution**, Versailles, v. 45, n. 1, p. 1-12, 2013.

DENG, G.F.; LIU, S.J.; SUN, X.S.; SUN, W.W.; ZHAO, Q.H.; LIAO, W.P.; YI, Y.H.; ET AL. A conserved region in the 3' untranslated region of the human LIMK1 gene is critical for proper expression of LIMK1 at the post-transcriptional level. **Neuroscience Bulletin**, Beijing, v.3, n. 29, p. 348–354, 2013

DODGSON, J. B.; DELANY, M.E.; CHENG, H.H. Poultry genome sequences: progress and outstanding challenges. **Cytogenetic and Genome Research**, Basel, v. 1, n. 134, p. 19–26, 2011.

DORAN, A.G.; CREEVEY, C.J. Snpdat: Easy and rapid annotation of results from de novo snp discovery projects for model and non-model organisms. **BMC Bioinformatics**, London, v. 1, n. 14, p. 1-6, 2013.

ECK, S.H.; BENET-PAGÈS, A.; FLISIKOWSKI, K.; MEITINGER, T.; FRIES, R.; STROM, T.M. Whole genome sequencing of a single *Bos taurus* animal for single nucleotide polymorphism discovery. **Genome Biology**, London, v. 8, n. 10, p. 1-8, 2009.

EITAN, Y.; SOLLER, M. Selection induced genetic variation. *Evolutionary Theory and Processes: Modern Horizons*. **Papers in Honor of Eviatar Nevo**, Haifa, p.153–176, 2004.

EPSTEIN, A.C.R.; GLEADLE, J.M.; MCNEILL, L.A.; HEWITSON, K.S.; O'ROURKE, J.; MOLE, D.R.; MUKHERJI, M.; METZEN, E.; WILSON, M.I.; DHANDA, A.; TIAN, Y.-M.; MASSON, N.; HAMILTON, D.L.; JAAKKOLA, P.; BARSTEAD, R.; HODGKIN, J.; MAXWELL, P.H.; PUGH, C.W.; SCHOFIELD, C.J.; RATCLIFFE, P.J.C. *elegans* EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. **Cell**, Cambridge, v. 107, p. 43-54, 2001.

FAN, W.L.; NG, C.S.; CHEN, C.F.; LU, M.Y. J.; CHEN, Y.H.; LIU, C.J.; WU, S.M.; CHEN, C.K.; CHEN, J.J.; MAO, C.T.; LAI, Y.T.; LO, W.S.; CHANG, W.H.; LI, W.H. Genome-wide patterns of genetic variation in two domestic chickens. **Genome Biology And Evolution**, Oxford, v. 7, n. 5, p. 1376–1392, 2013.

PINTO, L.F.B.; PACKER, I.U.; LEDUR, M.C.; MOURA A.S.A.M.T.; NONES, K.; COUTINHO, L.L. Mapping quantitative trait loci in *Gallus gallus* using principal components utilização de componentes principais. **Revista Brasileira de Zootecnia**, Viçosa, p. 2434–2441, 2010.

FURTADO, D; MOTA, J.K.M.; NASCIMENTO, J.W.B.; SILVA, V.R.; TOTA, L.C.A. Produção de ovos de matrizes pesadas criadas sob estresse térmico Egg production of broiler breeders reared under heat stress. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v.15, n.7, p.748–753, 2011.

GAYA, L.D.G.; MOURÃO, G.B.; FERRAZ, J.B.S. Estudo genético da deposição de gordura abdominal e de características de desempenho, carcaça e composição corporal em linhagem macho de frangos de corte. **Ciência Rural**, Santa Maria, v.36, n.2, p. 709-716, 2006.

GODBOUT, R. Identification and characterization of transcripts present at elevated levels in the undifferentiated chick retina. **Experimental Eye Research**, Amsterdam, n. 56, p. 95–106, 1993.

GROENEN, M.A.; WAHLBERG, P.; FOGGIO, M.; CHENG, H.H.; MEGENS, H-J.; CROOIJMANS, R.P.M.A.; BESNIER, F.; LATHROP, M.; MUIR, W.M.; WONG, G.K-S.; GUT, I.; ANDERSSON, L. A highdensity SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. **Genome Research**, New York, n. 19, p. 510–519, 2009.

GUNNARSSON, U; KERJE, S; BED'HOM, B; SAHLQVIST, A.S; EKWALL, O; TIXIER-BOICHARD, M; KA'MPE, O; ANDERSSON, L. The Dark brown plumage color in chickens is caused by an 8.3 kb deletion upstream of SOX10. **Pigment Cell & Melanoma Research**, Yale, n. 24, p. 268-274, 2011.

HILL, W.G. Acentury of corn selection. **Science**, Washington, n. 307, p. 683–684, 2005.

HILLIER, L.W.; MILLER, W.; BIRNEY, E.; WARREN, W.; HARDISON, R.C.; PONTING, C.P.; BORK, P.; BURT, D.W.; GROENEN, M.A.; DELANY, M.E. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. **Nature**, London, n. 432, p. 695–716, 2004.

HUANG, D.W.; SHERMAN, B.T.; LEMPICKI, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. **Nature Protocols**, London, v. 1, n. 4, p. 44–57, 2009.

HUANG, X.; FENG, Q.; QIAN, Q.; ZHAO, Q.; WANG, L.; WANG, A.; GUAN, J.; et al. High-throughput genotyping by whole-genome resequencing. **Genome research**, New York, p. 1068–1076, 2009.

IKEOBI, C.O.N.; WOOLLIAMS, J.A.; MORRICE, D.R.; LAW, A.; WINDSOR, D.; BURT, D.W.; HOCKING, P.M. Quantitative trait loci affecting fatness in the chicken. **Animal Genetics**, New York, v. 6, n. 33, p. 428–435, 2002.

IMSLAND, F.; FENG, C.G.; BOIJE, H.; BED'HOM, B.; FILLON, V.; DORSHORST, B.; RUBIN, C.J.; LIU R.R.; GAO, Y.; GU, X.R. The Rose-comb mutation in chickens constitutes a structural rearrangement causing both altered comb morphology and defective sperm motility. **PLOS Genetics**, San Francisco, v. 8, n. 6, p. 1-12, 2012.

JENNEN, D.G.J.; VEREIJKEN, A L.J.; BOVENHUIS, H.; CROOIJMANS, R. P. M.; VEENENDAAL, A.; VAN DER POEL, J.J.; GROENEN, M.A.M. Detection and localization of quantitative trait loci affecting fatness in broilers. **Poultry Science**, Savoy, v. 3, n. 83, p. 295–301, 2004.

JIANG, H.; WONG, W.H. SeqMap: mapping massive amount of oligonucleotides to the genome. **Bioinformatics**, Oxford, v. 24, p. 2395–2396, 2008.

KATO, K. Impact of the next generation DNA sequencers. **International Journal Of Clinical And Experimental Medicineis**, Madison, v. 2, p. 193–202, 2009.

KESSLER, A.M.; SNIZEK JR.; P.N.; BRUGALLI, I. **Manipulação da quantidade de gordura na carcaça de frangos**. In: CONFERÊNCIA APINCO DE CIÊNCIA e TECNOLOGIA AVÍCOLAS, 2000. Campinas, SP **Anais...** Campinas:APINCO SP, 2000. P 107-133.

KRANIS, A.; GHEYAS, A. A; BOSCHIERO, C.; TURNER, F.; YU, L.; SMITH, S.; TALBOT, R.; ET AL. Development of a high density 600K SNP genotyping array for chicken. **BMC Genomics**, London, v. 1, n. 14, p. 59, 2013.

KUERSTEN, S.; GOODWIN, E.B. The power of the 3' UTR: translational control and development. **Nature Reviews**, London, v. 8, n. 4, p. 626–637, 2003.

KUMAR, S.; BANKS, T.W.; CLOUTIER, S. SNP Discovery through Next-Generation Sequencing and Its Applications. **International Journal Of Plant Genomics**, Ohio, v. 2012, p. 1-15, 2012.

LAGARRIGUE, S.; PITEL, F.; CARRE, W.; ABASHT, B.; LE ROY, P.; NEAU, A.; AMIGUES, Y.; SOURDIOUX, M.; SIMON, J.; COGBURN, L.; AGGREY, S.; LECLERCQ, B.; VIGNAL, A.; DOUAIRE, M. Mapping quantitative trait loci affecting fatness and breast muscle weight in meat-type chicken lines divergently selected on abdominal fatness. **Genetics Selection Evolution**, Paris, n. 38, p. 85–97, 2006.

LANGMEAD, B.; SALZBERG, S.L. Fast gapped-read alignment with Bowtie 2. **Nature methods**, London, v. 4, n. 6, p. 357–359, 2012.

LECLERCQ, B. Genetic selection of meat-type chickens for high or low abdominal fat content. **Leanness in Domestic Birds, Genetic, Metabolic and Hormonal Aspects**, London, p. 25–40, 1988.

LEDUR, M.C.; PEIXOTO, J.D.O.; NONES, K.; COUTINHO, L.L. **Applied genomics: The Brazilian experience**. Disponível em: <
<http://www.alice.cnptia.embrapa.br/bitstream/doc/938031/1/final6890.pdf>>.
Acesso em: 22 nov. 2012.

LEENSTRA, F.R.; PIT, R. Fat deposition in a broiler sire line: heritability of and genetic correlations among body weight, abdominal fat, and feed conversion. **Poultry Science**, Savoy, n. 67, p. 1-9, 1988.

LERNER, H.; FLEISCHER, R. Prospects for the Use of Next-Generation Sequencing Methods in Ornithology. **The Auk**, Manitoba, v. 1, n. 127, p. 4–15, 2010.

LI, D.; KANG, Q.; WANG, D.-M. Constitutive coactivator of peroxisome proliferator-activated receptor (PPAR-gamma), a novel coactivator of PPAR-gamma that promotes adipogenesis. **Molecular Endocrinology**, Amsterdam, v. 21, p. 2320-2333, 2007.

LI, H.; RUAN, J.; DURBIN, R. 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. **Genome Research**, New York, n. 18, p. 1851–1858, 2008.

LI, H.; DURBIN, R. Fast and accurate short read alignment with Burrows-Wheeler transform. **Bioinformatics**, Oxford, v. 14, n. 25, p. 1754–1760, 2009.

LIN, H.; ZHANG, Z.; ZHANG, M.Q.; MA, B.; LI, M. ZOOM! Zillions of oligos mapped. **Bioinformatics**, Oxford, n. 24, p. 2431–2437, 2008.

LIU, D.; NAGAN, N.; JUST, W.W.; RODEMER, C.; THAI, T.P.; ZOELLER, R. A. Role of dihydroxyacetonephosphate acyltransferase in the biosynthesis of plasmalogens and nonether glycerolipids. **Journal of Lipid Research**, Bathesda, v. 4, n. 46, p. 727–735, 2005.

MADEIRA, L.A.; SARTORI, J.R.; ARAUJO, P.C.; CACHONI, C.; SALGADO, É.; BRAGA, P.; PEZZATO, A.C. Avaliação do desempenho e do rendimento de carcaça de quatro linhagens de frangos de corte em dois sistemas de criação. **Revista Brasileira de Zootecnia**, Viçosa, p. 2214–2221, 2010.

MARKLJUNG, E.; JIANG, L.; JAFFE, J.D.; MIKKELSEN, T.S.; WALLERMAN, O.; LARHAMMAR, M.; ZHANG, X.; WANG, L.; SAENZ-VASH, V.; GNIRKE, A.; et al. ZBED6, a novel transcription factor derived from a domesticated DNA transposon regulates IGF2 expression and muscle growth. **PLoS Biology**, San Francisco, v. 7, n. 12, p. 1-13, 2009.

MARKLUND, S.; CARLBORG, O. SNP detection and prediction of variability between chicken lines using genome resequencing of DNA pools. **BMC Genomics**, London, v. 1, n. 11, p. 665, 2010.

MATOS, R.; COUTINHO, C.M.; THULER, L.C.S.; FONSECA, F.P.; SOARES, F.; SILVA, E.; GIMBA, E.R. Expression analysis of thrombospondin 2 in prostate cancer and benign prostatic hyperplasia. **Experimental and Molecular Pathology**, San Diego, v. 3, n. 94, p. 438–444, 2013.

MATTE, C.H.F. **Indels autossômicos e do cromossomo X em uma amostra da população do Rio Grande do Sul para possíveis aplicações forenses**. 2011. 26p. Dissertação (Mestrado em Biologia Celular e Molecular) - Universidade Católica do Rio Grande do Sul, Porto Alegre, 2011.

MCELROY, J.P.; KIM, J.J.; HARRY, D.E.; BROWN, S.R.; DEKKERS, J.C. M.; LAMONT, S.J. Identification of trait loci affecting white meat percentage and other growth and carcass traits in commercial broiler chickens. **Poultry Science**, Savoy, v. 4, n. 85, p. 593–605, 2006.

MCKENNA, A.; HANNA, M.; BANKS, E.; SIVACHENKO, A.; CIBULSKIS, K.; KERNYTSKY, A.; GARIMELLA, K.; ALTSHULER, D.; GABRIEL, S.; DALY, M.; DEPRISTO, M. A. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. **Genome Research**, New York, n. 20, p. 1297–1303, 2010.

MCLAREN, W.; PRITCHARD, B.; RIOS, D.; CHEN, Y.; FLICEK, P.; CUNNINGHAM, F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. **Bioinformatics**, Oxford, v. 16, n. 26, p. 2069–2070, 2010.

MICHELAN FILHO, T. **Seleção para diminuição do conteúdo de gordura em frangos**. Concórdia: EMBRAPACNPISA, 1986. p. 19.

MILLS, R.E.; WALTER, K.; STEWART, C.; HANDSAKER, R.E.; CHEN, K.; ALKAN, C.; ABYZOV, A.; YOON, S.C.; YE, K.; CHEETHAM, R.K.; [...]; KORBEL, G.O. Mapping copy number variation by population-scale genome sequencing. **Nature**, London, n. 470, p. 59–65, 2011.

MIZUNO, Y.; NINOMIYA, Y.; NAKACHI, Y.; ISEKI, M.; IWASA, H.; AKITA, M.; TSUKUI, T.; [...]; OKAZAKI, Y. Tysnd1 deficiency in mice interferes with the peroxisomal localization of PTS2 enzymes, causing lipid metabolic abnormalities and male infertility. **PLOS Genetics**, San Francisco, v. 2, n. 9, 2013.

MORAES, V.G. de; CAPANEMA L. A genética de frangos e suínos – a importância estratégica de seu desenvolvimento para o Brasil. **Agroindústria BNDS**, Rio de Janeiro, v. 35, p. 119 –154, 2012.

MU, J.C.; JIANG, H.; KIANI, A.; MOHIYUDDIN, M.; BANI ASADI, N.; WONG, W. H. Fast and accurate read alignment for resequencing. **Bioinformatics**, Oxford, v. 18, n. 28, p. 2366–2373, 2012.

MURAKAMI, K.T.; PINTO, M.F.; HELENA, E.; PONSANO, G.; NETO, G. Desempenho produtivo e qualidade da carne de frangos alimentados com ração contendo óleo de linhaça. **Pesquisa Agropecuária Brasileira**, Brasília, v. 1, p. 401–407, 2010.

NADAF, J.; PITEL, F.; GILBERT, H. DUCLOS, M.J; VIGNOLES, F.; BEAUMONT, C.; VIGNAL, A.; PORTER, T.E.; COGBURN, L.A.; AGGREY, S.E; SIMON, J.; LE BIHAN-DUVAL E. QTL for several metabolic traits map to loci controlling growth and body composition in an F2 intercross between high- and low-growth chicken lines. **Physiological Genomics**, Bathesda, n. 38, p. 241–249, 2009.

NASSAR, M.K.; GORAGA, Z.S.; BROCKMANN, G. Quantitative trait loci segregating in crosses between New Hampshire and White Leghorn chicken lines: III. Fat deposition and intramuscular fat content. **Animal Genetics**, New York, v. 1, n. 44, p. 62–68, 2012.

NEPA – UNICAMP. **Tabela brasileira de composição de alimentos**. 4 ed. Campinas, 2011. 161p.

NEUMAN, J.; ISAKOV, O.; SHOMRON, N. Analysis of insertion-deletion from deep-sequencing data: software evaluation for optimal detection. **Briefings in Bioinformatics**, London, v. 1, n. 14, p. 46–55, 2013.

NG, P.C. SIFT: predicting amino acid changes that affect protein function. **Nucleic Acids Research**, London, v. 13, n. 31, p. 3812–3814, 2003.

NIE, H.; CROOIJMANS, R.P.M.A.; LAMMERS, A.; VAN SCHOTHORST, E. M.; KEIJER, J.; NEERINCX, P.B.T.; LEUNISSEN, J.A.M.; MEGENS, H-J.; GROENEN, M.A.M. Gene expression in chicken reveals correlation with structural genomic features and conserved patterns of transcription in the terrestrial vertebrates. **PloS one**, San Francisco, v. 5, n. 8, p. 1-7, 2010.

NIE, Q.; FANG, M.; XIE, L.; PENG, X.; XU, H.; LUO, C.; ZHANG, D.; ZHANG, X. Molecular characterization of the ghrelin and ghrelin receptor genes and effects on fat deposition in chicken and duck. **Journal of Biomedicine and Biotechnology**, New York, v. 2009, p. 1-12, 2009.

NINOV, K.; LEDUR, M.C.; NONES, K.; CAETANO, A.R.; COLDEBELLA, A.; BERTOL, T.M.; COUTINHO, L.L. **Mining of polymorphisms in the leptin receptor gene in two chicken lines and their association with performance and carcass traits**. In: INTERNATIONAL CONFERENCE ON ANIMAL GENETICS, 30., 2006. Porto Seguro. **Proceedings** Porto Seguro, BA, 2006. 1 CD-ROM

NOLAN, C.M.; KILLIAN, J.K.; PETITTE, J.N.; JIRTLE, R.L. Imprint status of M6P/IGF2R and IGF2 in chickens. **Development Genes and Evolution**, Berlin, v. 4, n. 211, p. 179–183, 2001.

NONES, K.; LEDUR, M.C.; ZANELLA, E.L.; KLEIN, C.; PINTO, L.F.B.; MOURA, A.S.A.M.T.; COUTINHO, L.L. Quantitative trait loci associated with chemical composition of the chicken carcass, **Animal Genetics**, New York, v. 43, p. 570–576, 2012.

OGURO, R.; KAMIDE, K.; KATSUYA, T.; AKASAKA, H.; SUGIMOTO, K.; CONGRAINS, A.; ARAI, Y.; HIROSE, N.; SAITOH, S.; OHISHI, M.; MIURA, T.; RAKUGI, H. A single nucleotide polymorphism of the adenosine deaminase, RNA-specific gene is associated with the serum triglyceride level, abdominal circumference, and serum adiponectin concentration. **Experimental Gerontology**, Oxford, v. 2, n. 47, p. 183–187, 2012.

OGUZ, I.; ALTAN, O.; KIRKPINAR, F.; SETTAR, P. Body weights, carcass characteristics, organ weights, abdominal fat, and lipid content of liver and carcass in two lines of Japanese quail (*Coturnix coturnix japonica*), unselected and selected for four week body weight. **British Poultry Science**, London, v. 37, p. 579–588, 1996.

PABINGER, S.; DANDER, A.; FISCHER, M.; SNAJDER, R.; SPERK, M.; EFREMOVA, M.; KRABICHLER, B.; ET AL. A survey of tools for variant analysis of next-generation genome sequencing data. **Briefings in Bioinformatics**, London, p 1-23, 2013.

PARK, H.-B.; JACOBSSON, L.; WAHLBERG, P.; SIEGEL, P.B.; ANDERSSON, L. QTL analysis of body composition and metabolic traits in an intercross between chicken lines divergently selected for growth. **Physiological Genomics**, Bathesda, v. 2, n. 26, p. 216–223, 2006.

Patthy, L. Exons and introns. **Current Opinion in Structural Biology**, Philadelphia, v. 4, p. 383–392, 1994.

PENA, S.D.J. Segurança Pública: Determinação de identidade genética pelo DNA. In: SEMINÁRIOS TEMÁTICOS PARA A 3ª CONFERÊNCIA NACIONAL DE C, T & I, 2005. Brasília, DF **Proceedings...** 2005. Brasília:Parcerias Estratégicas p .13

PLOTKIN, J.B.; KUDLA, G. Synonymous but not the same: the causes and consequences of codon bias. **Nature Reviews**, London, v. 1. n. 12, p. 32–42, 2011.

RAO, Y.S.; WANG, Z.F.; CHAI, X.W.; NIE, Q.H.; ZHANG, X.Q. Relationship between 5' UTR length and gene expression pattern in chicken. **Genetics**, Bathesda, v. 141, p. 311–318, 2013.

RAO, Y. S.; WANG, Z.F.; CHAI, X.W.; WU, G.Z.; NIE, Q.H.; ZHANG, X.Q. Indel segregating within introns in the chicken genome are positively correlated with the recombination rates. **Hereditas**, Lund, v.2, n. 147, p. 53–57, 2010.

RESNYK, C.W.; CARRÉ, W.; WANG, X.; PORTER, T.E.; SIMON, J.; LE BIHAN-DUVAL, E.; DUCLOS, M.J.; AGGREY, S.E.; COGBURN, L.A. Transcriptional analysis of abdominal fat in genetically fat and lean chickens reveals adipokines, lipogenic genes and a link between hemostasis and leanness. **BMC genomics**, London, v. 14, p. 1-26, 2013.

RODEMER, C.; THAI, T.P.; BRUGGER, B.; KAERCHER, T.; WERNER, H.; NAVE, K.A.; WIELAND, F.; GORGAS, K.; JUST, W. W. Inactivation of ether lipid biosynthesis causes male infertility, defects in eye development and optic nerve hypoplasia in mice. **Human Molecular Genetics**, Oxford, v. 12, p. 1881-1895, 2003.

ROSÁRIO, M.F. do; LEDUR, M.C.; MOURA, A.S.A.M.T.; COUTINHO, L.L.; GARCIA, A.A.F. Genotypic characterization of microsatellite markers in broiler and layer selected chicken lines and their reciprocal F1s. **Scientia Agricola**, Piracicaba, v. 2, n. 66, p. 150–158, 2009.

ROSMOND, R.; CHAGNON, M.; BOUCHARD, C.; BJORNTORP, P. G-308A polymorphism of the tumor necrosis factor alpha gene promoter and salivary cortisol secretion. **Journal of Clinical Endocrinology & Metabolism**, Baltimore, v. 86, p. 2178-2180, 2001.

RUBIN, C.J.; ZODY, M.C.; ERIKSSON, J.; MEADOWS, J.R.S.; SHERWOOD, E.; WEBSTER, M.T.; JIANG, L.; INGMAN, M.; SHARPE, T.; KA, S.; HALLBOOK, F.; BESNIER, F.; CARLBORG, O.; BED'HOM, B.; TIXIER-BOICHARD, M.; JENSEN, P.; SIEGEL, P.; LINDBLAD-TOH, K.; ANDERSSON, L. Whole-genome resequencing reveals loci under selection during chicken domestication. **Nature**, London, v. 464, n. 7288, p. 587–591, 2010.

RUY, D.C. **Mapeamento de QTL para desempenho e características de carcaça, nos cromossomos 3 e 5 de *Gallus gallus***. 2004. 122p. Tese (Doutorado em Ciência Animal e Pastagens) – Universidade de São Paulo, Piracicaba, 2004.

ORDINE, A.G.; ALMEIDA, A.; DIAS, Z. ALGAe: A test-bench environment for a gene-tic algorithm-based multiple sequence aligner. In: BRAZILIAN SYMPOSIUM ON BIOINFORMATICS, 2011. Brasília, DF **Digital Proceedings...** Brasília, DF, 2011. p 57-60

SANTOS, A.L.; SAKOMURA, N.K.; FREITAS, E.R.; FORTES, C.M.L.S.; CARRILHO, E.N.V.M.; FERNANDES, J.B.K. Estudo do Crescimento , Desempenho , Rendimento de Carcaça e Qualidade de Carne de Três Linhagens de Frango de Corte. **Revista Brasileira de Zootecnia**, Viçosa, p.1589–1598, 2005.

SANTOS, C.R.; SCHULZE, A. Lipid metabolism in cancer. **The FEBS journal**, Oxford, v. 15, n. 279, p. 2610–2623, 2012.

SAUNA, Z.E.; KIMCHI-SARFATY, C. Understanding the contribution of synonymous mutations to human disease. **Nature Reviews**, London, v. 12, n. 10, p. 683–691, 2011.

SCHATZ, M. Cloudburst: highly sensitive read mapping with mapreduce. **Bioinformatics**, Oxford, v. 25, p. 1363–1369, 2009.

SCHUTZ, K.; KERJE, S.; CARLBORG, O.; JACOBSSON, L.; ANDERSSON, L. JENSEN, P. QTL analysis of a red jungle fowl × White Leghorn intercross reveals trade-off in resource allocation between behavior and production traits. **Behavior Genetics**, New York, v. 32, p. 423–433, 2002.

SEIT-NEBI, A.; CHENG, W.; XU, H.; HAN, J. MLK4 has negative effect on TLR4 signaling. **Cellular & Molecular Immunology**, Philadelphia, v. 9, p. 27–33, 2012.

SHITAYE, H.S; TERKHORN, S.P; COMBS, J.A; HANKENSON, K.D. Thrombospondin-2 is an endogenous adipocyte inhibitor. **Matrix Biology**, Stuttgart, v. 29, p. 549–556, 2010.

SIEGEL, P.B.; DODGSON, J.B.; ANDERSSON, L. Progress from chicken genetics to the chicken genome. **Poultry Science**, Savoy, v. 85, n. 12, p. 2050–2060, 2006.

SILVA, F.E.D.J.; GUIDO, L.N.; GAZAFFI, R.; GARCIA, A.A.F.; LEDUR, M. C.; COUTINHO, L.L.; ROSÁRIO, M.F. do. Regiões genômicas associadas a características de desempenho e carcaça no cromossomo 5 de linhagens Brasileiras de galinha. **Pesquisa Agropecuária Brasileira**, Brasília, v. 46, n. 3, p.229–238, 2011.

SILVA, V.H.; PANDOLFI, J.R.C.; GODOY, T.F.; PEIXOTO, J.O.; TESSMANN, A.L.; LEDUR, M.C. A SNP in the hepatocyte nuclear factor 4 gene associated with carcass traits in males from a paternal broiler line. In: WORLD'S POULTRY CONGRESS, 24., 2012. Salvador, BA **Proceedings...** Salvador, BA, 2012. P 1-9.

SMITH, A.D.; XUAN, Z.; ZHANG, M. Using quality scores and longer reads improves accuracy of Solexa read mapping. **BMC Bioinformatics**, London, v. 9, p.1-8, 2008.

SOUZA, C.A. **Investigação de polimorfismos nos genes dos fatores miogênicos e miostatina como marcadores moleculares para características quantitativas em *Gallus gallus***. 2004. 108p. Dissertação (Mestrado em Ciência Animal e Pastagens) – Escola Superior de Agricultura “Luiz de Queiróz”, Universidade de São Paulo, Piracicaba, 2004.

SUN, Y.; ZHAO, G.; LIU, R.; ZHENG, M.; HU, Y.; WU, D.; ZHANG, L.; et al. The identification of 14 new genes for meat quality traits in chicken using a genome-wide association study. **BMC genomics**, London, v. 14, n. 1, p. 1-11, 2013.

TATSUDA, K.; FUJINAKA, K. Genetic mapping of the QTL affecting body weight in chickens using a F-2 family. **British Poultry Science**, London, v.42, p. 333–337, 2001.

TEIGLER, A.; KOMLJENOVIC, D.; DRAGUHN, A.; GORGAS, K.; JUST, W. W. Defects in myelination, paranode organization and Purkinje cell innervation in the ether lipid-deficient mouse cerebellum. **Human Molecular Genetics**, Oxford, v.18, p. 1897-1908, 2009.

TIAN, D.; WANG, Q.; ZHANG, P.; ARAKI, H.; YANG, S.; KREITMAN, M.; NAGYLAKI, T.; HUDSON, R.; BERGELSON, A.; QUN, J.C. Single-nucleotide mutation rate increases close to insertions/deletions in eukaryotes. **Nature**, London, n. 455, p. 105–108, 2008.

UBABEF - União Brasileira de Avicultura. **Relatório anual 2012**. Disponível em: <<http://www.ubabef.com.br/publicacoes?m=75&date=2012-04>>. Acesso em: 20 Abril 2012.

USA - United States Department of Agriculture. 2013. **Indian Beef Exports Surge To Continue in 2013 World's Leading Beef Exporters Shipments to Asia Fuel Indian Beef Exports**. Disponível em: <<http://www.usda.gov/wps/portal/usda/usdahome>>. Acesso em: 20 Abril 2012.

USDA – Food composition table. **National Nutrient Database for Standard Reference, Release**. 2013. Disponível em: <<http://ndb.nal.usda.gov/ndb/search/list?format=&sort=&fg=Poultry+Products&offset=50>>. Acesso em: 21 Abril 2013.

WANG, K.; LI, M.; HAKONARSON, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. **Nucleic Acids Research**, London, v. 38, n. 16, p. 1-7, 2004.

WANG, M.C.; BOHMANN, D.; JASPER, H. JNK extends life span and limits growth by antagonizing cellular and organism-wide responses to insulin signaling. **Cell**, Cambridge, v. 121, n.1, p. 115–125, 2005.

WANG, S.Z.; HU, X.X.; WANG, Z.P.; LI, X.C.; WANG, Q.G.; WANG, Y.X. Quantitative trait loci associated with body weight and abdominal fat traits on chicken. **Genetics and Molecular Research**, Ribeirão Preto, v.11, n. 2, p. 956–965, 2012.

WANG, Y.; LI, H.; GU, Z.L.; ZHAO, J.G.; WANG, Q.G.; WANG, Y. X. Correlation analysis between single nucleotide polymorphism of the leptin receptor intron 8 and fatness traits in chickens. **Yi Chuan Xiu Bao**, Chaoyang, v. 31, n. 3, p. 265-269, 2004.

WATERS, S.M.; MCCABE, M.S.; HOWARD, D.J.; GIBLIN, L.; MAGEE, D.; MACHUGH, D.E.; BERRY, D.P. Associations between newly discovered polymorphisms in the *Bos taurus* growth hormone receptor gene and performance traits in Holstein-Friesian dairy cattle. **Animal Genetics**, New York, v. 42, n. 1, p. 39–49, 2010.

WEBER, J.L.; DAVID, D.; HEIL, J.; FAN, Y.; ZHAO, C.; MARTH, G. Human diallelic insertion/deletion polymorphisms. **American Journal of Human Genetics**, Chicago, v. 71, p. 854–862, 2002.

WEIVODA, M.M.; HOHL, R.J. Geranylgeranyl pyrophosphate stimulates PPAR γ expression and adipogenesis through the inhibition of osteoblasts differentiation. **Bone**, New York, v. 50, n. 2, p. 467–476, 2011.

WONG, G.K.S.; LIU, B.; WANG, J.; ZHANG, Y.; YANG, X.; ZHANG, Z.; MENG, Q.; [...]; YANG, H. A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. **Nature**, London, v. 432, n. 7018, p. 717–722, 2004.

WRIGHT, D.; BOIJE, H.; MEADOWS, J.R.S.; BED'HOM, B.; GOURICHON, D.; VIEAUD, A.; TIXIER-BOICHARD, M.; RUBIN, C.J.; IMSLAND, F.; HALLBOOK, F. et al. Copy number variation in intron 1 of *SOX5* causes the Pea-comb phenotype in chickens. **PLoS Genetics**, San Francisco, v. 5, n. 6, p. 1–10, 2009.

WU, G. Q.; DENG, X. M.; LI, J. Y.; LI, N.; YANG, N. A potential molecular marker for selection against abdominal fatness in chickens. **Poultry Science**, Savoy, v. 85, p. 1896–1899, 2006.

XU, Z.; NIE, Q.; ZHANG, X. Overview of Genomic Insights into Chicken Growth Traits Based on Genome-Wide Association Study and microRNA Regulation. **Current Genomics**, Hilversum, v. 14, n. 2, p. 137–146, 2013.

ZEREHDARAN, S.; VEREIJKEN, A.L.; VAN ARENDONK, J.A.; VAN DER WAAIJT, E.H. Estimation of genetic parameters for fat deposition and carcass traits in broilers. **Poultry Science**, Savoy, v. 4, p. 521–525, 2004.

ZHANG, S.; LIU, S.; YU, N.; XIANG, L. RNA released from necrotic keratinocytes upregulates intercellular adhesion molecule-1 expression in melanocytes. **Archives of Dermatological Research**, Heidelberg, v. 303, n. 10, p. 771–776, 2011.

ZHBANNIKOV, I.Y.; HUNTER, S.S.; SETTLES, M.L. **SeqyClean User Manual**. Disponível em: < <http://sourceforge.net/projects/seqclean/files/>>. Acesso em: 03 Jun. 2013.

ZHOU, H.; EVOCK-CLOVER, C.M.; MCMURTRY, J.P.; ASHWELL, C.M.; LAMONT, S.J. Genome-wide linkage analysis to identify chromosomal regions affecting phenotypic traits in the chicken. **Poultry Science**, Savoy, v. 86, n. 2, p. 267-276, 2007.

APÊNDICE

APÊNDICE A – Protocolo de extração de DNA (proteínase k)

MÉTODO RÁPIDO DE EXTRAÇÃO E PURIFICAÇÃO DE DNA EXTRAÇÃO COM SAL

Protocolo modificado para extração de DNA de sangue de galinha

1. Misturar 50µL de sangue total colhido em EDTA com 1mL de tampão de lise de células vermelhas;
2. Centrifugar por 1 minuto a 13000rpm. Descartar o sobrenadante. Repetir o passo 1 e 2 (repetimos por 4 vezes).
3. Lavar o pellet uma vez com 0,5mL de água destilada (13000 rpm, 1min);
4. Ressuspender o pellet em 80 µL do tampão de proteinase K (5x), 7 µL de proteinase K (20mg/ml), 10 µL de SDS 20% e 282 µL de água destilada.
5. Incubar a 55°C por 1 hora;
6. Adicionar 120 µL de NaCl 5M e agitar no vortex por 15 segundos. Centrifugar a 13000 rpm por 5min.
7. Transferir o sobrenadante para um novo tubo e adicionar 1 mL de etanol absoluto. Homogeneizar e centrifugar por 5 minutos a 13000 rpm.
8. Descartar o sobrenadante, secar o pellet, ressuspender em 100 µL H₂O UP e dissolver passando no vortex por 30 segundos.

Tampão de Lise

0,32M sucrose

12mM Tris-HCl pH7,5

5mM MgCl₂

1% Triton X

Tampão de Proteinase K (5x)

0,375M NaCl

0,12M EDTA (pH 8.0)