UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS JABOTICABAL

EFEITO DA ENDOGAMIA NA SELEÇÃO GENÔMICA EM POPULAÇÕES SIMULADAS DE AVES POEDEIRAS

Guilherme Batista do Nascimento

Biólogo

2014

UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS JABOTICABAL

EFEITO DA ENDOGAMIA NA SELEÇÃO GENÔMICA EM POPULAÇÕES SIMULADAS DE AVES POEDEIRAS

Guilherme Batista do Nascimento

Orientador: Prof. Dr. Danísio Prado Munari

Coorientador: Dra. Mônica Correa Ledur

Dr. Ricardo Vieira Ventura

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias - Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Genética e Melhoramento Animal.

2014

Nascimento, Guilherme Batista do
N244e Efeito da endogamia na seleção genômica em populações simuladas de aves poedeiras / Guilherme Batista do Nascimento. – – Jaboticabal, 2014
viii, 42 p. : il. ; 28 cm

Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, 2014
Orientadora: Danísio Prado Munari
Co-orientadores: Mônica Correa Ledur, Ricardo Vieira Ventura
Banca examinadora: Ricardo da Fonseca, Mauricio de Alvarenga Mudado Mudadu
Bibliografia

1. Acurácia. 2. Desequilíbrio de ligação. 3. Herdabilidade. 4.Valor genético genômico. I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 575.17:636.5

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação – Serviço Técnico de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.

DADOS CURRICULARES DO AUTOR

Guilherme Batista do Nascimento – Filho de Gildo Divino Soares do Nascimento e Deise Mara Batista, nasceu em Tupã – SP, no dia 12 de outubro de 1988. Iniciou a graduação em Ciências Biológicas em março de 2008, na Faculdade de Ciências Agrárias e Veterinárias – UNESP - Câmpus de Jaboticabal e obteve o Título de Bacharel em Ciências Biológicas em 2012. Durante os anos de 2010 e 2011, desenvolveu o trabalho de Iniciação Científica na mesma instituição de ensino, sob orientação do Prof. Dr. Danísio Prado Munari e Dr. Rodrigo Pelicioni Savegnago, o qual lhe rendeu o prêmio Ciências Biológicas, pelo melhor trabalho de conclusão de curso. Em março de 2012, ingressou no Curso de Mestrado do Programa de Genética e Melhoramento Animal na Faculdade de Ciências Agrárias e Veterinárias – UNESP - Câmpus de Jaboticabal, sob a orientação do Prof. Dr. Danísio Prado Munari e co-orientação da Pesquisadora Dra. Mônica Corrêa Ledur e Prof. Dr. Ricardo Vieira Ventura.

“O sábio não se exibe, e por isso brilha.
Ele não se faz notar, e por isso é notado.
Ele não se elogia, e por isso tem mérito.
E, porque não está competindo, ninguém no mundo
pode competir com ele”.

(LaoTsé - Tao Te Ching)

AGRADECIMENTOS

Ao meu orientador prof. Dr. Danísio Prado Munari, pelo profissionalismo, dedicação para com este trabalho e seu exemplo como profissional.

Ao amigo Dr. Rodrigo Pelicioni Savegnago, profissional exemplar que muito contribuiu para realização deste trabalho e da minha formação acadêmica.

Aos meus coorientadores Pesq. Drª. Mônica Corrêa Ledur e Dr. Ricardo Vieira Ventura, pelas valiosas contribuições e ensinamentos necessários para chegar a conclusão deste trabalho.

Ao Dr. Roberto Carvalheiro, pelas inúmeras ideias e contribuições nas análises.

Ao Dr. Leonardo de Oliveira Seno, por dispor de toda sua capacidade, atenção e amizade em cada valiosa opinião.

Ao Programa de Pós-Graduação em Genética e Melhoramento Animal da Faculdade de Ciências Agrárias e Veterinárias "Júlio de Mesquita Filho" – UNESP Câmpus de Jaboticabal, CAPES e FAPESP (processo 2012/23384-6), por acreditarem na minha capacidade em realizar esse projeto de pesquisa.

A todos os colegas e funcionários do departamento de Ciências Exatas, por ajudas profissionais e conversas, que tornaram a convivência uma agradável experiência a cada café compartilhado.

Aos queridos amigos da República "Kilombo" e seus agregados, que nunca faltaram quando os dias ficavam mais longos e uma boa amizade se fazia necessária. Vocês são a família que me foi permitido escolher.

Ao amigo e irmão Pedro, pelo companheirismo no dia-a-dia, pelos Beatles, pelo Floyd e pelo Clif!

A Renata, por ser o porto seguro e a palavra de otimismo quando mais precisei. Seu companheirismo e seu sorriso sempre me fortaleceram.

As amigas Ale e Marina, a cada encontro uma risada, uma cerveja um aprendizado diferente. O inesquecível está mesmo na intensidade e não no tempo.

A meu Pai, Gildo. Seu estimulo para eu ser um curioso é peça fundamental em cada dia do meu trabalho, seu amor é fundamental para que eu acredite que todo dia vai ser um dia melhor. A minha mãe, Deise. Sua força me faz acreditar que tudo é

possível, seu amor me faz sentir a pessoa mais importante do mundo. A minha irmã Nali. Por acreditar e me amar de uma forma como só quem tem um irmão pode saber. A minha avó Ivone, sempre doce e companheira e a meu avô Celso (*In memoriam*), onde estiver estará sempre comigo. A vocês dedico esse e todos os trabalhos de minha vida.

Aos queridos bizarros da Bio 08, todas as conquistas têm um pouco de vocês, que permaneçamos sempre unidos nessa amizade.

A Deus, pela experiência diária da fé, onde sou capaz de acreditar no que ainda não posso ver.

SUMÁRIO

Página

RESUMO.....II

ABSTRACT- IV

CAPÍTULO 1 - CONSIDERAÇÕES GERAIS.....1

1 INTRODUÇÃO 1

3.1 SOFTWARE DE SIMULAÇÃO 3

3.2 POLIMORFISMO DE NUCLEOTÍDEO ÚNICO – SNP 4

3.3 DESEQUILÍBRIO DE LIGAÇÃO.....4

3.4 ENDOGAMIA 6

3.5 PREDIÇÃO GENÔMICA.....8

4 REFERÊNCIAS.....10

CAPÍTULO 2 – ACURÁCIA DOS VALORES GENÉTICOS GENÔMICOS EM POPULAÇÕES COM DIFERENTES TAXAS DE ENDOGAMIA 16

1 INTRODUÇÃO 16

2 OBJETIVOS.....17

3 MATERIAL E MÉTODOS 17

3.1 Simulação das populações de aves 17
3.2 Estrutura da população histórica 18
3.3 Estrutura da população recente 18
3.4 Genoma 21
3.5. Tendência genética 23
3.6. População de treino e validação 24

4. RESULTADOS E DISCUSSÃO 27

4.1 Endogamia, homozigose e Ne 27
4.2 Tendência genética 30
4.4 Acurácia dos VGGP 36

5. CONCLUSÕES 38

6. REFERÊNCIAS.....38

EFEITO DA ENDOGAMIA NA SELEÇÃO GENÔMICA EM POPULAÇÕES SIMULADAS DE AVES POEDEIRAS

RESUMO–O objetivo do presente trabalho foi avaliar a acurácia de predição dos valores genéticos genômicos para características de diferentes herdabilidades, em populações simuladas de aves com diferentes níveis de endogamia. Os dados fenotípicos e genotípicos foram simulados com base na estrutura populacional de uma população experimental de aves poedeiras. Foram simulados os fenótipos e os genótipos de aves para características de taxa de postura total de ovos (PTO) e peso dos ovos as 32 semanas de idade (PO), com herdabilidades de 0,15 e 0,37 respectivamente. Foi simulada uma população histórica a fim de gerar desequilíbrio de ligação na população e esta deu origem as populações recentes em que foram simulados três cenários populacionais, visando maximizar (REC1), minimizar (REC2) e aleatorizar (REC3) os acasalamentos endogâmicos. Ao longo de 10 gerações recentes, os animais foram selecionados com base nos maiores valores genéticos preditos (VGP), utilizando o BLUP (best linear unbiased prediction) tradicional. O genoma das aves foi simulado ao longo dos 958 Mb do genoma *Gallus gallus* 4.0 com 3.747 QTL (loci de caracteres quantitativos) aleatoriamente distribuídos e 49.978 marcadores SNP uniformemente distribuídos. A fim de alterar as frequências alélicas e gerar variabilidade genética ao longo das gerações, foram simulados eventos de deriva genética, taxa de recombinação e mutação recorrente. Para avaliar os efeitos da endogamia nas populações recentes, foram calculados os desequilíbrios de ligação (DL), o tamanho efetivo da população (Ne) e as tendências genéticas em todos os cenários de populações recentes em ambas as características simuladas. Para predizer os valores genéticos genômicos preditos (VGGP), as populações recentes foram subdivididas em populações de treinamento e validação. Nas subpopulações de treinamento, os 960 animais com maiores valores de acurácia para os VGP foram genotipados, colhidas informações fenotípicas e tiveram seus valores genéticos preditos deregredidos (VGPd). As populações de validação foram apenas genotipadas e formadas pelos 1.120 animais da décima geraçõe recente. Com base nos VGPd dos animais da subpopulação de treinamento, o efeito dos marcadores foram obtidos por meio do modelo RR-BLUP (*random regression best linear unbiased prediction*). As maiores diferenças ao longo das 10 gerações recentes com relação ao coeficiente de endogamia e homozigose foram entre o cenário REC1 e os demais cenários. A média do coeficiente de endogamia variou de 0,06 ± 0,03 a 0,22 ± 0,12 nos cenários para PTO e 0,05 ± 0,03 a 0,20 ± 0,12 para PO. O cenário de populações recentes REC1 em PTO e PO, em comparação com REC2 e REC3, apresentou acréscimos mais acentuados no coeficiente de endogamia ao longo das gerações. Isso ocorreu pela seleção dentro de famílias causada pelo BLUP e pelo sistema de acasalamento adotado na simulação de REC1, gerando maior fixação de alelos, elevando as diferenças entre os valores de DL entre REC1 e os demais cenários. O incremento da endogamia elevou na magnitude do DL, mas alterou a redução do mesmo ao longo do distanciamento entre os marcadores adjacentes. Os maiores valores de acurácia

para PTO e PO foram 0,61 ± 0,11 e 0,73 ± 0,02, respectivamente, ambos em REC1. Porém, não houve diferenças significativas (p > 0,05) entre os valores de acurácia nas populações de validação. As acurácias dos VGGP obtidas nas populações de validação indicaram que diferentes taxa de endogamia e estimativas de herdabilidade não alteraram a capacidade dos modelos estatísticos em predizer o VGGP nesses cenários.

Palavras-chave: Acurácia, Desequilíbrio de ligação, Herdabilidade, Valor genético genômico

Inbreeding effect on genomic selection in simulated laying hens populations

ABSTRACT-The objective of this study was to evaluate the prediction accuracy of genomic breeding values for traits of different heritability in simulated populations with different inbreeding. Phenotypic and genotypic data were simulated based on the population structure of an experimental population of White Leghorn hens at Embrapa Suínos e Aves. The phenotypes and genotypes were simulated for the rate of total egg production (PTO) and egg weight to 32 weeks of age (PO) with heritability of 0.15 and 0.37, respectively. The historical population was simulated to generate linkage disequilibrium in the population. Three scenarios in recent populations were simulated for each trait: REC1, REC2 and REC3 to maximize inbreeding, minimize inbreeding and random mating, respectively. The animals were selected based on the largest breeding values along 10 generations. The genome of the birds was simulated with eight macro-chromosomes and 19 micro-chromosomes with 3.747 QTL randomly distributed and 49.978 SNPs markers evenly spaced along the 958 cM. Recombination, random drift and recurrent mutation were simulated in order to generate genetic variability. The linkage disequilibrium (LD), effective population (Ne) and genetic trends were calculated for all scenarios. Each recent population was divided in training and validation sets In order to predict the genomic breeding values. The training set included the genotypes and phenotypes of 960 animals, which had higher breeding values' accuracy. The validation set had 1120 animals of the last generation of the recent population. The average inbreeding ranged from 0.06 ± 0.30 to 0.22 ± 0.12 for PTO and 0.05 ± 0.03 to 0.20 ± 0.12 for PO. The REC1 populations had higher inbreeding along generations compared, both for PTO and PO, compared to REC2 and REC 3, and consequently higher level of LD. The highest accuracy for PTO and PO were 0.61 ± 0.11 and 0.73 ± 0.02, respectively, both in REC1. However, there were significant differences ($p>0.05$) between accuracies in validation sets among recent populations. The accuracies of the genomic breeding values of he validation populations indicated that inbreeding and heritability did not affected the ability of the statistical models to predict the genomic breeding values in all scenarios, considering the number of replicates used in this study.

Keywords: Accuracy, Linkage Disequilibrium, Heritability, Genomic breeding values

CAPÍTULO 1 - CONSIDERAÇÕES GERAIS

1 INTRODUÇÃO

O Brasil produz aves de postura atendendo totalmente à demanda do mercado interno, o que representou produção superior a 26 mil quilos de ovos em 2012 (UBABEF, 2012). Com uma população de 85,546 milhões de cabeças de aves de postura, o país é o 7º maior produtor mundial de ovos (ANUALPEC, 2013). Programas de melhoramento genético de aves poedeiras são realizados levando em conta características de interesse econômico que visam o aumento da produtividade, tais como a produção e o peso dos ovos.

Dentre os importantes fatores que precisam ser controlados nas populações sob seleção, o direcionamento dos acasalamentos garante o não aparecimento de problemas correlatos ao aumento da endogamia, uma vez que tornar os animais mais aparentados ao longo das gerações é passível de perda da variação genética na população (FRANKHAM et al., 2002). A inclusão de dados genômicos pode ser estratégica para o aumento na produção de ovos, entre os benefícios, informações mais precisas sobre a taxa de endogamia das populações de aves poedeiras.

A galinha foi o primeiro animal com seu genoma sequenciado (HILLIER et al., 2004), sendo identificados 2,8 milhões de polimorfismos de nucleotídeo único (WONG et al., 2004), tornando possível a realização de estudos de associação e seleção genômica em grande escala. Há dificuldades em se obter dados genômicos reais, o que torna a simulação de dados importante ferramenta para melhor compreender os efeitos da endogamia e da seleção genômica.

Ao predizer os parâmetros genéticos é possível conhecer as associações genéticas entre as características quantitativas, e quanto da variação fenotípica pode ser atribuída aos efeitos aditivos dos genes. Com essas informações é possível decidir sobre os melhores métodos de seleção e quais genótipos devem ser selecionados para atingir os objetivos de seleção (LEDUR et al., 1993). A predição dos parâmetros genéticos pode ser mais acurada com o uso de informações advindas de *chips* genômicos, que ao utilizar painéis de SNP (polimorfismo de nucleotídeo único) de alta densidade é capaz de predizer grande parte da variação

genética aditiva por meio da predição dos QTL (loci de características quantitativas) (HAYES et al., 2009). Os valores genéticos preditos com uso dos efeitos dos SNP são chamados de valores genéticos genômicos preditos (**VGGP**), enquanto que o valor genético predito (**VGP**) é obtido com informações fenotípicas dos animais e sua estrutura de *pedigree*. Meuwissen et al., (2001) foram os primeiros a usar simultaneamente todas as informações dos marcadores SNP como covariáveis na predição dos VGGP. Com o desenvolvimento *do chip* Illumina Chicken SNP60K (GROENEN et al., 2011), foi possível fazer estudos de associação genômica ampla e de seleção genômica, compreendendo melhor a segregação de alelos e o desequilíbrio de ligação (**DL**) em aves.

O uso da metodologia RR-BLUP (*random regression best linear unbiased prediction*), amplamente utilizada e acurada para predição dos VGGP (MEUWISSEN et al., 2001; BRITO et al., 2011), faz uso da matriz de parentesco estimada pelos marcadores SNP (matriz de parentesco genômica), mais acurada que a matriz de parentesco baseada no *pedigree*, usada no BLUP (best linear unbiased prediction) tradicional, pois informa a matriz de parentesco realizada e não uma matriz de parentesco médio associada ao *pedigree.* As informações geradas a partir da matriz de parentesco genômica resultam em maior precisão e confiabilidade no cálculo da taxa de endogamia da população em estudo. Diversas pesquisas relataram o uso de informações genômicas em animais de produção (ROOS et al., 2008; LUAN et al., 2009). No entanto, poucos estudos têm sido conduzidos em aves. Assim, a variação dos fatores que afetam a predição dos VGGP, tais como, herdabilidade, endogamia e DL deve ser melhor explorada.

2 OBJETIVOS

Objetivo geral

O objetivo do presente trabalho foi avaliar a acurácia de predição dos valores genéticos genômicos para características com diferentes herdabilidade, em populações simuladas de aves poedeiras com diferentes níveis de endogamia.

Objetivos específicos

(1) Avaliar a influência da taxa de endogamia na acurácia dos VGGP;
(2) Avaliar a influência da herdabilidade na capacidade do modelo de predição em predizer os efeitos dos marcadores;
(3) Avaliar a sensibilidade do DL sob elevadas taxas de endogamia.

3 REVISÃO DE LITERATURA

3.1 Software de simulação

A simulação de dados vem sendo utilizada para verificar o desempenho de modelos e detecção de QTL (CRUZ, 2006). Embora os custos de genotipagem tenham diminuído substancialmente, estudos de seleção genômica em larga escala ainda são de valor expressivo. Nas últimas décadas várias metodologias têm sido utilizadas neste segmento. Euclydes (1996) desenvolveu o sistema GENESYS, visando simular dados e avaliar metodologias de seleção, além das associações dos marcadores moleculares com características fenotípicas. Hoggart (2007) fez uso de ferramentas para simular dados genéticos de populações, propondo o *software* FREGENE. No entanto, grande parte dos *softwares* disponíveis publicamente não possuem as funcionalidades necessárias para estudos em programas de melhoramento de animais domésticos. Diferente dos outros *softwares,* o QMSim (SARGOLZAEI; SCHENKEL, 2009) foi projetado para simular análises baseadas em grandes sequências genômicas e complexas estruturas de pedigree, simulando assim populações próxima às reais. O programa considera características evolutivas

como o DL, mutação, gargalos genéticos, expansões populacionais e taxas de recombinação.

3.2 Polimorfismo de nucleotídeo único – SNP

Os SNP são um tipo de polimorfismo em que há alteração de um único nucleotídeo. De todos os marcadores genéticos, os mais abundantes são os SNP, uma vez que estão distribuídos homogeneamente ao longo de todo genoma, sendo responsáveis por 90% de toda a variação genética humana (LI et al., 2009). Os marcadores SNP em geral são bi-alélicos, ou seja, apenas duas variantes alélicas são encontradas em um mesmo *loci*. Os SNP podem ocorrer em regiões codificadoras ou regulatórias, porém, na maior parte das vezes são encontrados em espaços intergênicos, sem função determinada (SANTOS, 2011).

ORITA et al. (1989) relataram a existência de SNP em seu estudo, porém, restrições tecnológicas limitaram o uso deste tipo de variação à apenas regiões específicas do genoma. O desenvolvimento de metodologias para genotipagem de milhares de SNP simultaneamente, deu origem aos *chips* de SNP que permitiram maior acesso aos diferentes polimorfismos em animais de produção, tornando realidade o uso destes marcadores de forma ampla em modelos de predição (MEUWISSEN et al., 2001; SCHAEFFER, 2006; GODDARD; HAYES, 2007; RESENDE et al., 2008; DE LOS CAMPOS et al., 2009; HAYES, et al., 2009; SHEPHERD et al., 2010). Recentemente, Kraniset al. (2013) desenvolveram um *chip* com 600k marcadores SNP para a espécie *Gallus gallus*, que permitiu ampliar os estudos sobre o uso da informação genômica no processo de seleção em aves. Assim, estudos com marcadores SNPs podem contribuir para identificar alelos favoráveis para ganhos genéticos em programas de melhoramento genético.

3.3 Desequilíbrio de ligação

O desequilíbrio de ligação é definido como associação não-aleatória de alelos em dois ou mais loci, não necessariamente no mesmo cromossomo (FALCONER; MACKAY, 1996). Lander e Schork, (1994) relataram que eventos de migração, mutação, seleção, tamanho efetivo da população, entre outros eventos genéticos,

são causadores de variação no DL. Existem vários cálculos para se obter o DL, Hill e Robertson, (1968) propuseram a medida de desequilíbrio de ligação como sendo:

$$r^2 = \frac{D^2}{f(A)\,f(a)\,f(B)\,f(b)}$$

em que D é a f(AB)-f(A)f(B) e f(AB), f(A), f(a), f(B), f(b) são as frequências dos alelos marcadores. Esta medida do DL é a mais utilizada, por considerar a correlação entre os alelos de dois loci para indivíduos da população e foi demonstrado ser a medida mais adequada para estimar o DL entre marcadores bialélicos, como os SNP, por ser menos sensível à frequência alélica e ao tamanho da amostra (ZHAO et al., 2007; BOHMANOVA et al., 2010). Valores de r² próximos a zero indicam o estado de equilíbrio de ligação entre os marcadores, ou seja, há segregação independente entre os alelos de diferentes marcadores e valores de r² próximos de um indicam desequilíbrio de ligação total, em que há segregação em conjunto entre alelos de diferentes genes (RESENDE et al., 2008).

A extensão do DL decresce em função do distanciamento entre os marcadores e o aumento na taxa de recombinação entre os marcadores (BRITO, 2011). Apesar de esta tendência ser comumente observada, marcadores próximos não estão sempre em DL (ARDLIE et al, 2001) e, ao contrário, DL foi observado entre marcadores localizados a grandes distâncias um do outro (COLLINS et al., 1999; STEPHENS et al., 2001). Os estudos realizados para descrever a extensão e o padrão do DL em aves se tornam mais detalhados à medida que os *chips* de marcadores SNP foram se tornando mais densos. Além de sua utilidade no mapeamento de genes e na seleção, o DL pode também ser útil na inferência sobre a história reprodutiva de uma população (RESENDE, 2008).

Qanbari et al. (2010) encontraram diferenças entre o DL quando compararam aves da raça White Leghorn e da linhagem GmbH mantidas em sistemas comerciais e experimentais, relatando valores máximos de r² de 0,73 ± 0,36 para aves comerciais da raça White Leghorn e mínimo de 0,28 ± 0,30 para aves comerciais F2 GmbH em distâncias de 0,0-0,1 Mb. Sargolzaei et al., (2008) relataram em bovinos r^2 máximo de 0,58 ± 0,37 em distâncias de 0,0-0,1 Mb e mínimo de 0,08 ± 0,10 em distâncias de 0,9-1,0 Mb, utilizando *chip* de 10k marcadores SNPs. O DL pode se

estender por mais de 5-10 Mb de distância entre os marcadores (FARNIR et al., 2000), no entanto, há rápido decréscimo com aumento da distância física entre os marcadores. Qanbari et al., (2010) em populações comerciais de White Leghorn, registraram r^2 de 0,03 ± 0,06 a distâncias de 5-10 Mb.

Marcadores em desequilíbrio de ligação com os QTL, tanto com grandes ou pequenos efeitos, explicarão quase a totalidade da variação genética de um caráter quantitativo. Para isso, o número de SNPs deve ser suficientemente grande para que a probabilidade de se encontrar um QTL em desequilíbrio de ligação com ao menos um marcador seja muito alta. Este aspecto é importante uma vez que somente os marcadores em desequilíbrio de ligação com os QTL serão úteis na determinação dos fenótipos das características e na explicação da variação genética aditiva (RESENDE et al. 2008).

3.4 Endogamia

A endogamia de uma população é resultante do acasalamento entre indivíduos mais aparentados do que seriam se tivessem sido escolhidos ao acaso (VAN VLECK et al., 1987). O aumento na taxa de endogamia eleva a frequência de alelos em homozigose, diminuindo a taxa de heterozigotos na população. Muir (2000) relatou que o primeiro impacto da endogamia é a perda de alelos, decorrente da oscilação genética, que é diretamente proporcional aos níveis de endogamia. Quando a endogamia ultrapassa os 10%, há grande aumento no número de *loci* em homozigose, elevando a chance de expressão de genes recessivos indesejáveis (ALCALÁ et al., 1995). Altas taxas de endogamia podem provocar perda no ganho genético em animais sob seleção, reduzindo o valor fenotípico médio de características reprodutivas ou a eficiência fisiológica do animal (FALCONER; MACKAY, 1996).

O coeficiente de endogamia (F), proposto por Wright (1922), expressa a probabilidade de que, em dois indivíduos, dois alelos amostrados aleatoriamente sejam idênticos por descendência. Em populações com *pedigree* incompleto, é possível predizer a endogamia dos animais mesmo que sejam desconhecidos ambos os pais, no entanto, esta predição irá ser subestimada ou superestimada (LUTAAYA et al., 1999). Boichard et al. (1997), em bovinos de corte, observaram

que o coeficiente de endogamia subestimado influencia e superestima o tamanho efetivo da população.

A ausência de acasalamentos eficientes e o uso de rebanhos fechados causa elevação nos níveis de animais endogâmicos (SILVA et aI., 2001). No entanto, Oliveira et aI. (1999) ressaltam a dificuldade em reduzir os coeficientes de endogamia em pequenas populações quando a seleção é praticada para apenas uma característica de interesse econômico. A endogamia é importante parâmetro para ser monitorado e controlado em um programa de melhoramento genético (SORENSEN et al., 2005), uma vez que é considerada o principal fator genético que gera entrave, em curto prazo, a sobrevivência das espécies, enquanto que a deriva genética é considerada a principal causa de perda de variação genética em longo prazo.

A seleção dos animais tem gerado progresso genético das características de interesse econômico, mas a seleção focada em poucos animais geneticamente superiores pode levar ao aumento da probabilidade na ocorrência de animais endogâmicos, prejudicando o ganho genético em programas de melhoramento genético (VERRIER et al., 1993). Oliveira et al. (2006), em populações simuladas de frango de corte, verificaram a importância no controle da endogamia quando se utiliza os valores genéticos preditos pelo BLUP, já que o mesmo leva ao rápido incremento da endogamia. A falta de controle sobre os acasalamentos pode acarretar em maior seleção de indivíduos de uma mesma família, elevando a endogamia e subsequente diminuição a resposta a seleção (CARNEIRO et al., 2008).

Toro et aI. (1988), utilizando-se do BLUP tradicional em dados simulados e com diferentes esquemas de acasalamento, relataram que utilizar informações de *pedigree* para planejar esquemas de acasalamento, antes de se obter as avaliações genéticas, reduz as taxas de endogamia na população. O uso da matriz de parentesco genômica pode ser determinante no sucesso de programas de melhoramento genético e no controle da taxa de endogamia nas populações sob seleção, o que poderia ser conseguido com o maior uso dos marcadores genômicos em programas de melhoramento animal.

3.5 Predição genômica

A predição genômica consiste na obtenção dos valores genéticos dos animais com base na avaliação genética, utilizando marcadores SNP ao longo de todo o genoma para serem empregados na seleção (MEUWISSEN el al., 2001). Para poder predizer os efeitos dos genes responsáveis pelas alterações fenotípicas em uma determinada característica, a seleção genômica considera que a posições dos QTL não são conhecidas, sendo necessária grande quantidade de marcadores ao longo de todo o genoma para predizer a variação dos QTLs. Baseada no desequilíbrio de ligação entre os marcadores e QTL, essa metodologia tem a finalidade de explicar os efeitos do QTL sobre as características de interesse nos programas de melhoramento genético. O DL garante que haja associação entre SNP e QTL, mesmo que em longas distâncias no genoma e essa associação é útil a programas de melhoramento se for mantida ao longo das gerações, estando associada a toda população e dependente das características avaliadas. A densidade de SNPs necessária para um programa de melhoramento genético depende desta extensão do desequilíbrio de ligação que, por sua vez, depende do tamanho efetivo da população (BRITO, 2011).

O uso de ferramentas moleculares tem possibilitado grande avanço na compreensão do funcionamento do genoma de plantas e animais, possibilitando reais aplicações em diversas áreas, da medicina à produção animal. Na predição genômica todos os SNPs e QTL são considerados na análise, não sendo necessária a construção de mapa de ligação (LEDUR; PEIXOTO, 2010). Solberg et al. (2006) compararam o uso direto dos efeitos dos marcadores com o uso de haplótipos (blocos de DNA transmitidos em conjunto para progênies) e verificaram melhores valores de acurácia com a utilização apenas dos marcadores. Os valores genéticos genômico predito tem suas acurácias dependentes do: (1) do nível de DL entre os marcadores e os QTL; (2) número de animais com fenótipo e genótipo na população treinamento, na qual os efeitos dos SNPs serão estimados; (3) herdabilidade da característica e da acurácia dos valores genéticos preditos (VGP) no treino e validação; (4) distribuição dos efeitos dos QTL (HAYES et al., 2009), (5) tamanho do genoma e (6) tamanho efetivo da população (Ne). Por meio de dados simulados, Meuwissen et al. (2001) demonstraram que é possível calcular valores genéticos

genômicos com alta acurácia a partir dos efeitos preditos pelos marcadores, para isso os marcadores devem estar em DL com os QTL. Os autores simularam uma população com r^2= 0,20 e encontraram níveis de acurácia dos VGGP próximos a 0,85. Calus et al. (2008) encontraram acurácia de 0,68 e 0,82 para os VGGP quando r^2 foi igual a 0,1, e 0,2, respectivamente, indicando a alta relação das taxas de DL com a acurácia.

Para predizer os efeitos dos marcadores SNP deve-se primeiramente formar um conjunto de animais com informações genotípicas e fenotípicas (população de treinamento). Por meio de modelos estatísticos são preditos os efeitos de cada SNP e baseado nessa predição é possível predizer os VGGP do conjunto de treinamento e de outros conjuntos de animais que necessitam ter apenas informações genotípicas (população de validação). A fim de obter informações acuradas na predição dos VGGP nas populações de validação é necessário que o DL entre marcadores e QTLs persista entre as populações de treinamento e validação. Quanto mais geneticamente distante for a população de treinamento da população de validação, menores serão as acurácias das predições genômicas para os indivíduos da população de validação.

A herdabilidade é um dos fatores que juntamente com o número de fenótipos, afetam a confiabilidade dos valores genéticos genômicos (RESENDE et al. 2008). Características de baixa herdabilidade necessitam de maior número de fenótipos e genótipos para estimar acuradamente os efeitos dos marcadores (MEUWISSEN et al., 2001; CALUS et al., 2008; GODDARD, 2009). Segundo Warkup (2005), a seleção genômica é vantajosa quando usada para características de alto custo de produção, de difícil mensuração, baixa herdabilidade, expressões tardias no animal e medidas em apenas um dos sexos. Há muitos desafios na implementação da seleção genômica, entre os quais estão o aumento das acurácias dos VGGP, a integração das informações genômicas e aumento da população referência (HAYES et al., 2009; GODDARD, 2009).

4 REFERÊNCIAS

ALCALÁ, A. M.; FRANGANILLO, A. R.; CÓRDOBA, M. V. Analisis genético de los niveles de consanguinidad en la raza Retinta. **Archivos Latinoamericanos de Producción Animal**, v. 44, p. 257-265, 1995.

ANUALPEC; anuário da pecuária brasileira . Agra FNP, 2013. p. 261.

ARDLIE, K. G.; LIU-CORDERO, S. N.; EBERLE, M. A. DALY, M.; BARRETT, J.; WINCHESTER, E.; LANDER, E.S.; KRUGLYAK, L. Lower-thanexpected linkage disequilibrium between tightly linked markers in humans suggests a role for gene conversion. **American Journal of Human Genetics**, v. 69, p. 582-589, 2001.

BOHMANOVA, J.; SARGOLZAEI, M.; SCHENKEL, F. S. Characteristics of linkage disequilibrium in North American Holsteins. **BMC Genomics**, v. 11, p. 421, 2010.

BOICHARD, D.; MAIGNEL, L.; VERRIER, E. The value of using probabilities of gene origin to measure genetic variability in a population. **Genetics Selection Evolution**, v. 29, p.5 -23, 1997.

BRITO, F. V.; NETO, J. B.; SARGOLZAEI, M.; COBUCI, J. A.; SCHENKEL, F. S. Accuracy of genomic selection in simulated populations mimicking the extent of linkage disequilibrium in beef cattle. **BMC Genetics**, v. 12, p. 80, 2011.

BRITO, F.V. **Diversidade Genética e Acurácia da Informação Genômica Em Bovinos de Corte**. 2011. 134f. Tese (Doutorado) – Universidade Federal Do Rio Grande Do Sul. Faculdade De Agronomia, Porto Alegre, RS, 2011.

CALUS, M. P. L.; MEUWISSEN, T. H. E.; ROOS, A. P. W.; VEERKAMP, R. F. Accuracy of genomic selection using different methods to define haplotypes. **Genetics**, v. 178, p. 553-561, 2008.

CARNEIRO, P. L. S. **Oscilação Genética e Comparação de métodos de seleção tradicionais e associados a marcadores moleculares**. 2002. 79p. (Dissertação) – Universidade Federal de Viçosa, Viçosa, MG, 2002.

CARNEIRO, P. L. S.; MALHADO, C. H. M.; AFFONSO, P. R. A. M.; EUCLYDES, R. F.; CARNEIRO, A. P. S.; CUNHA, E. E.; SOUZA, L. G. R. Comparação de metodologias de seleção sob oscilação genética. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia.** V. 60, p. 932-942, 2008.

COLLINS, A.; LONJOU, C.; MONTON, N. E. Genetic epidemiology of singlenucleotide polymorphisms. **Proceedings of the National Academy of Sciences**, v.96, p.15173-15177, 1999.

CRUZ, E. M. **Efeito da saturação e do tamanho de populações F2 e de retrocruzamento sobre a acurácia do mapeamento genético**. 2006. 119p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa, 2006.

DE LOS CAMPOS, G.; NAYA, H.; GIANOLA, D.; CROSSA, J.; LEGARRA, A.; MANFREDI, E.; WEIGEL, K.; COSTES J. M. Predicting quantitative traits with regression models for dense molecular markers and pedigree. **Genetics,** v. 182, p. 375-385, 2009.

EUCLYDES, R. F. **Uso do sistema para simulação Genesys na avaliação de métodos de seleção clássicos e associados a marcadores moleculares**. Viçosa:UFV, 1996.

FALCONER, D. S; MACKAY, T. F. C. **Introduction to Quantitative Genetics**. Harlow, Essex: Addison Wesley Longman, 1996.

FARNIR, F,; COPPIETERS, W.; ARRANZ, J. J,; BERZI, P.; CAMBISANO, N.; GRISART, B.; KARIM, L.; MARCQ, F.; MOREAU, L.; MNI, M.; NEZER, C.; SIMON, P.; VANMANSHOVEN, P.; WAGENAAR, D.; GEORGES, M.; Extensive genome-wide linkage disequilibrium in cattle. **Genome Research**, v.10, p.220-227, 2000.

FRANKHAM, R.; BRISCOE D.A.; BALLOU J.D. **Introduction to Conservation Genetics**. Cambridge: Cambridge University Press, 2002.

GODDARD, M. E.; HAYES, B. J. Genomic selection. **Journal of Animal Breeding and Genetics**, v. 124, p. 323-330, 2007.

GODDARD, M. Genomic selection: prediction of accuracy and maximization of long term response **Genetica**, v.136, p. 245 -257, 2009.

GROENEN, M. A.; MEGENS, H. J.; ZARE, Y.; WARREN, W. C.; HILLIER, L. W.; CROOIJMANS, R. P.; VEREIJKEN, A.; OKIMOTO, R.; MUIR, W. M.; CHENG, H. H. The development and characterization of a 60K SNP chip for chicken. **BMC Genomics,** v. 12, p 274-283, 2011.

HAYES, B. J.; BOWMAN, P. J.; CHAMBERLAIN, A. J.; GODDARD M. E. Invited review: Genomic selection in dairy cattle: progress and challenges. **Journal Dairy Science**, v.92, p. 433 -443, 2009.

HILL, W. G.; ROBERTSON, A. Linkage disequilibrium in finite populations. **Theoretical and Applied Genetics**, v.38, p.226-231, 1968.

HILLIER, L. W.; et al. Sequence and comparative analysis of the chicken genome provide unique perspectives of vertebrate evolution. **Nature**, v. 716 p. 432-695, 2004.

HOGGART, C. J.; CHADEAU-HYAM, M.; CLARK, T. G.; LAMPARIELLO, R.; WHITTAKER, J. C.; DE IORIO, M.; BALDING, D. J. Sequence-level population simulations over large genomic regions. **Genetics**, v.177, p.1725-1731. 2007.

KRANIS, A.; GHEYAS, A. A.; BOSCHIERO, C. Development of a high density 600K SNP genotyping array for chicken. **BMC Genomics,** v. 14, p. 59-62, 2013.

LANDER, E. S.; SCHORK, N. J. Genetic dissection of complex traits. **Science**, v.265, p.2037-2048, 1994.

LEDUR, M. C.; SCHMIDT, G. S.; FIGUEIREDO, E. A. P.; AVILA, V. S.; BALEN, L. Genetic and phenotypic parameters for productive traits in White egg layer stocks. **Pesquisa Agropecuária Brasileira**, v.28 p.1031–1037, 1993.

LEDUR, M.C.; PEIXOTO, J.O. Aplicações e Perspectivas da Genômica no Melhoramento Genético de Aves. In: SIMPÓSIO BRASILEIRO DE MELHORAMENTO ANIMAL, 8., 2010, Maringá. **Anais eletrônicos**... Maringá: SBMA,2010. Disponível em: <http://sbmaonline.org.br/anais/viii/palestras/pdfs/1.pdf> Acesso em: 21 jul. 2013.

LI, R.; LIM Y.; FANG, X.; YANG, H.; WANG, J.; KRISTIANSEN, K.; WANG, J. SNP detection for massively parallel whole-genome resequencing. **Genome Research**, v. 19, p. 1124-32, 2009.

LUAN, T.; WOOLLIAMS, J. A.; LIEN, S.; KENT, M.; SVENDSEN, M.; MEUWISSEN, T .H. E. The accuracy of genomic selection in norwegian red cattle assessed by cross-validation. **Genetics**, v. 183, p. 1119-1126, 2009.

LUTAAYA, E.; MISZTAL, I.; BERTRAND, J. K.; Mabry, W. Inbreeding in populations with incomplete pedigrees. **Journal of Animal Breeding and Genetics**, v.116, p.475-480, 1999.

MEUWISSEN, T. H. E.; GODDARD, M. E.; HAYES, B. J. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, p. 1819-1829, 2001.

MUIR, W. M. The interaction of selection intensity, inbreeding depression, and randon genetic drift on short and long-term response to selection: Results using finite locus and finite population size models incorporating directional dominance. **Journal Animal Science**, v. 79, p. 1-11, 2000.

OLIVEIRA, C. G.; CARNEIRO, P. L. S.; MALHADO, C.H.M; CUNHA, E. E.; EUCLYDES, R. F.; CARNEIRO, A. P. S.; SOUZA, L. G. R. Endogamia em populações de frangos de corte submetidas a diferentes metodologias de seleção. Revista Científica de Produção Animal, v. 8, p. 29-39.

OLIVEIRA, J. A.; BASTOS, J. F.; TONHATI, H. Endogamia em um rebanho da raça Guzerá. **Revista Brasileira de Zootecnia**, v.28, p.721-728, 1999.

ORITA, M.; IWAHANA, H.; KANAZAWA, H.; HAYASHI, K.; SEKIYA, T. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. **Proceedings of the National Academy of Sciences USA**, v.86, n.8, p.2766-70, 1989.

QANBARI, S.; PIMENTEL, E. C. G.; TETENS, J.; THALLER, G.; LICHTNER P.; SHARIFI, A. R.; SIMIANER, H., The pattern of linkage disequilbriom in German Holstein cattle. **Animal Genetics**, v. 41 p. 346-356, 2010.

RESENDE, M. D. V. **Genômica Quantitativa e Seleção no melhoramento de Plantas e Animais**. Colombo: Embrapa Florestas, 2008. p. 21.

RESENDE, M. D. V.; LOPES, P. S. L.; SILVA, R. L.; PIRES, I. E. Seleção genômica ampla (GWS) e maximização da eficiência do melhoramento genético. **Pesquisa Florestal Brasileira**, n.56, p.63-77, 2008.

ROOS, A. P. W.; HAYES, B. J.; SPELMAN, R. J.; GODDARD, M. E. Linkage disequilibrium and persistence of phase in holstein–friesian, jersey and angus cattle. **Genetics**, v. 179, p. 1503-1512, 2008.

SANTOS, L. G. **Seleção genômica em populações simuladas**. 2011. 59f. Tese (Doutorado) – Universidade Federal Do Rio Grande Do Sul. Faculdade De Agronomia, Porto Alegre, 2011.

SARGOLZAEI, M.; SCHENKEL, F. S. Extent and pattern of linkage disequilibrium in North American Holstein cattle evaluated using a 50k SNP panel. In: Kennedy Conference on Quantitative Genetics And Animal Breeding, 2008, **Resumos...** Guelph, 2008, p. 125-126.

SARGOLZAEI, M.; SCHENKEL, F. S. QMSim: a large-scale genome simulator for livestock. **Bioinformatics**, v.25, p. 680–681, 2009.

SCHAEFFER, L. R. Strategy for applying genome-wide selection in dairy cattle. **Journal of Animal Breeding and Genetics**, v. 123, p. 218-223, 2006.

SHEPHERD, R. K.; MEUWISSEN, T. H. E.; WOOLLIAMS, J. A. Genomic selection and complex trait prediction using a fast EM algorithm applied to genome-wide markers. **BMC Bioinformatics,** v. 11 p. 529, 2010.

SILVA, M. Y. G. B.; COBUCI, W. J. F. J. A.; GUARAGNA, G. P.; OLIVEIRA, P. R. P. Efeito da endogamia sobre características produtivas e reprodutivas de bovinos do ecótipo Mantiqueira. **Revista Brasileira de Zootecnia**, v. 30, p.1236-1242, 2001.

SOLBERG, T. R.; SONESSON, A.; WOOLIAMS, J.; MEUWISSEN, T. H. E. Genomic selection using different marker types and density. In: World Congress Of Genetics Applied To Livestock Production, 8., 2006. Proceedings. Belo Horizonte: Ed. da UFMG, 2006. 1 CD-ROM.

SORENSEN, A. C.; SORENSEN M. K.; BERG, P. Inbreeding in Danish dairy cattle breeds. **Journal of Dairy Science**, v. 88, p.1865 -1872, 2005.

STEPHENS, J. C.; SCHNEIDER, J. A.; TANGUAY, D. A.; et al. Haplotype variation and linkage disequilibrium in 313 human genes. **Science**, v. 293, p.489-493, 2001.

TORO, M. A.; Nieto, B.; Salgado, C. A note on minimization of inbreeding in small-scale selection programmes. **Livestock Production Science**, v.20, p.317-323, 1988.

UBABEF, União Brasileira de Avicultura. **Relatório Anual**. São Paulo, 2012. Disponível em: <http://www.ubabef.com.br> Acesso em: 01 mai. 2013.

VAN VLECK, L. D.; POLLAK, E. J.; OLTENACU, E. A. B. **Genetics for the animal sciences**. 1 ed. New York: Freeman and Company, 1987. 391p.

VERRIER, E.; COLLEAU, J. J.; FOULLEY, J. L. Long-term effects of selection based on the animal model BLUP in a finite population. **Theoretical and Applied Genetics**, v.87, p.446-454, 1993.

WARKUP, C. The impact of developments in genetics and genomics on livestock breeding and animal health. **Royal Agricultural Society of England**, v.169**,** 2008.

WONG, G. K.; et al. A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. **Nature**, v. 722 p. 432-717, 2004.

WRIGHT, S. Coefficients of inbreeding and relationship. **American Society of Naturalists**, v.56, n.645, p.330-338, 1922.

ZHAO, H.; NETTLETON, D.; DEKKERS, J. C. M. Evaluation of linkage disequilibrium measures between multi-allelic markers as predictors of linkage disequilibrium between single nucleotide polymorphisms. **Genetics Research**, v. 89, p.1-6, 2007.

CAPÍTULO 2 – ACURÁCIA DOS VALORES GENÉTICOS GENÔMICOS EM POPULAÇÕES COM DIFERENTES TAXAS DE ENDOGAMIA

1 INTRODUÇÃO

A perda de variabilidade genética pelo aumento nas taxas de endogamia da população deve ser controlada nos programas de melhoramento genético, sendo que o primeiro impacto da endogamia é a perda de alelos decorrente da oscilação genética (Muir, 2000), fator agravante para o progresso genético em programas de melhoramento animal. A estrutura da população, as estratégias de seleção, o tamanho efetivo da população (Ne) e a extensão do desequilíbrio de ligação (DL), permitem conhecer detalhadamente a estrutura genética de aves de postura e a resposta genética na seleção em características quantitativas em populações sob seleção.

Com a implementação da seleção genômica, estratégias de seleção se tornaram mais eficientes, promovendo maiores progressos genéticos em características de interesse econômico (VAN RADEN et al., 2009). A seleção genômica considera que as posições dos QTL (*loci* de características quantitativas) não são conhecidas, sendo necessária grande quantidade de marcadores SNP (polimorfismo de nucleotídeo único) ao longo de todo o genoma em DL, e que a associação entre SNP e QTL persista com o passar das gerações. Para obter os efeitos dos marcadores SNP, a metodologia RR-BLUP (*random regression best linear unbiased prediction*), tem sido amplamente utilizada e acurada para obter os valores genéticos genômicos preditos (**VGGP**) (MEUWISSEN et al., 2001; BRITO et al., 2011). O VGGP é obtido com o uso dos efeitos dos marcadores SNP, enquanto que o valor genético predito (**VGP**) é o obtido com informações fenotípicas dos animais e sua estrutura de *pedigree* tradicional.

A simulação de dados vem sendo utilizada para verificar o desempenho de modelos e detecção de QTL (CRUZ, 2006). Muito embora os custos de genotipagem tenham diminuído substancialmente, estudos de seleção genômica em larga escala ainda são de alto custo e grande parte dos *softwares* de simulação disponíveis publicamente não possuem as funcionalidades necessárias para estudos em

programas de melhoramento de animais domésticos. Diferente dos outros *softwares,* o QMSim (SARGOLZAEI; SCHENKEL, 2009) foi projetado para simular análises baseadas em grandes sequências genômicas e complexas estruturas de pedigree, simulando assim populações próxima às reais. O programa leva em conta características evolutivas da população como o DL, a ocorrência de mutação e de gargalos genéticos e as expansões populacionais e taxas de recombinação.

2 OBJETIVOS

O objetivo do presente trabalho foi avaliar a acurácia de predição dos valores genéticos genômicos para características de diferentes herdabilidades, em populações simuladas de aves com diferentes níveis de endogamia.

3 MATERIAL E MÉTODOS

3.1 Simulação das populações de aves

As simulações foram baseadas na estrutura populacional de uma população experimental de aves poedeiras White Leghorn, denominada CC, desenvolvida e mantida sob seleção desde 1987 pela Embrapa Suínos e Aves em Concórdia, SC. Esta linhagem tem como principal objetivo de seleção a melhoria na taxa de produção de ovos. Foram simulados dados fenotípicos e genotípicos para taxa de postura total de ovos (**PTO**), expressa em porcentagem, colhida da 17ª a 70ª semana de idade (totalizando 54 semanas de produção) e peso dos ovos às 32 semanas de idade (**PO**), expresso em gramas.

As estimativas de herdabilidade e dos componentes de variância utilizados como pressupostos para simulação destas características foram obtidas por Savegnago et al. (2011). Os cenários de simulação variaram quanto à magnitude da herdabilidade e quanto aos esquemas de acasalamento das populações sob seleção. Para realizar as simulações foi utilizado o programa QMSim (SARGOLZAEI; SCHENKEL, 2009). O programa foi escrito em linguagem C++ e

executado em sistema operacional Debian GNU/Linux versão 7.0 "Wheezy" 64-bit. O computador utilizado para gerar as simulações, está locado no Departamento de Ciências Exatas da FCAV/UNESP, Câmpus de Jaboticabal (Processo CNPq 475608/2011-4). O servidor possui processador intel® Core™ i-7-3930k com 12 núcleos de 3.20GHz cada e memória RAM de 64 Gb. Para a simulação dos dados foram utilizados simultaneamente os 12 núcleos de processamento.

3.2 Estrutura da população histórica

Foi simulada uma população histórica com tamanho efetivo (NE) calculado segundo Wright, (1931), de 1.100 reprodutores (550 machos e 550 fêmeas), considerando a união aleatória dos gametas, durante 1.000 gerações de tamanho constante. Após a milésima geração, foi provocado gargalo genético, diminuindo gradualmente o Ne até 640 ao longo de 20 gerações (Figura 1). O gargalo genético foi importante para gerar deriva genética, e assim o histórico de desequilíbrio de ligação. A redução do Ne é justificada pelo processo de domesticação, em que ocorre redução no tamanho efetivo da população (HENSON, 1992). Após o gargalo genético, a fim de gerar número suficiente de animais para os três cenários de populações recentes, a população histórica sofreu expansão por três gerações (geração 1.021 até 1.023), sendo selecionados aleatoriamente dos 640 animais 80 machos e 560 fêmeas, mantendo proporção de um reprodutor para sete reprodutrizes, com quatro progênies por reprodutriz. Ao final da população histórica (1.023 gerações), foram obtidos 2.240 animais (Figura 1).

3.3 Estrutura da população recente

Após a formação da população histórica, foram criados três cenários de populações recentes, divergentes apenas nos esquemas de acasalamento. Os cenários tiveram intuito de maximizar (REC1) ou minimizar (REC2) a endogamia e acasalamentos aleatórios (REC3). Em cada cenário foram simuladas duas características limitadas ao sexo: PO e PTO, com herdabilidades de 0,37 e 0,15, respectivamente, e variância fenotípica de 16,11g² e 130,32%², respectivamente (SAVEGNAGO et al., 2011).

A média dos coeficientes de endogamia (F) em cada cenário foi calculada segundo Falconer e Mackay (1996):

$$F_t = 1 - (1 - \frac{1}{2Ne})^t$$

em que **F** é o coeficiente de endogamia, **Ne** o tamanho efetivo inicial da população e **t** o número de gerações recentes. Para gerar populações divergentes quanto aos níveis de endogamia, o *software* QMSim utiliza o método de *simulated annealing*, adaptação do algoritmo Metropolis-Hastings para o problema de optimização global (SONESSON e MEUWISSEN, 2000).

Todos os cenários recentes tiveram como fundadores diferentes animais provenientes dos 2.240 da última geração da população de expansão, mantendo a proporção de um reprodutor para sete reprodutrizes, com quatro progênies por fêmea. Essas proporções foram mantidas ao longo de dez gerações discretas nos cenários das populações recentes, ou seja, não houve sobreposição de gerações (Figura 1).

Os animais das gerações recentes foram selecionados ao longo de 10 gerações com base nos maiores valores genéticos preditos, utilizando o método da máxima verossimilhança restrita, via BLUP (*best linear unbiased prediction*), usando equações de modelos mistos (HENDERSON, 1975) em modelo animal:

$$y = \mu + Za + \varepsilon$$

Em que **y** é o vetor de características fenotípicas, ***μ*** é a média geral, ***Z*** é a matriz de incidência dos efeitos aleatórios, ***a*** é o vetor de efeitos genéticos aditivos dos animais e $\boldsymbol{\varepsilon}$ é o vetor de erros aleatórios.

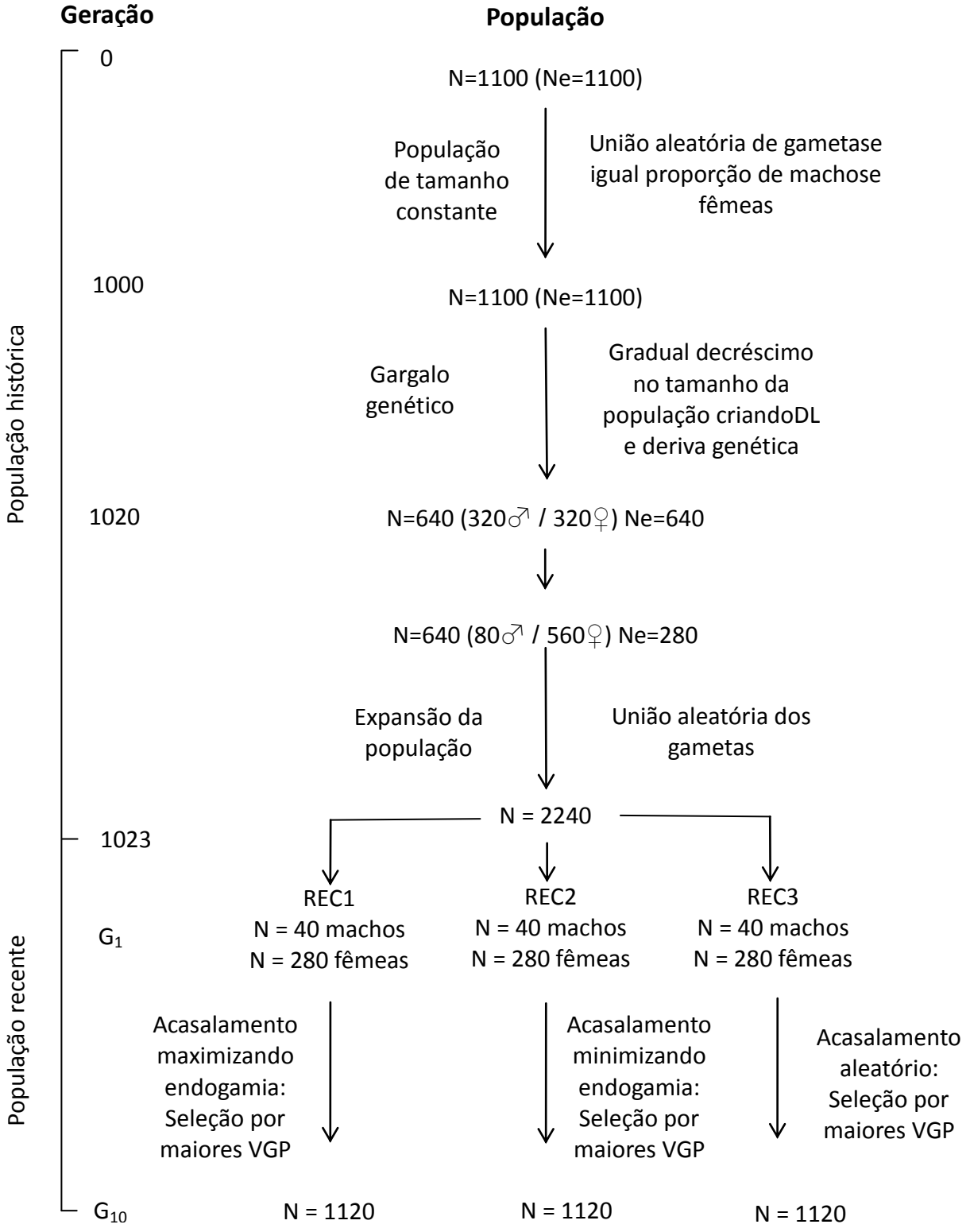


Figura 1. Estrutura populacional. G_0 a G_{10}: populações recentes da geração zero a dez, respectivamente; REC1, REC2 e REC3 representam os cenários das populações recentes; DL: desequilíbrio de ligação; N: número total de animais;Ne: tamanho efetivo da população; VGP: valor genético predito.

3.4 Genoma

O genoma da galinha doméstica é composto por 38 pares de cromossomos autossômicos e um par de cromossomos sexuais, totalizando 78 cromossomos (2n), consistindo em cinco pares de macrocromossomos, cinco pares de cromossomos intermediários, 28 pares de microcromossomos e um par de cromossomos sexuais (CHENG, 1997; BURT, 2013). O genoma das aves foi simulado com todos os cromossomos autossômicos a partir do genoma de referência *Gallus gallus* 4.0 (NCBI, 2013), para todos os cenários (REC1, REC2 e REC3), totalizando 958 cM. Apenas os cromossomos autossômicos foram simulados e o microcromossomo 16 foi omitido, por ter tamanho menor que 1 Mb, limite inferior para ser usado no QMSim. Foram simulados ao longo do genoma, 3.747 QTL, que representaram todos os QTL publicados para *Gallus gallus* (HU et al., 2013) e 49.978 marcadores SNPs igualmente espaçados em uma proporção de 52,17 marcadores por 1 cM.

A fim de alterar as frequências alélicas e gerar variabilidade genética ao longo das gerações, foram simulados eventos de deriva genética, que alteram a frequência dos alelos na população; recombinação, gerada a partir de uma distribuição de Poisson e distribuída aleatoriamente ao longo do genoma e mutação recorrente, em que é assumida alteração de um estado alélico para o outro e, consequentemente, não são criados novos alelos. Os eventos de deriva genética e mutação recorrente ocorreram apenas na população histórica, até a geração 1.020 (Figura 1), considerando taxa de mutação recorrente de $2,5 \times 10^{-5}$ para marcadores SNP e QTLs. Os parâmetros detalhados para gerar as populações simuladas estão apresentados na Tabela 1. A simulação foi repetida cinco vezes.

Tabela 1. Parâmetros considerados para gerar as populações simuladas de aves poedeiras

Estrutura da população	Parâmetros
População histórica (PH)	
Número de animais (Número de gerações constantes)	1100(1000)
Número de animais (Número de gerações - gargalo genético)	640(20)
Expansão da população (EXP)	
Número de machos fundadores (PH)	80
Número de fêmeas fundadoras (PH)	560
Número de gerações	3
Número de filhos por mãe	4
Populações recentes (RECs)	
Número de machos fundadores (EXP)	40
Número de fêmeas fundadoras (EXP)	280
Número de gerações	10
Número de filhos por mãe	4
Porcentagem de machos por geração de filhos	50%
Esquemas de acasalamento	MaxF/MinF/Aleatório
Sobreposição de geração	Discreta
Seleção	Valor genético predito
Método de estimação dos valores genéticos	BLUP
Herdabilidade das características	0,15(PTO)/0,37(PO)
Variância fenotípica	130,32 (PTO)/16,11(PO)
Genoma	
Número de cromossomos	27
Tamanho total	958 Mb
Número total de marcadores	49.978
Número total de QTL	3779
Distribuição dos QTL	Aleatória
MAF dos marcadores	0,05
Efeito alélico aditivo dos QTL	Distribuição Gamma (0,40)
Taxa de marcador genético perdido	0,01
Taxa de erro de genotipagem	0,005
Taxa de mutação recorrente do marcador	$2{,}5.10^{-5}$
Taxa de mutação recorrente do QTL	$2{,}5.10^{-5}$

REC1, REC2 e REC3: cenários de populações recentes 1, 2 e 3; MinF: acasalamentos realizados minimizando o efeito da endogamia; MaxF: acasalamentos realizados maximizando o efeito da endogamia; QTL: *Quantitative Trait Loci* (*loci* de característica quantitativa); MAF: *Minor Allele Frequency (*alelo menos frequente); PTO: produção total de ovos; PO: peso de ovos as 32 semanas de idade;BLUP: best linear unbiased prediction. A simulação foi repetida cinco vezes.

Em estudos com simulações é possível obter o valor genético verdadeiro (**VGV**) dos animais para cada característica. O VGV foi calculado como a somatória dos efeitos de todos os QTL de cada animal:

$$VGV_i = \sum_{j=1}^{3747} (x_{ijk} g_{jk})$$

em que x_{ijk} é o número de cópias do alelo ***k*** que o individuo ***i*** tem na posição ***j*** e g_{jk} é o efeito do QTL do alelo ***k*** na posição ***j***. Para gerar os efeitos dos QTLs foi utilizada a distribuição gamma com parâmetro de forma igual a 0,40 (HAYES; GODDARD, 2001).

O desequilíbrio de ligação (DL) foi mensurado por r^2, que é o coeficiente de correlação das frequências dos alelos em dois ou mais loci (HILL; ROBERTSON, 1968).

$$r^2 = \frac{D^2}{f(A)\,f(a)\,f(B)\,f(b)}$$

em que D = f(AB)-f(A)f(B) e f(AB), f(A), f(a), f(B), f(b) são as frequências dos alelos marcadores. Foi calculado o DL para todas as populações nas gerações genotipadas, considerando todos os 27 cromossomos em cinco repetições.

O Ne das populações recentes foi calculado segundo Wright (1931), citada por Falconer e Mackay (1996) :

$$\frac{1}{N_e} = \frac{1}{4N_m} + \frac{1}{4N_f}$$

em que **Ne** é o tamanho efetivo da população, $\mathbf{N_m}$ o número de machos e $\mathbf{N_f}$ o número de fêmeas acasaladas.

3.5. Tendência genética

Tendências genéticas para cada característica foram obtidas a partir da regressão linear simples dos respectivos valores genéticos em função das gerações.

Foi utilizada a estatística t, cuja hipótese de nulidade considerada foi que o coeficiente de regressão é igual a zero

3.6. População de treino e validação

Para cada cenário de população recente (REC1, REC2 e REC3), foram criados conjuntos de treinamento e validação dos modelos estatísticos. Primeiro os fenótipos e genótipos dos conjuntos de treinamento foram utilizados para estimar os efeitos dos marcadores SNP. Em seguida, apenas os genótipos dos conjuntos de validação foram considerados na predição dos valores genéticos genômicos destes animais. Foi utilizada a divisão no tempo para criar os conjuntos de treino e validação, ou seja, animais pertencentes às gerações 7, 8, 9 foram utilizados para o treinamento dos modelos e a geração dez para a validação. Os conjuntos de treino para cada cenário de população recente (REC1, REC2 e REC3) foi formado por 960 animais com os maiores valores de acurácia para os VGP nas gerações sete, oito e nove. O conjunto de validação de cada cenário conteve 1.120 animais da décima geração recente (Figura 2). Foi assumido que os animais desta última geração não tiveram a expressão de seus fenótipos, sendo estes apenas genotipados.

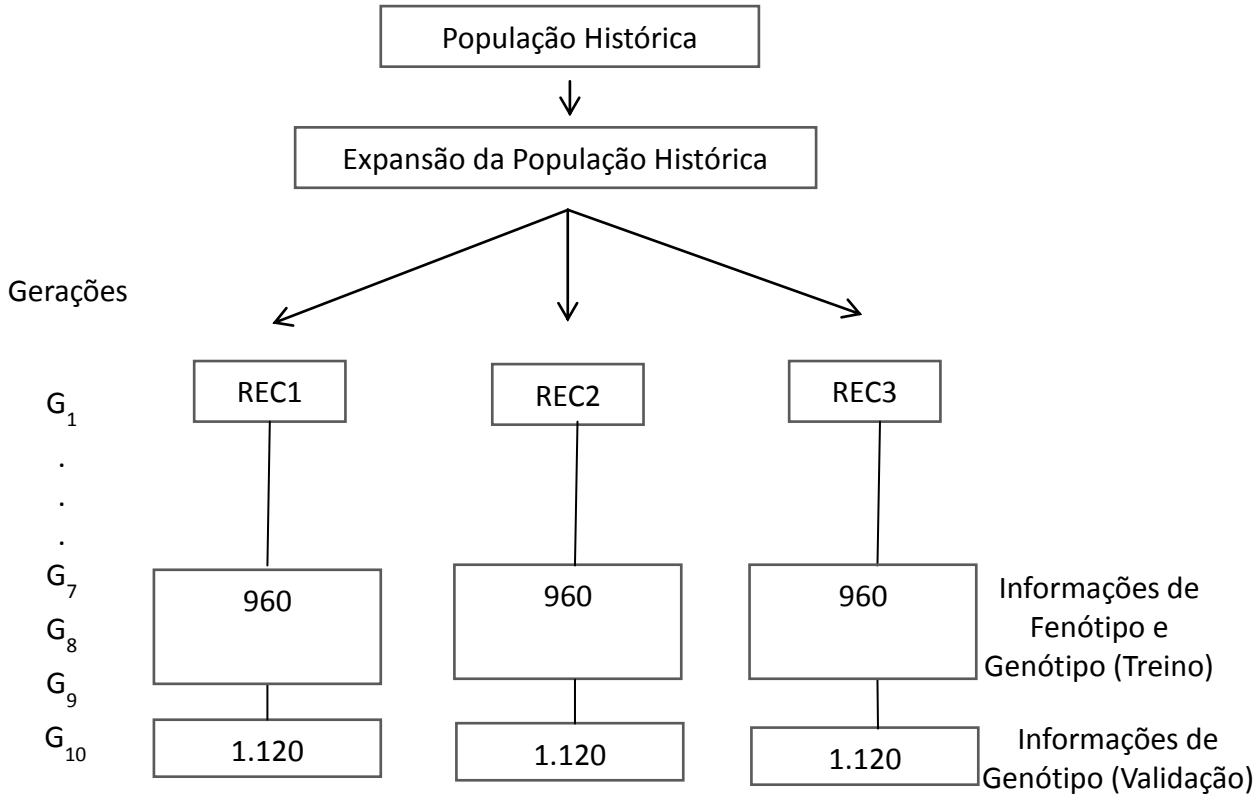


Figura 2. Estrutura de treino e validação dos cenários das populações recentes. G: gerações recentes; REC1, REC2 e REC3: cenários das populações recentes 1, 2 e 3.

Para obter os valores fenotípicos (y_{ij}) dos animais ***i*** para a característica ***j*** (PTO e PO) foi utilizada a seguinte equação:

$$y_{ij} = VGV_{ij} + \sigma_{ij}$$

Sendo VGV_{ij} o valor genético verdadeiro do animal ***i*** para a característica ***j*** e σ_{ij} a variâncias aleatória residual do animal ***i*** para a característica ***j***.

Os VGP, com os respectivos erros-padrão, foram estimados para todos os animais das gerações recentes pelo software AIREMLF90 (MISZTAL et al., 2002). A acurácia do valor genético (Acc) de cada animal foi obtida pela equação descrita por Gilmour, (2009):

$$Acc = \sqrt{1 - \frac{s_i^2}{(1 + f_i)\sigma_A^2}}$$

em que S_i é o erro-padrão calculado via BLUP para o individuo i, f_i é ocoeficiente de endogamia e σ_A^2 é a variância genética aditiva.

Para predizer os efeitos dos SNPs, nos conjuntos de treinamento, foi utilizado o modelo estatistico RR-BLUP (*random regression best linear unbiased prediction*), em que é assumido efeito aleatório igual para todos os SNP (MEUWISSEN et al., 2001). Na simulação, toda variação genética aditiva foi atribuida apenas aos QTLs. Assim, nesta simulação, foi pressuposto que os marcadores foram capazes de explicar toda variação genética aditiva, não havendo efeito poligênico residual (variação genética não explicada pelos marcadores). Como as caracteristicas PTO e PO apresentam valores fenotípicos apenas em fêmeas, obteve-se os valores genéticos preditos deregredidos (***VGPd***) como proposto por Garrick et al., (2009), aumentando o número de informações na variável resposta, o que elevou o poder de predição do modelo RR-BLUP. As variâncias dos marcadores foram consideradas como desconhecidas e preditas pela máxima verosimilhança restrita (REML), sendo as informações genotípicas ajustadas usando o modelo:

$$VGPd = 1 + Zm + \varepsilon$$

$$m \sim (0, \mathrm{G} = I\sigma_m^2)$$

$$\varepsilon \sim (0, R = I\,{}^{\sigma_e^2}/_{\omega_i})$$

em que ***VGPd*** são os valores genéticos preditos deregredidos (GARRICK et al., 2009), ***m*** é o vetor de efeitos aleatórios dos marcadores, ε é um vetor de erros aleatórios, **Z** a matriz de incidência *d*o conjunto de treinamento, σ_m^2 variância comum a cada efeito de marcador, σ_e^2 variância residual e ω ponderador para o animal ***i***. Efeitos fixos não foram simulados, portanto não foram incluídos no modelo que também não incluiu informações do *pedigree* tradicional.

Com os efeitos dos marcadores preditos pelo modelo RR-BLUP, foi possível obter o VGGP dos animais ***j*** do conjunto de treinamento e validação. O modelo dos animais de treinamento foi:

$$VGGP_j = \sum_{i}^{n} Z_{ij}\hat{m}_i$$

em que ***i*** é o alelo específico do i-ésimo marcador do animal ***j***, ***n*** o total de marcadores, $\hat{\boldsymbol{m}}$ é o vetor de efeitos dos marcadores SNPs e **Z** matriz de marcadores genéticos de cada animal do conjunto de treinamento. Igualmente foram obtidos os VGGP para o conjunto de validação, apenas substituindo a matriz **Z** pela matriz **X** de marcadores genéticos de cada animal do conjunto de validação. As análises foram realizadas utilizando o *software* GS3 (LEGARRA et al., 2011), com 100.000 iterações, excluindo as primeiras 10000, em um intervalo de 90. Foi feito o controle de qualidade genômico dos dados, descartando marcadores com MAF menores que 5%. Segundo Lewontin (1995) e Qanbari et al. (2010), marcadores SNP com MAF menor que 5% tem baixo poder em predizer o DL. A acurácia do VGGP foi obtida pela correlação simples de Pearson entre o VGGP e o VGV, em cinco repetições para cada população em cada característica. Diferenças estatísticas entre as acurácias médias foram obtidas pelo teste de Tukey, com nível de significância de 5%.

4. RESULTADOS E DISCUSSÃO

4.1 Endogamia, homozigose e Ne

Na geração zero das populações recentes os valores de coeficiente de endogamia, desequilíbrio de ligação e taxa de homozigose foram iguais nas características simuladas (PTO e PO). A partir da primeira geração sob seleção, esses parâmetros variaram em razão dos diferentes esquemas de acasalamentos (REC1, REC2 e REC3). As maiores diferenças, ao longo das 10 gerações recentes, foram entre o cenário REC1 e os demais cenários, com relação ao coeficiente de endogamia e homozigose (Figura 2). A média do coeficiente de endogamia das 10 gerações recentes variou de 0,06 ± 0,03 a 0,22 ± 0,12 nos cenários simulados para

PTO e 0,05 ± 0,03 a 0,20 ± 0,12 nos relativos à PO (Tabela 2). Os valores para os coeficientes de endogamia dos cenários REC2 e REC3 foram semelhantes aos relatados por Rosa et al. (2013), cujo coeficiente de endogamia foi 0,07 para uma linhagem de White Leghorn e metade dos 0,12 de Oliveira et al. (2006) em dados simulados para frango de corte em populações com acasalamento aleatório e Ne de 166,67. As médias da homozigose deste trabalho (Tabela 2) foram semelhantes às obtidas por Qanbariet al. (2010), que observaram média de 0,63 ± 0,15 em linhagem comercial de White Leghorn.

Tabela 2. Descrição do painel de marcadores das populações recentes genotipadas.

	PTO			PO		
	REC1	REC2	REC3	REC1	REC2	REC3
SNPs total	49.978	49.978	49.978	49.978	49.978	49.978
QTL total	3.747	3.747	3.747	3.747	3.747	3.747
DM[1] (Mb)	52,17	52,17	52,17	52,17	52,17	52,17
SNPs com MAF >5%[1]	31.239	38.414	38.063	33.859	38.792	38.283
Ne	140	140	140	140	140	140
Média de homozigose[2]	0,75(0,04)	0,69(0,01)	0,70(0,02)	0,74(0,05)	0,69(0,01)	0,70(0,02)
Média de endogamia[2]	0,22(0,12)	0,04(0,01)	0,06(0,03)	0,20(0,12)	0,03(0,04)	0,05(0,03)

REC1, REC2 e REC3: cenários de populações recentes 1, 2 e 3; DM: Densidade de marcadores SNPs ao longo do genoma; PTO: postura total de ovos; PO: peso dos ovos as 32 semanas de idade; QTL: *loci* de características quantitativas; MAF: alelo menos frequente; SNP: polimorfismos de nucleotídeos únicos; Ne: tamanho efetivo da população.
[1]Médias a partir das 10 gerações recentes.
[2]Médias e seus respectivos erros-padrão em parênteses obtida a partir das 10 gerações recentes.

As médias dos coeficientes de endogamia e de homozigose ao longo das gerações recentes em REC1, para PTO e PO, foram maiores que as observadas em REC2 e REC3 (Figura 2).

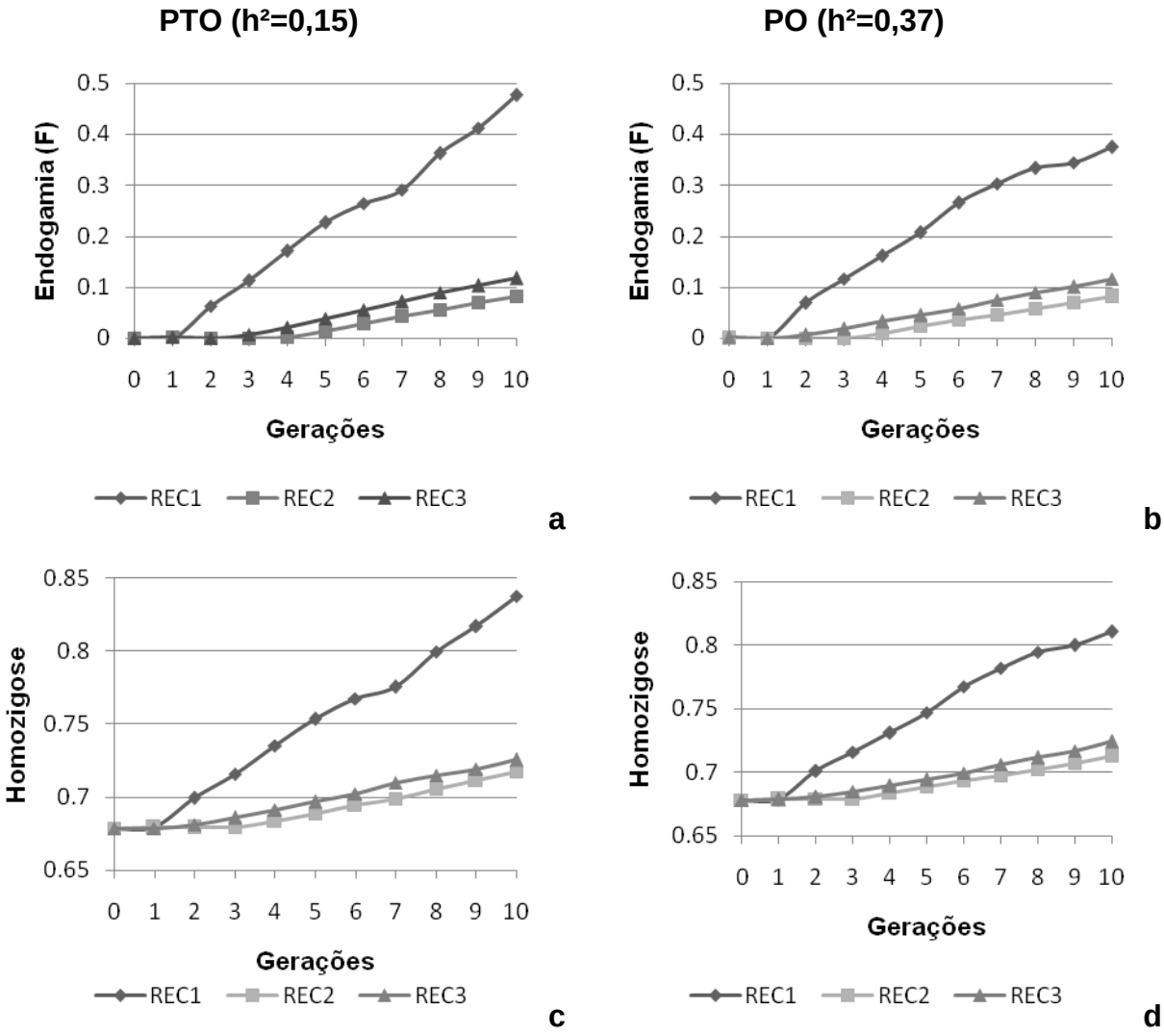


Figura 2. Variação das taxas de endogamia e homozigose nas populações recentes das características de produção total de ovos (PTO) e peso dos ovos (PO), respectivamente, **a**, **c** e **b**, **d**; h²: herdabilidade.

A seleção com base nos valores genéticos preditos pelo BLUP tradicional causa implicações como a seleção de indivíduos mais aparentados que se acasalados ao acaso, ocorrendo formação de linhagens ou famílias com endogamia elevada na população (OLIVEIRA et al., 2006). No entanto, os valores de endogamia de REC2 (acasalamentos visando minimizar a endogamia) e REC3 (acasalamentos aleatórios) foram muito próximos entre si e diferentes de REC1 (acasalamentos visando maximizar a endogamia) (Figura 2a e 2b), ou seja, em REC1 o incremento da endogamia ocorreu principalmente pelo acasalamento direcionado para animais mais aparentados e não apenas pelo efeito que a metodologia BLUP tem em

selecionar animais da mesma família, proporcionando aumento na taxa de endogamia dos animais sob seleção. Um dos principais efeitos da endogamia foi a redução na frequência dos genótipos heterozigotos o que causou a elevação na frequência dos genótipos em homozigose, diretamente proporcional ao aumento da endogamia nas populações recentes (Figura 2c e 2d).

4.2 Tendência genética

As tendências genéticas por geração nos cenários de populações REC1, REC2 e REC3 para PTO e PO foram apresentadas na Figura 3.

Os coeficientes de regressão (b) foram significativamente diferentes de zero ($p<0,05$) pelo teste t. As diferenças nos ganhos genéticos entre os cenários REC1, REC2 e REC3 dentro de PTO, indicaram maior ganho genético nas populações com maior endogamia, seguida por REC2 e REC3. Para PO, populações com acasalamento aleatório (REC3) apresentaram maiores coeficientes de regressão, diferente dos resultados apresentados para PTO. O acasalamento dos animais nas populações recentes foi realizado para os maiores VGP. No entanto, nos cenários REC1 e REC2, os acasalamentos foram realizados visando primeiramente maximizar (REC1) e minimizar (REC2) os efeitos da endogamia nestas populações. A necessidade de atender primeiramente estas premissas no acasalamento pode ter diminuído o ganho genético médio por geração destas populações, evidenciado pelo coeficiente de regressão, já que animais com maiores VGP podem não ter sido considerados nos acasalamentos por não priorizarem a necessidade de maximizar e minimizar os efeitos da endogamia em REC1 e REC2, respectivamente, o que resultou em ganho genético médio inferior em relação a REC3. Assim, o ganho genético médio por geração para PO foi de 23,90%; 21,30% e 23,15% em REC1, REC2 e REC3, respectivamente e para PTO foi de 21,72%; 19,49% e 22,83% em REC1, REC2 e REC3, respectivamente, indicando ganhos genéticos maiores em PO, dado que esta característica tem maior herdabilidade. Segundo Meuwissen e Woolliams (1994), populações com Ne entre 30 e 250 não apresentam decréscimos no valor adaptativo. Assim, o Ne igual para todos os cenários, a variação do ganho genético pode ser explicada apenas pelas diferentes taxas de endogamia e herdabilidade.

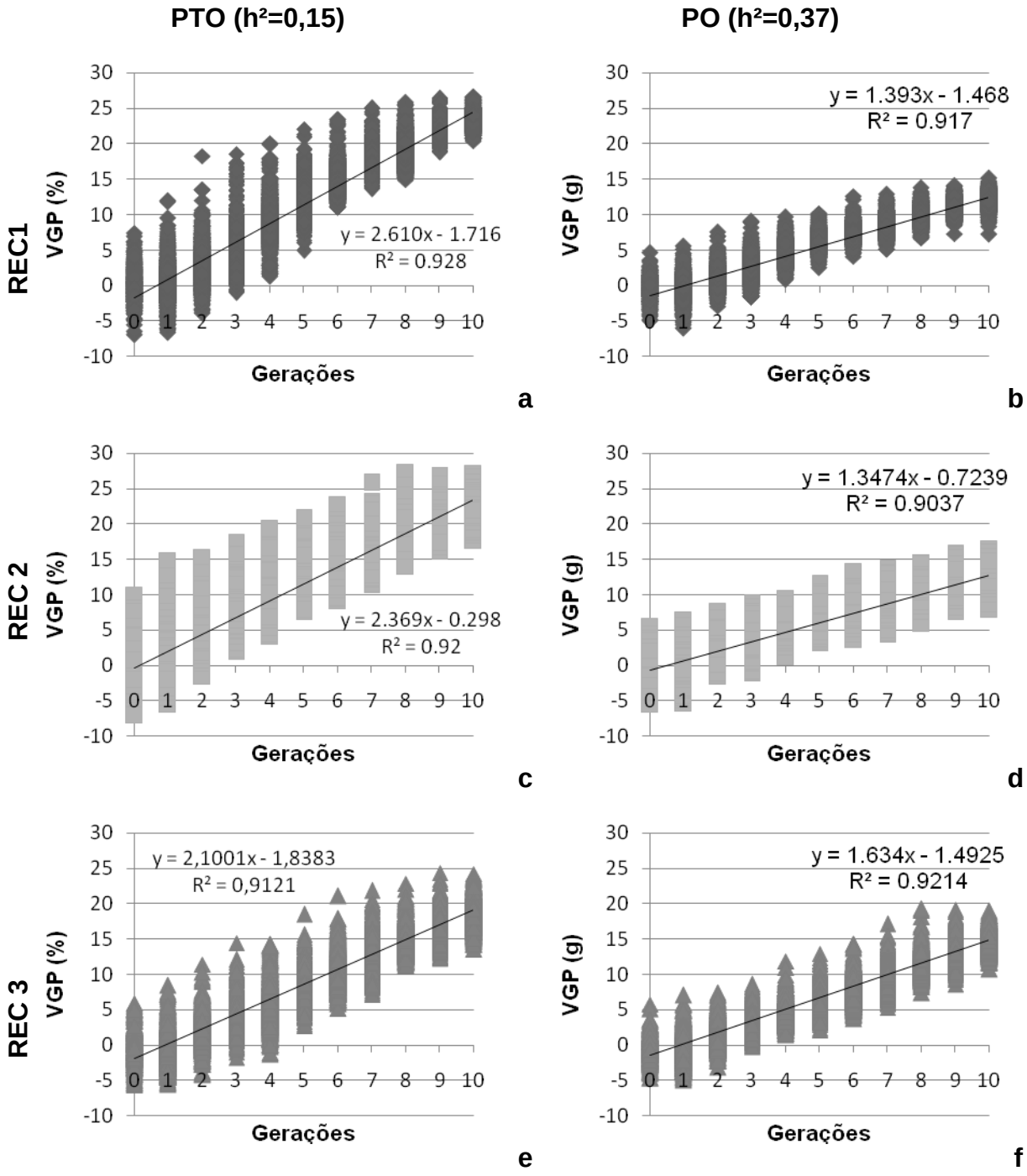


Figura 3. Tendências genéticas das características produção total de ovos (PTO) em REC1 (**a**), REC2 (**c**) e REC3 (**e**) e peso dos ovos (PO) em REC1 (**b**), REC2 (**d**) e REC3(**f**). VGP: valor genético predito; **R²**: coeficiente de determinação;REC1, REC2 e REC3: cenários de populações recentes 1, 2 e 3; h²: herdabilidade.

4.3 Desequilíbrio de ligação

O número de marcadores em DL e a média de r^2 entre os marcadores nas gerações genotipadas são apresentados na Tabela 3. O DL ou a correlação entre alelos em diferentes posições é geralmente dependente da taxa de recombinação entre os *loci* polimórficos em função da distância entre eles (RESENDE et al., 2008). Bohmanova et al., (2010) demonstraram que r^2 é a medida mais apropriada para estimar DL de marcadores bialélicos, tais como os SNP, pois r^2 é menos sensível à frequência alélica. Qanbariet al. (2010) trabalharam com populações comerciais de poedeiras das raças White Leghorn e GmbH, obtendo DL à distâncias 0-0,05 Mb de 0,67±0,38 para White Leghorn e 0,28±0,30 para GmbH. Estas diferenças no DL são explicadas na formação da população histórica das aves White Leghorn, que tem origem em poucas aves, gerando uma população com menor tamanho efetivo em comparação com as aves GmbH. Na última geração histórica (geração 1023 – Figura 1), o DL teve valor de 0,29 ± 0,28 à distância de 0,00-0,05 Mb entre SNP adjacentes. Esse valor inicial foi o mesmo utilizado no início dos cenários REC1, REC2 e REC3. Após 10 gerações sob seleção, a maior média de r^2 foi 0,52 ± 0,37 à distância entre 0,00-0,05 Mb entre SNP adjacentes em REC1, para a característica PTO (Tabela 3). O alto valor de DL em curtas distâncias e a acentuada diminuição como o aumento da distância fisica dos marcadores é comum a todos os cenários de população recente e esta de acordo com os resultados obtitos por Andreescu, (2007) que, em estudos com aves de corte, encontrou cerca de 10% dos marcadores SNP com r^2 superior a 0,8 em distâncias de 0,25-0,5 cM entre os marcadores adjacentes. Esta porcentagem caiu para menos de 1% com distâncias maiores que 1 cM. Quantidade menor do que a relatada por Heifetzet al., (2005), em aves poedeiras, entretanto neste estudo foi utilizado microssatélite e outra mensuração do DL.

Tabela 3. Médias dos desequilíbrios de ligação das gerações 7, 8, 9 e 10. Para 27 cromossomos simulados em 5 repetições. As herdabilidades 0,15 e 0,37 representam as características produção total de ovos (PTO) e peso do ovo (PO), respectivamente. Entre parênteses estão os desvios-padrão.

h^2	Distância (Mb)	REC1		REC2		REC3	
		Pares	r^2 (DP)	Pares	r^2 (DP)	Pares	r^2 (DP)
0,15	0,00-0,05	4142	0,52(0,37)	3833	0,37(0,34)	3815	0,38(0,34)
	0,05-0,10	6086	0,47(0,36)	5675	0,32(0,31)	5678	0,33(0,31)
	0,10-0,20	9932	0,42(0,34)	9367	0,27(0,28)	9372	0,27(0,28)
	0,20-0,30	9721	0,38(0,32)	9289	0,22(0,24)	9287	0,23(0,25)
	0,30-0,40	9771	0,35(0,31)	9422	0,19(0,22)	9417	0,20(0,23)
	0,40-0,50	10837	0,33(0,30)	10493	0,17(0,20)	10459	0,18(0,21)
	0,50-0,60	9497	0,31(0,29)	9219	0,16(0,19)	9232	0,17(0,20)
	0,60-0,70	9448	0,30(0,28)	9201	0,15(0,18)	9191	0,16(0,19)
	0,70-0,80	9515	0,29(0,27)	9243	0,14(0,17)	9245	0,15(0,18)
	0,80-0,90	9977	0,29(0,27)	9752	0,14(0,16)	9741	0,15(0,17)
	0,90-1,00	9889	0,26(0,21)	9651	0,13(0,16)	9632	0,14(0,17)
0,37	0,00-0,05	3951	0,45(0,36)	3799	0,36(0,33)	3807	0,37(0,34)
	0,05-0,10	5826	0,41(0,35)	5659	0,31(0,31)	5676	0,32(0,31)
	0,10-0,20	9577	0,35(0,32)	9384	0,26(0,27)	9394	0,26(0,27)
	0,20-0,30	9448	0,31(0,30)	9283	0,21(0,24)	9265	0,22(0,24)
	0,30-0,40	9531	0,28(0,28)	9386	0,19(0,21)	9378	0,19(0,22)
	0,40-0,50	10613	0,26(0,27)	10478	0,17(0,20)	10452	0,17(0,20)
	0,50-0,60	9327	0,25(0,26)	9219	0,15(0,18)	9217	0,16(0,19)
	0,60-0,70	9294	0,24(0,25)	9197	0,14(0,17)	9197	0,15(0,18)
	0,70-0,80	9361	0,23(0,24)	9249	0,14(0,16)	9247	0,14(0,17)
	0,80-0,90	9840	0,22(0,24)	9734	0,13(0,16)	9725	0,14(0,16)
	0,90-1,00	9732	0,22(0,23)	9646	0,12(0,15)	9642	0,13(0,16)

REC1, REC2 e REC3: cenários de populações recentes 1, 2 e 3; r^2:Desequilíbrio de ligação; h^2: herdabilidade.

Brito (2011) verificou que a seleção causa DL apenas em regiões específicas do genoma, resultando em pequena influência no DL total do genoma. A redução do DL é mais relacionada a fatores como o Ne, taxa de recombinação entre os marcadores e aumento na distância física entre os marcadores. Quanto mais próximos estiverem dois locos, menor a frequência de recombinações entre esses e mais tempo permanecerão ligados, após eventos como a seleção e redução do tamanho efetivo da população, fatores que elevam o DL na população (RESENDE, 2008). Porém, marcadores próximos não estão sempre em DL (ARDLIE et al., 2001) e, ao contrário, DL foi observado entre marcadores localizados a grandes distâncias um do outro (COLLINS et al., 1999; STEPHENS et al., 2001).

Foi observado que, quanto maior o coeficiente de endogamia da população, maior foi a extensão do DL nestas populações. Nos cenários REC1 para ambas as herdabilidades, o aumento da endogamia justificou os maiores valores de DL. Carneiro (2002) relatou que o aumento na taxa de endogamia elava a frequência de locos fixados no genoma e García-Gámez et al. (2012), observaram que são esperados altos valores de homozigose e DL em populações com locos fixados. Diferenças entre os valores de DL ao longo do genoma, para os animais sob seleção (Figura 4 e 5), indicam que o incremento de endogamia em uma população eleva os valores de DL da mesma, interferindo na magnitude do DL, no entanto, não altera a gradual redução do DL ao longo do distanciamento entre os marcadores adjacentes. A média de r^2 em PTO, para todas as populações genotipadas, foi de 0,26 ± 0,11, superior à apresentada por PO de 0,23 ± 0,09, justificada pelos maiores valores de endogamia em PTO.

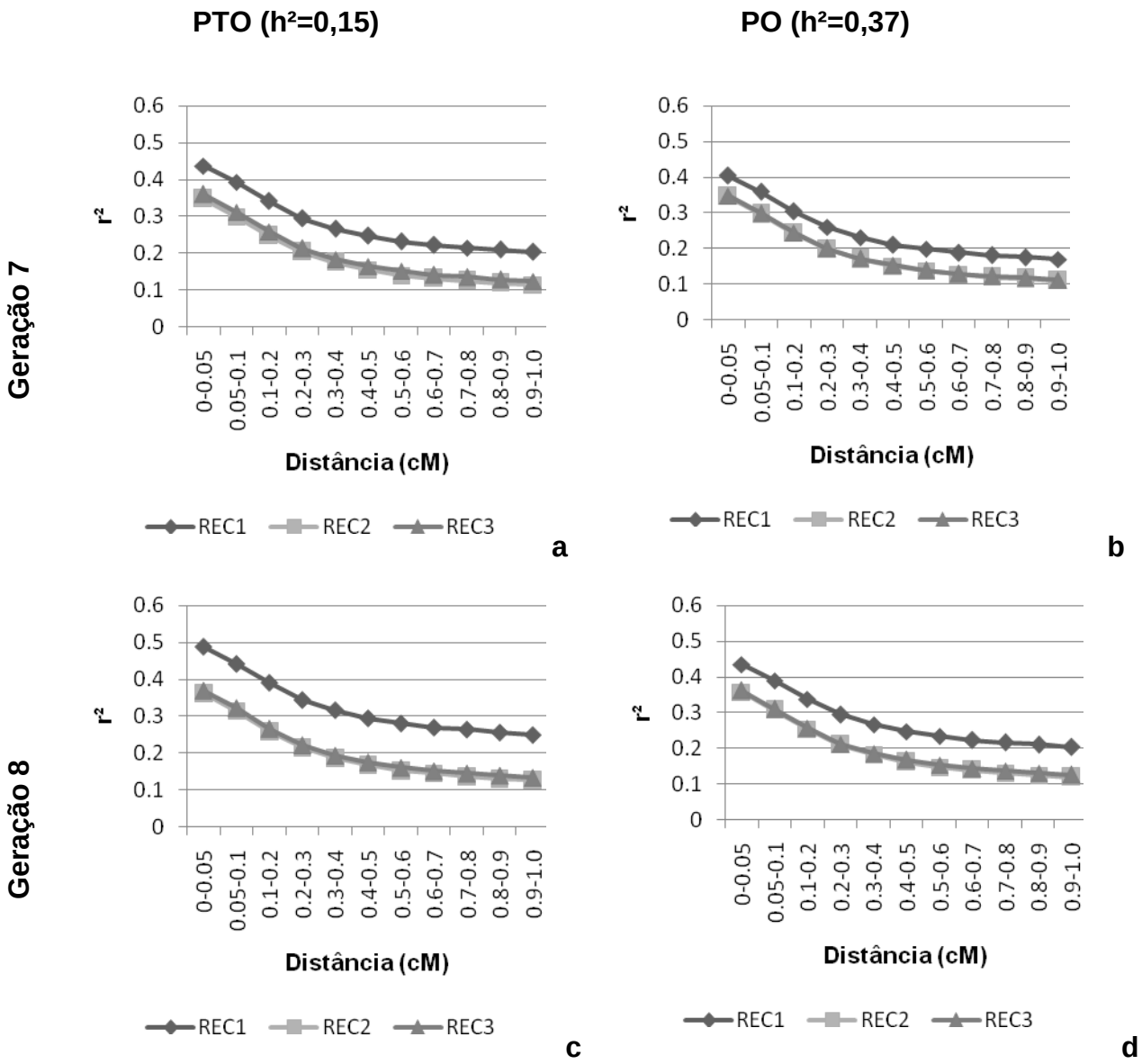


Figura 4. Variação do desequilíbrio de ligação (r^2) das populações recentes genotipadas (gerações 7 e 8) das características produção total de ovos (**PTO**) e peso dos ovos (**PO**) respectivamente, **a**,**c** e **b**,**d**, h^2:herdabilidade.

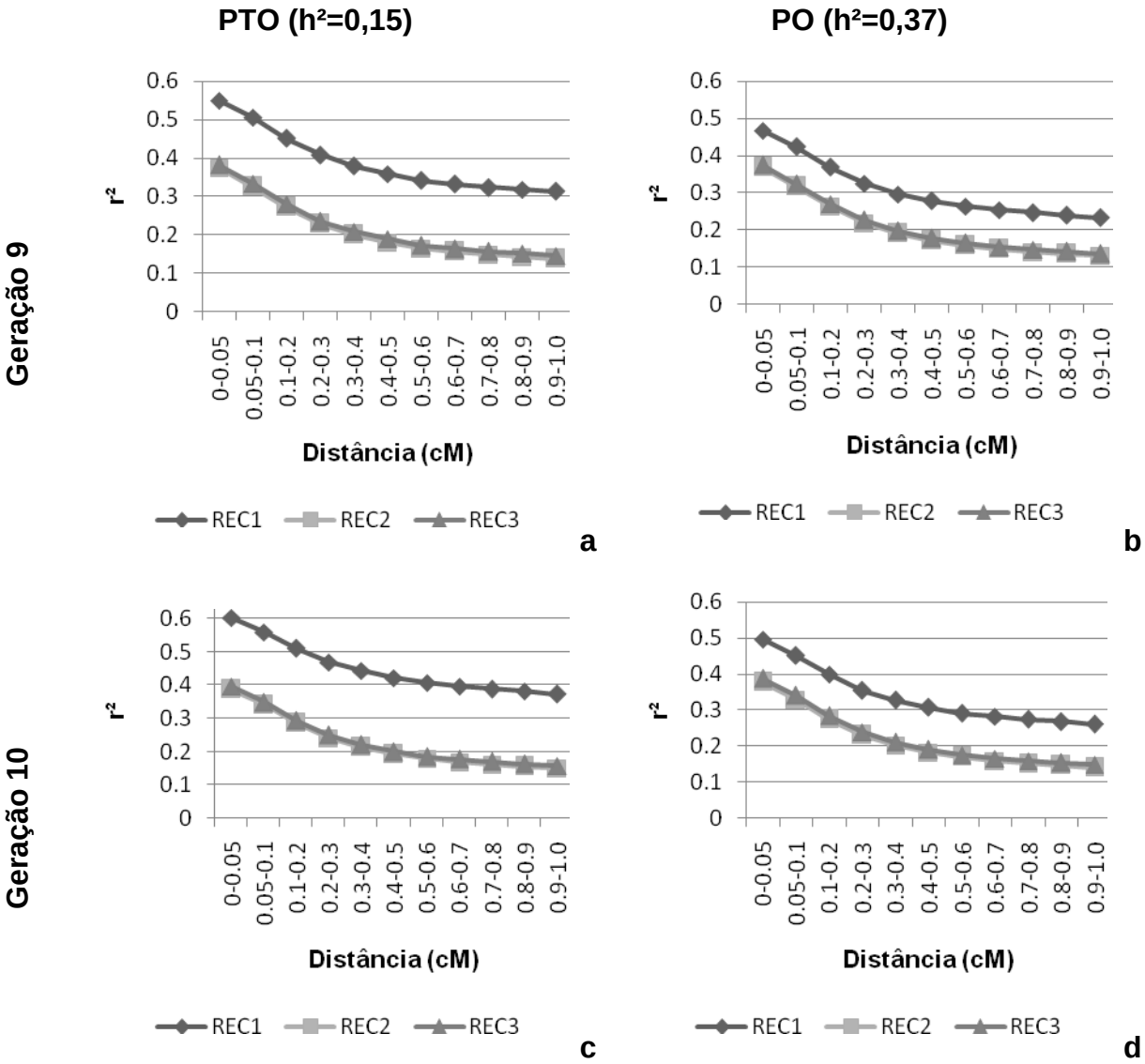


Figura 5. Variação do desequilíbrio de ligação (r^2) das populações recentes genotipadas (gerações 9 e 10) das características produção total de ovos (**PTO**) e peso dos ovos (**PO**) respectivamente, **a**,**c** e **b**,**d**, h^2:herdabilidade.

4.4 Acurácia dos VGGP

Os valores de acurácia dos VGGP para PTO e PO apresentaram diferença significativa ($p < 0,05$) entre as populações de treinamento dos cenários REC1, REC2 e REC3. Os maiores valores de acurácia para PTO e PO foram 0,61 ± 0,11 e 0,73 ± 0,02, respectivamente, ambos em REC1. Porém, não houve diferenças significativas ($p > 0,05$) entre os valores de acurácia nas populações de validação (Tabela 4). Meuwissenet al. (2001) obtiveram correlações entre os VGGP e os VGV

próximos à 0,73 ± 0,03 em uma população de treinamento, similar ao observado para PO em REC1. Brito et al. (2011) relataram acurácia de 0,45 para uma característica simulada com herdabilidade de 0,10, com média de r^2 = 0,26 e 960 animais na população de treinamento.

As acurácias dos VGGP obtidas nas populações de validação sugerem que diferentes taxa de endogamia e herdabilidade não alteram o poder dos modelos estatísticos em predizer o VGGP nesses cenários. As baixas acurácias dos VGP podem ter influenciado na predição dos efeitos dos marcadores, justificando os valores de acurácia dos VGGP. Dos fatores que podem influenciar as estimativas das acurácias, tanto nas populações de treino quanto nas de validação, destacamos o desequilíbrio de ligação. Os cenários de REC1 apresentaram valores de DL superiores aos de REC2 e REC3 em todas as distâncias físicas e ambas as características simuladas (Figura 4 e 5), justificando os maiores valores de acurácia em relação aos demais cenários.

Mesmo com grande diferença entre os coeficientes de endogamia de REC1 para as demais, a variação dos valores genéticos de todos os cenários foi semelhante. A estimativa de herdabilidade pouco influenciou na acurácia dos VGGP, não havendo diferença significativa ($p > 0,05$) entre as características PTO e PO em todos os cenários simulados das populações de validação, mas tendo influência nos cenários de treinamento.

Tabela 4. Correlação entre as médias dos valores genéticos genômicospreditos (VGGP) eas médias dos valores genéticos verdadeiros (VGV) das populações de treino e validação.

		N[1]	Ac. Média[2]	Acc. Treino[3]	N[5]	Ac. Média[2]	Acc. Validação[4]
PTO	REC1	960	0,48	0,61(0,11)[ab]	1120	0,48	0,37(0,07)[a]
	REC2	960	0,38	0,51(0,04)[b]	1120	0,38	0,36(0,04)[a]
	REC3	960	0,39	0,54(0,05)[b]	1120	0,39	0,39(0,07)[a]
PO	REC1	960	0,48	0,73(0,02)[a]	1120	0,48	0,45(0,06)[a]
	REC2	960	0,38	0,62(0,05)[ab]	1120	0,38	0,42(0,03)[a]
	REC3	960	0,39	0,61(0,05)[b]	1120	0,39	0,45(0,04)[a]

[1]Número de animais da população de treinamento; [2]Acurácia média dos VGPd; [3]Acurácia das populações de treinamento e respectivos erros-padrão; [4]Acurácia das populações de validação e respectivos erros-padrão; [5]Número de animais da população de validação; Médias seguidas da mesma letra, na coluna, não diferem entre si pelo teste Tukey a 5% de probabilidade; PTO: produção total de ovose PO peso dos ovos; REC1, REC2 e REC3: cenários de populações recentes 1, 2 e 3.

5. CONCLUSÕES

As variações na magnitude das estimativas de herdabilidade das características estudadas e os diferentes cenários de acasalamento não interferiram nos valores de acurácia dos VGGP na população de validação. No entanto, nas populações de treinamento, concluiu-se que as diferenças nas acurácias dependem da herdabilidade da característica e também do cenário de acasalamento, fortemente influenciado pelo desequilíbrio de ligação.

O desequilíbrio de ligação entre marcadores SNP é sensível ao aumento das taxas de endogamia, quando marcadores adjacentes estão em curtas distâncias de centimorgan.

6. REFERÊNCIAS

ANDREESCU, C.; AVENDANO, S.; BROWN, S. R.; HASSEN, A.; LAMONT, S. J.; DEKKERS; J. C. M. Linkage Disequilibrium in Related Breeding Lines of Chickens. **Genetics,** v. 177 p. 2161–2169. 2007.

ARDLIE, K. G.; LIU-CORDERO, S. N.; EBERLE, M. A. et al. Lower-thanexpected linkage disequilibrium between tightly linked markers in humans suggests a role for gene conversion**. American Journal of Human Genetics,** v. 69, p. 582-589, 2001.

BOHMANOVA, J.; SARGOLZAEI, M.; SCHENKEL, F. S. Characteristics of linkage disequilibrium in North American Holsteins.**BMC Genomics**, v. 11 p. 421, 2010.

BRITO, F. V.; NETO, J. B.; SARGOLZAEI, M.; COBUCI, J. A.; SCHENKEL, F. S. Accuracy of genomic selection in simulated populations mimicking the extent of linkage disequilibrium in beef cattle. **BMC Genetics**, v. 12, p. 80, 2011.

BRITO, F.V. **Diversidade genética e acurácia da informação genômica em bovinos de corte**. 2011. 134f. Tese (Doutorado) – Universidade Federal Do Rio Grande Do Sul. Faculdade De Agronomia, Porto Alegre, RS, 2011.

CHENG, H. H. Mapping the chicken genome. **Poultry Science**, v. 76 p. 1101-1107, 1997.

COLLINS, A.; LONJOU, C.; MONTON, N. E. Genetic epidemiology of singlenucleotide polymorphisms **Proceedings of the National Academy of Sciences**. v. 96, p. 15173-15177, 1999.

CRUZ, E.M. **Efeito da saturação e do tamanho de populações F2 e de retrocruzamento sobre a acurácia do mapeamento genético**. 2006. 119p. Tese (Doutorado) - Universidade Federal de Viçosa, Viçosa, MG, 2006.

FALCONER, D. S; MACKAY, T. F. C. **IntroductiontoQuantitativeGenetics**. Harlow, Essex: Addison Wesley Longman, 1996.

GARCÍA-GÁMEZ, E.; SAHANA, G.; GUTIÉRREZ-GIL, B.; ARRANZ, J. J. Linkage disequilibrium and inbreeding estimation in Spanish Churra sheep, **BMC Genetics** p. 13:43, 2012

GARRICK, D. J.; TAYLOR, J. F.; FERNANDO, R. L. Deregressing estimated breeding values and weighting information for genomic regression analyses.**Genetics Selection Evolution**, v.41, p. 55. 2009.

GILMOUR, A. R.; GOGEL, B.J.; CULLIS, B.R.; THOMPSON, R., ASReml User Guide Release 3.0 VSN International Ltd, **Hemel Hempstead**, 2009.

HAYES, B.; GODDARD, M. E.The distribution of the effects of genes affecting quantitative traits in livestock.**Genetics Selection Evolution**, v. 33, n.3, p.209-229, 2001.

HEIFETZ, E. M.; FULTON, J. E.; O'SULLIVAN, N.; ZHAO, H.; DEKKERS; J. C. M.; SOLLER, M. Extent and consistency across generations of linkage disequilibrium in commercial layer chicken breeding populations. **Genetics,** v. 171 p. 1173–1181. 2005.

HENDERSON, C. R.; Best linear unbiased estimation and prediction under a selection model.**Biometrics**, v. 31, p. 423-447, 1975.

HENSON, E. L. In situ Conservation of Livestock and Poultry. In: (2nd ed.), **FAOAnimal Production and Health**, 1992.

HILL, W. G.; ROBERTSON A. Linkage disequilibrium in finite populations.**Theoretical and Applied Genetics** v. 38, p.226-231, 1968.

HU, Z.L.; PARK, C. A.; WU, X.L.; REECY, J. M., Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. **Nucleic Acids Research**, v. 41 p. D871-D879, 2013.

LEGARRA, A.; RICARDI, A.; FILANGI, O. In: **GS3: Genomic Selection, Gibbs Sampling, Gauss-Seidel** (and RR-Blup1), 2011.

LEWONTIN, R. C. The detection of linkage disequilibrium in molecular sequence data.**Genetics**, v. 140 p. 377-388, 1995.

MEUWISSEN, T. H. E.; GODDARD, M. E.; HAYES, B. J. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, p. 1819-1829, 2001.

MEUWISSEN, T. H. E.; WOOLIAMS, l. A. Effective population sizes of livestock populations to prevent a decline in fitness.**Theoretical and Applied Genetics**, v. 89, p.1019-1026, 1994

MISZTAL, I.; Tsuruta, T.; Strabel, T.; Auvray, B.; Druet, T.; Lee, D. H., BLUPF90 and related programs (BGF90). In:World Congress on Genetics Applied to Livestock Production,8., 2002, Montpellier. **Anais eletrônicos…** , Montpellier: GALP, 2002. Disponível em: <http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=28-07.pdf> Acesso em: 14 mai. 2013.

MUIR, W. M. The interaction of selection intensity, inbreeding depression, and randon genetic drift on short and long-term response to selection: Results using finite locus and finite population size models incorporating directional dominance. **Journal Animal Science**, v. 79, p. 1-11, 2000.

NCBI, National Center for Biotechnology Information.Disponível em <http://www.ncbi.nlm.nih.gov/genome/111?project_id=10808, 2013> Acessado em 07 mar. 2013.

OLIVEIRA, C. G.; CARNEIRO, P. L. S.; MALHADO, C.H.M; CUNHA, E. E.; EUCLYDES, R. F.; CARNEIRO, A. P. S.; SOUZA, L. G. R. Endogamia em populações de frangos de corte submetidas a diferentes metodologias de seleção. Revista Científica de Produção Animal, v. 8, p. 29-39.

QANBARI, S.; PIMENTEL, E. C. G.; TETENS, J.; THALLER, G.; LICHTNER P.; SHARIFI, A.R.; SIMIANER, H., The pattern of linkage disequilbriom in German Holstein cattle. **Animal Genetics**, v. 41 p. 346-356, 2010.

RESENDE, M. D. V. **Genômica Quantitativa e Seleção no melhoramento de Plantas e Animais**. Colombo: Embrapa Florestas, 2008. p. 21.

RESENDE, M. D. V.; LOPES, P. S.; SILVA, R. L.; PIRES, I. E. Seleção genômica ampla (GWS) e maximização da eficiência do melhoramento genético. **Pesquisa florestal brasileira**, v. 56, p. 63-77, 2008.

ROSA, J. O.; VARGAS, G.; BERNARDES, P. A.; FIGUEIREDO, E. A. P.; LEDUR, M. C.; MUNARI, D. P. Avaliação da endogamia de uma população de aves de postura White Leghorn. In: Simpósio Brasileiro de Melhoramento Animal, 10., 2013.

Uberaba. **Anais eletrônicos...** Uberaba: SBMA, 2013. Disponível em: <http://sbmaonline.org.br/anais/x/trabalhos/pdf/6NHZ.pdf>Acesso em: 01 jul. 2013.

SARGOLZAEI, M.; SCHENKEL, F. S. QMSim: a large-scale genome simulator for livestock. **Bioinformatics**, v. 25, p. 680–681, 2009.

SAVEGNAGO, R. P.; CAETANO, S. L.; RAMOS, S. B.; NASCIMENTO, G. B.; SCHMIDT, G. S.; LEDUR, M. C.; MUNARI, D. P. Estimates of genetic parameters, and cluster and principal components analyses of breeding values related to egg production traits in a White Leghorn population. **Poultry Science**, v. 90, p. 2174–2188, 2011.

SONESSON, A. K.; MEUWISSEN, T. H. E., Mating schemes for optimum contribution selection with constrained rates of inbreeding**.Genetics Selection Evolution**, v. 32, p. 231-248, 2000.

STEPHENS, J. C.; Schneider, J. A.; Tanguay, D. A. et al. Haplotype variation and linkage disequilibrium in 313 human genes. **Science**, v. 293, p. 489-493, 2001.

VAN RADEN, P. M., VAN TASSELL, C. P.; WIGGANS, G. R.; SONSTEGARD, T. S.; SCHNABEL, R. D.; TAYLOR, J. F.; SCHENKEL, F. S. Invited review: Reliability of genomic predictions for North American Holstein bulls. **Journal Dairy Science**, v. 92 p.16–24, 2009.

WRIGHT, S. Evolution in Mendelian Populations. **Genetics**, v. 16, p. 97-159, 1931.