

DELVAN ALVES DA SILVA

**INFLUÊNCIA DA CONEXIDADE E DA ESTRUTURA DOS GRUPOS DE
CONTEMPORÂNEOS NA AVALIAÇÃO GENÉTICA DE BOVINOS
NELORE VIA INFERÊNCIA BAYESIANA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Zootecnia, para obtenção do título de Magister Scientiae.

VIÇOSA
MINAS GERAIS – BRASIL
2015

Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa

T

S586i
2015

Silva, Delvan Alves da, 1990-

Influência da conectividade e da estrutura dos grupos de contemporâneos na avaliação genética de bovinos Nelore via inferência bayesiana / Delvan Alves da Silva. – Viçosa, MG, 2015.

xi, 38f. : il. ; 29 cm.

Inclui anexos.

Orientador: Paulo Sávio Lopes.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f.24-28.

1. Bovino. 2. Ganho genético. 3. Herdabilidade. 4. Peso. 5. Métodos estatísticos. I. Universidade Federal de Viçosa. Departamento de Zootecnia. Programa de Pós-graduação em Zootecnia. II. Título.

CDD 22. ed. 636.20852

DELVAN ALVES DA SILVA

**INFLUÊNCIA DA CONEXIDADE E DA ESTRUTURA DOS GRUPOS DE
CONTEMPORÂNEOS NA AVALIAÇÃO GENÉTICA DE BOVINOS
NELORE VIA INFERÊNCIA BAYESIANA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Zootecnia, para obtenção do título de Magister Scientiae.

APROVADA: 27 de julho de 2015.


Prof. Antônio Policarpo Souza
Carneiro


Prof. Fabyano Fonseca e Silva
(Coorientador)


Prof. Paulo Sávio Lopes
Orientador

Aos meus pais, David e Geralda, meus irmãos Denival
e João Paulo, minha irmã Dalva e a minha
companheira Alessandra e sua família.

Dedico.

“Se eu vi mais longe, foi por estar de pé sobre ombros de gigantes.”

Isaac Newton

AGRADECIMENTOS

A Deus pelo dom da vida e pela força do dia-a-dia para mais essa conquista.

À Universidade Federal de Viçosa – UFV, especialmente ao Departamento de Zootecnia, pela estrutura e oportunidade de realização deste curso. Ao Cluster UFV pela estrutura computacional para análise dos dados.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de estudos.

À Associação Brasileira dos Criadores de Zebu (ABCZ), pelo fornecimento dos dados para realização deste trabalho.

Ao Professor Paulo Sávio Lopes, pelos ensinamentos, conselhos, paciência, amizade e orientação. Principalmente pelo apoio e orientação na elaboração deste trabalho.

Ao Dr. Henrique Torres Ventura, pelo apoio e conselhos na execução do trabalho.

Ao Professor Fabyano Fonseca e Silva, pelos ensinamentos, apoio e conselhos para elaboração do trabalho.

Ao Professor Antônio Policarpo Souza Carneiro, pela disponibilidade e participação na banca avaliadora.

Ao professor Robledo de Almeida Torres, pelo incentivo e acolhida em Viçosa para realização do estágio.

Ao Laboratório de Biometria da UFV, que forneceu a estrutura física e computacional para a execução das análises, e à FAPEMIG por proporcionar recursos financeiros para o estabelecimento da estrutura computacional. Ao professor Leonardo Bhering pelo apoio.

Ao corpo docente e técnicos do curso de Zootecnia que participaram de forma direta ou indireta na minha formação.

À minha família, minha Mãe Geralda e ao meu Pai David, meus irmãos Denival, Dalva e João Paulo, por todo apoio, suporte, compreensão e carinho durante todo o curso, principalmente nos momentos mais difíceis.

À Alessandra Alves da Silva pela amizade, cumplicidade e companheirismo que forneceu apoio e amparo durante esse período, por tudo que conquistamos juntos e pelos momentos de felicidades que vivemos e compartilhamos.

A seu Adail, Dona Joana, Adail Junior, Mário, Rosimary, Victor e Janete pela amizade criada e pelo apoio.

Às amigas Denise, Patrícia e Nayanne, pelo apoio. E todos que de forma direta ou indireta torceram e incentivaram durante o curso.

Ao Vinícius Silva Junqueira, pelo apoio e ensinamentos para realização deste trabalho, principalmente pela paciência e disponibilidade.

Aos amigos Giovani Caetano e Rosa, pelo apoio e incentivo durante o curso.

Aos amigos da pós-graduação Nadson Souza, André Mauric, Sirlene Lázaro, Laís Costa, Hinayah Rojas, Hugo Teixeira, Edson Vinícius, Rodrigo Mota, Aline Crispim, Luciano Pinheiro, Jeferson Corrêa, Matilde Pessoa, Felipe Gomes, Renata Veroneze, Daniele Diniz, Leonardo Glória, Rodrigo Godinho, Ederson Camargo, Rogério Veloso, Raphael Rocha e José Teodoro pelas discussões e trocas de conhecimentos.

Aos amigos da graduação Bárbara Félix, William Mariano, Geraldo Iria, Otávio, Mariana Ribeiro, Naila Carvalho, Bia Nicolau, Ana Luiza Vecchia, Anderson Mota, Eduardo Gonzaga, Ricardo Lucas, Vinícius Saltarelli, Jordana Veríssimo, Cléria Monteiro e Paulo Henrique.

BIOGRAFIA

DELVAN ALVES DA SILVA, filho de David dos Santos Silva e Geralda Alves da Silva, nasceu em 29 de Outubro de 1990, na cidade de Juara, Mato Grosso.

Em março de 2009 iniciou o curso de Zootecnia pela Universidade Federal de Mato Grosso, campus de Sinop – MT, dedicando a atividades acadêmicas relacionadas ao melhoramento animal.

Em Julho de 2013 graduou-se em Zootecnia pela Universidade Federal de Mato Grosso.

Em Agosto de 2013 iniciou Mestrado no Programa de Pós-Graduação em Zootecnia da Universidade Federal de Viçosa, Minas Gerais, na área de melhoramento animal.

Em 27 de Julho de 2014, submeteu-se aos exames finais de defesa de dissertação para obtenção do título de Magister Scientiae em Zootecnia.

SUMÁRIO

RESUMO	viii
ABSTRACT	x
INTRODUÇÃO	1
REVISÃO DE LITERATURA	2
Grupos de contemporâneos (GC).....	2
Conexidade de dados.....	4
Metodologias para determinação da conexidade de dados	5
Importância da conexidade de dados na avaliação genética	8
MATERIAL E MÉTODOS	9
RESULTADOS	12
Análises dentro de desvio-padrão	13
Análises entre os níveis de desvios-padrão	14
Porcentagem de animais em comum, acurácias e ganhos genéticos.....	15
DISCUSSÃO	19
CONCLUSÕES	23
REFERÊNCIAS BIBLIOGRAFICAS	24
ANEXOS	29

RESUMO

SILVA, Delvan Alves da, M.Sc., Universidade Federal de Viçosa, Julho de 2015. **Influência da conectividade e da estrutura dos grupos de contemporâneos na avaliação genética de bovinos Nelore via inferência Bayesiana.** Orientador: Paulo Sávio Lopes. Coorientador: Fabyano Fonseca e Silva.

Objetivou-se avaliar a influência da estrutura de grupos de contemporâneos (GC) e da conectividade de dados na avaliação genética de peso ao desmame padronizado para 210 dias de idade em bovinos da raça Nelore. Foram utilizados 713.474 registros de pesos de animais Nelore provenientes de 3.066 rebanhos das regiões Centro-Oeste e Norte do Brasil. Os dados fazem parte do Serviço de Registro Genealógico das Raças Zebuínas (SRGRZ) disponibilizados pela Associação Brasileira dos Criadores de Zebu – ABCZ. Os GC foram formados com os efeitos de rebanho, condição de criação, regime alimentar, sexo, data de pesagem, ano e estação de nascimento. Os GC e as classes de idade da vaca ao parto concatenada com sexo da progênie foram definidos como efeitos sistemáticos, a idade do animal na data de pesagem foi considerada como covariável (linear e quadrática). Foram utilizados três níveis de desvios-padrão em relação à média para remoção de outliers (2, 2,5 e 3,5) e GC que continham pelo menos 3, 7 e 15 animais, que possibilitou a combinação de nove diferentes estruturas de dados. A conectividade entre os dados foram mensuradas e avaliadas considerando o número total de ligações genéticas diretas entre GC (NTLGD). Foram consideradas 5, 10 e 15 ligações entre os GC, L5, L10 e L15, respectivamente. Empregou-se o modelo animal unicaracterístico via inferência Bayesiana para estimação dos valores e parâmetros genéticos. A estrutura de dados com 3,5 desvios-padrão e GC com no mínimo 15 animais apresentou maior valor de variância genética aditiva ($82,65 \pm 2,8575$). A estrutura de dados com 2 desvios-padrão e GC com no mínimo 15 animais apresentou menor valor de variância residual ($206,86 \pm 1,451$), e o maior valor foi verificado para estrutura de dados com 3,5 desvios-padrão e GC com no mínimo 3 animais ($276,42 \pm 1,4902$). O maior valor de herdabilidade ($0,18 \pm 0,006$) foi estimado ao considerar a estrutura com 2,5 desvios-padrão e GC com no mínimo 15 animais. As estruturas com 3,5 desvios-padrão e mínimo de 15 animais por GC proporcionaram maiores ganhos genéticos, de 5,04 e 3,34 kg para os 1% dos machos e 20% das fêmeas e para 5% dos machos e 40% das fêmeas, respectivamente. Ao avaliar a conectividade

de dados, não verificou-se diferença entre as estruturas de dados com as diferentes ligações genéticas, sendo a conectividade acima de 99% em todas as estruturas. Ao considerar a conectividade, os valores encontrados para as estimativas dos parâmetros genéticos nas diferentes estruturas de dados foram iguais aos encontrados na avaliação sem considerar a conectividade entre os dados. A alta conectividade entre os GC é reflexo da alta qualidade na estrutura de dados, justificada pela intensa comercialização de sêmen de touros provados e troca de material genético entre os rebanhos. Os diferentes critérios para formação de grupos de contemporâneos influenciaram na avaliação genética. O uso de critérios para remoção de outliers com 2 desvios-padrão e grupos de contemporâneos com mínimo de 3 animais proporcionaram menores estimativas de herdabilidade. O uso de critérios para remoção de outliers com 2,5 e 3,5 desvios-padrão e grupos de contemporâneos com mínimo de 15 animais proporcionaram maiores estimativas de herdabilidade, porém resulta em maior perda de registros. Os dados apresentam boa qualidade na estrutura em função da alta conectividade existente entre os grupos de contemporâneos.

ABSTRACT

SILVA, Delvan Alves da, M.Sc., Universidade Federal de Viçosa, July of 2015. **Contemporary groups connectivity and structure influences in genetic evaluation of Nellore cattle using Bayesian inference.** Adviser: Paulo Sávio Lopes. Co-Adviser: Fabyano Fonseca e Silva.

We aimed to evaluate the influence of contemporary groups (CG) structure and data connectivity in the genetic evaluation for body weight at weaning standardized for 210 days of age in Nellore cattle. A total of 713,474 records from 3,066 herds located at Middle West and North of Brazil were used. These herds are included at Genealogical Records Service of Zebu Breeds (SRGRZ) provided by Brazilian Association of Zebu Breeders – ABCZ. The CG were formed according herd, handling condition, diet, sex, date of weighting, year and birth season effects. Contemporary groups and classes of cow age at calving concatenated with offspring gender were defined as systematic effects, and the animal age at weighting was considered as covariate (linear and quadratic). Three standard deviations from the average (2, 2.5 and 3.5) were used for outliers removal, and CG with at least 3, 7 and 15 animals were included, thus allowing the formation of nine different data structures. Connectivity across the data was measured considering the total direct genetic linkage between CG (NTLGD). Five, ten and fifteen linkages between CG, L5, L10 and L15, respectively, were considered. Genetic evaluation was performed using a single trait animal model under a Bayesian approach to estimate the breeding values and genetic parameters. Data structure with 3.5 standard deviations and CG with at least 15 animals presented higher additive genetic variance values (82.65 ± 2.8575). Data structure with 2 standard deviations and CG with at least 15 animals presented lower residual variance values (206.86 ± 1.451), and the higher value (276.42 ± 1.4902) was verified for the data structure with 3.5 standard deviations and CG with at least 3 animals. The higher heritability value (0.18 ± 0.006) was estimated when considering the structure with 2.5 standard deviations and CG with at least 15 animals. Structures with 3.5 standard deviations and least 15 animals per CG had the highest genetic gains, 5.04 and 3.34 kg for the 1% of the males and 20 % of the females and 5% of the males and 40% of the females, respectively. The evaluation the data connectivity showed no difference between structures with different genetic linkage, resulting in more than 99% of connectivity for all considered structures. Genetic parameters estimates values of different data structures considering the data connectivity were the same as those without consider the data connectivity. The high connectivity between CG is due the good quality of the data structure explained by the intense marketing of proven bulls semen and genetic material exchange between herds. The different criteria for CG formation influenced the data structure and

genetic evaluation. Criteria usage for removing outliers with 2 standard deviations and CG with at least 3 animals had the lowest heritability estimates. The criteria used for removing outliers with 2.5 and 3.5 standard deviation and CG with at least 15 animals had the highest heritability estimates, but lead to a higher records loss. Data presents a good structure quality due the high existing connectivity between CG.

INTRODUÇÃO

Na avaliação genética de bovinos de corte a formação de grupos de contemporâneos (GC) apresenta um paradoxo conceitual quanto ao tamanho do grupo de animais e a definição das condições homogêneas comuns aos mesmos. O agrupamento dos animais que atende a maximização da homogeneidade dentro do grupo remete ao número reduzido de amostra que interfere diretamente na acurácia das diferenças esperadas nas progênies (DEP) e na maior necessidade de eliminação de registro para avaliação genética dos animais (Cobuci et al., 2006). Os erros na identificação dos GC, induzem a tomada de decisões incorretas em função de valores genéticos superestimados dos animais que tiveram melhores condições e subestimados para animais que tiveram poucas condições para expressar seu potencial genético (Ferraz e Eler, 1998).

O número específico de animais para compor um GC ideal é um dos principais dilemas nas avaliações genéticas, visando a abrangência do maior número de animais contemporâneos e a máxima homogeneidade das condições ambientais por grupo de contemporâneos (Oliveira, 1995). Na expectativa de identificar dados discrepantes, como desempenhos diferenciados devido a fatores não genéticos dentro dos GC, são aplicados critérios para remoção de outliers. O desvio-padrão é uma alternativa para remover os outliers. Todavia, não é especificada na literatura os critérios para a remoção de outliers. Assim, os valores de desvios-padrão para remoção de outliers são determinados de forma empírica pelos melhoristas nas avaliações genéticas, sem estudos científicos para determinar a mais adequada escala de desvio-padrão que implicaria maior acurácia dos valores genéticos.

Por outro lado, a acurácia das avaliações genéticas depende, dentre outros fatores, da estrutura dos dados e do grau de conexão entre os grupos de contemporâneos (Carneiro et al., 2001; Mathur et al., 2002; Roso et al., 2004; Fouilloux et al., 2006). Além da acurácia das avaliações genéticas, esses fatores influenciam também na estimação dos parâmetros genéticos que podem ser subestimados ou superestimados, o que reflete na incorreta classificação dos animais (Fouilloux et al., 2006; Fouilloux et al., 2008; Kuehn et al., 2009; Soga et al., 2010) e, conseqüentemente, na resposta à seleção (Lewis e Simm, 2000; Simm et al., 2001; Kuehn et al., 2008). A conexão de dados pode ser medida por meio das relações genéticas dos animais entre os GC; assim, dois GC podem ser conectados por

participação direta ou indireta das ligações genéticas (Fries & Roso, 1997, Roso et al., 2004).

Os objetivos com este trabalho foram avaliar a influência da estrutura de grupos de contemporâneos (considerando 2, 2,5 e 3,5 desvios-padrão para a remoção de outliers e mínimos de 3, 7 e 15 animais por grupos de contemporâneos) e a influência da conexidade de dados (considerando 5, 10 e 15 laços genéticos entre os grupos de contemporâneos) para peso ao desmame padronizado para 210 dias de idade em bovinos da raça Nelore.

REVISÃO DE LITERATURA

Grupos de contemporâneos (GC)

Nas avaliações genéticas de bovinos de corte há uma série de fatores de ambiente identificados como importantes fontes de variação. Entre as principais, pode ser citado rebanho, ano de nascimento, mês de nascimento, tipo de manejo a que estão submetidos os animais, a idade da vaca ao parto e o sexo (Oliveira, 1995).

Neste contexto, por definição, os grupos de contemporâneos (GC) permitem a remoção dos efeitos de ambiente e manejo diferenciados entre os animais, permitindo que os animais expressem seus fenótipos nas mesmas condições ambientais (Oliveira, 1995; Cobuci et al., 2006). Exemplos relacionados na Tabela 1 mostram algumas das formas mais usadas na formação dos GC nas avaliações genéticas de bovinos de corte.

Possíveis erros cometidos na identificação correta dos GC podem levar a tomada de decisões incorretas em função de valores genéticos superestimados dos animais que tiveram melhores condições e subestimados para os animais que tiveram poucas condições para expressar seu potencial genético (Ferraz e Eler, 1998).

Os grupos de contemporâneos são formados pelos responsáveis pelas avaliações genéticas a partir das informações obtidas em campo, após análises preliminares dos dados. Embora seja de responsabilidade do melhorista, a pessoa responsável pelas decisões do manejo na propriedade pode facilitar e aumentar a eficiência da formação dos grupos de contemporâneos, identificando corretamente os grupos de animais que são submetidos ao mesmo manejo ou que estão colocados em um mesmo lote desde o nascimento até o momento que foram coletadas as medidas (Oliveira, 1995).

Tabela 1. Algumas formas mais utilizadas para o agrupamento de animais contemporâneos nas avaliações genéticas de bovinos de corte

Formas de Agrupamento
Ano e Mês de nascimento
Ano e Estação de nascimento – Ano de reprodução
Ano e Mês de nascimento – Ano de reprodução
Rebanho – Ano de reprodução – Estação de monta – Grupo de manejo
Sexo – Ano de nascimento – Grupo de manejo
Sexo – Mês de nascimento – Grupo de manejo
Ano e Mês de nascimento – Sexo – Ordem de parto
Rebanho – Ano e Mês de nascimento – Grupo de manejo – Pasto
Ano e Estação de nascimento – Sexo
Sexo – Estação – Criador – Rebanho – Código alimentar
Rebanho – Grupo de manejo ao nascimento – Grupo de manejo ao desmame – Grupo de manejo ao sobreano
Ano – Grupo de manejo – Rebanho de nascimento e/ou do desmame

Fonte: Cobuci et al. (2006)

Porém, no contexto da definição de GC existe um paradoxo conceitual quanto ao tamanho do grupo de animais e as condições homogêneas comum para os animais. O agrupamento com a redução do intervalo de tempo, como exemplo, atende a maximização da homogeneidade dentro do grupo, mas remete ao número reduzido de amostra que interfere na acurácia das diferenças esperadas nas progênes (DEP) e maior necessidade de eliminação de registro para avaliação genética dos animais (Cobuci et al., 2006).

A eliminação de dados ocorre devido a formação incorreta dos GC relacionados ao número de animais contemporâneos, com animais isolados por não serem contemporâneos aos animais pertencentes a determinado grupo. Esses dados não devem constar nas análises e quando constam não fornecem informações úteis para o progresso genético, invalidando os esforços praticados por associações e produtores. Isto reforça a necessidade dos devidos cuidados para a perfeita formação dos GC (Cobuci et al., 2006).

Predições de DEP com maior acurácia podem ser alcançadas com maior número de animais compondo os GC, ou seja, em função da maior amostragem da população (Oliveira, 1995). Portanto, conforme descrito por Cobuci et al. (2006), não existe uma fórmula exata para definir a divisão dos grupos que fornecerá o melhor resultado, mas deve-se buscar um ponto de equilíbrio.

Assim, na expectativa de detectar os fenótipos extremos que não condizem com o intervalo aceito sob aspecto biológico e que podem causar viés na predição dos

valores genéticos, diversos estudos (Tabela 2) utilizam de medidas de dispersão como desvio-padrão para identificar e eliminar os outliers. Entretanto, não há na literatura um valor de consenso quanto ao desvio-padrão a ser utilizado para garantir a exclusão dos outliers e a interferência na estrutura dos dados.

Tabela 2. Trabalhos disponíveis na literatura com os respectivos valores de desvios-padrão para remoção de outliers e número mínimo de animais por grupo de contemporâneos

Trabalhos	Raça	DP	N/GC
Boligon et al. (2006)	Nelore	2,5	5
Yooko et al. (2007)	Nelore	3	8
Boligon et al. (2009)	Nelore	3	4
Boligon et al. (2010)	Nelore	3	4
Yooko et al. (2010)	Nelore	3	8
Pedrosa et al. (2010)	Nelore	3	3
Bignardi et al. (2011)	Nelore	3,5	9 e 6
Laureano et al. (2011)	Nelore	-	4
Faria et al. (2011)	Nelore	-	4
Shiotsuki et al. (2012)	Nelore	3,5	3
Santos et al. (2012)	Nelore	3	4
Boligon et al. (2013)	Nelore	3	4
Matos et al. (2013)	Nelore	3	7
Oliveira Júnior et al. (2013)	Nelore	-	20
Ferriani et al. (2013)	Nelore	3,5	4
Lopes et al. (2013)	Nelore	-	3
Silva et al. (2013)	Nelore	-	9
Ambrosini et al. (2014)	Nelore	3	5

DP: desvio-padrão utilizado para remoção de outliers; N/GC: número mínimo de animais por grupos de contemporâneos.

Conexidade de dados

A característica fundamental de dados desconectados é que cada um dos seus conjuntos de dados desconexos pode ser analisado separadamente dos demais conjuntos. Cada um tem as suas próprias equações normais que podem ser resolvidas sem referência aos demais grupos. Desconexão de dados significa não só que cada conjunto desconectado de dados pode ser analisado separadamente, mas também, que todos conjuntos não podem serem analisados como um único grupo de dados (Searle, 1971).

Em modelos mistos, se não for atendida a pressuposição de que as esperanças dos efeitos aleatórios dentro dos níveis de efeitos fixos sejam nulas, as comparações entre os efeitos aleatórios em diferentes níveis de efeitos fixos serão viesadas. Neste caso, a conexão entre os dados é definida em termos da correlação entre os efeitos aleatórios relativos às observações efetuadas em diferentes níveis de efeitos fixos. Se

a correlação existe, novamente as comparações são passíveis de predição não-viesada, mas a variância do erro de predição (PEV) destas comparações é afetada pelo grau de conexão dos dados. Nos modelos fixos, a conexão assume caráter descontínuo, ou seja, os dados são conectados (1) ou não (0), enquanto nos modelos mistos a conexão assume caráter contínuo, variando de 0 a 100%, (Martins, 1994).

A conexão de dados pode ser definida como uma medida estatística por influenciar diretamente a acurácia de predição. A relação genética existente entre os grupos de animais avaliados é determinante no aumento do grau de conexão. Mas se dois grupos de animais não tem relação genética, podem ser conectados se for enviado material genético de ambos os grupos para outros grupos que compartilham as mesmas condições de ambiente, o que aumenta o grau de conexão entre os rebanhos (Mathur et al., 1999).

As avaliações genéticas em bovinos de corte e de leite envolve grande volume de dados, obtidos de vários sistemas de produção, o que torna o controle sobre o número de observações nas subclasses de efeitos fixos praticamente impossível. Quando se trabalha com grande conjunto de dados, pouca importância é dada para determinar a conexão dos dados antes de se efetuar as análises dos dados (Carneiro et al., 2001).

Metodologias para determinação da conexão de dados

Em algumas avaliações, a determinação da conexão é exigida por razões estatísticas. A identificação dos subconjuntos de dados conectados permite definir quais os contrastes entre os fatores são estimáveis. Em outros casos, a determinação da conexão é exigida por razões computacionais. Programas computacionais que utilizam os métodos de quadrados mínimos e da máxima verossimilhança não podem ser executados, a menos que os dados sejam conectados (Fernando et al., 1983).

Dentre as metodologias para quantificar os níveis de conexão entre os GC descritas na literatura destacam-se: Médias das relações genéticas (MRGA) (Banos & Cady, 1988); Índice de conexão (IC) e Índice geral de conexão (IGC) (Foulley et al., 1992); Coeficiente de determinação (r^2) (Laloë, 1993); PEV das diferenças em valores genéticos entre os animais (PEV(x)), Fluxo de genes (FG), Variância da deriva genética (VDG) e Variância das diferenças entre os efeitos de GC (Kennedy & Trus, 1993); Número total de laços genéticos diretos entre GC (NTLGD) (Fries & Roso,

1997); Grau de conexidade (GrC) (Mathur et al., 2002); Correlação entre as PEV dos valores genéticos (r_{ij}) (Lewis e Simm, 2000).

Weeks e Williams (1964) descreveram um procedimento para identificar subgrupos de dados conectados para modelos com n fatores classificatórios sem interação. A aplicação direta do método pode ser complicada quando o número de classes de fatores for grande, além de não quantificar a conexidade de dados. Petersen (1978) desenvolveu um teste para conexão em uma classificação de duas vias. No entanto, o método envolve a absorção de um dos dois fatores e de trabalho com elementos da matriz dos coeficientes após absorção. Este trabalho apresenta um algoritmo para determinar conexidade em uma tabela de dupla entrada. Suas características são: 1) a identificação de todos os dados conectados nos subconjuntos de uma classificação em dois sentidos, 2) o método é mais simples computacionalmente, pois não é necessária a absorção de efeitos principais, e 3) o armazenamento é limitado a dois vetores de modo que o algoritmo pode ser aplicado mesmo quando o número de classes de cada um dos fatores é grande.

Fernando et al. (1983) apresentaram método computacional para identificar todos os subconjuntos conectados para modelo com dois fatores classificatórios sem interação. O algoritmo não requer a absorção de nenhum fator e nem o armazenamento de matrizes. No entanto, com um modelo fixo de dois sentidos, sem interação, o algoritmo pode ser usado para identificar contrastes estimáveis. Todas as classes que envolvem contrastes dos fatores contidos em um subconjunto conectado são estimáveis. Uma outra aplicação do algoritmo consiste em identificar as dependências lineares nas equações normais.

Um método geométrico simples para determinar quando os dados são conectados ou não, para modelos com dois fatores fixos quaisquer, foi apresentado por Searle (1987) considerando o modelo de reprodutor. Se a distribuição das progênes dos reprodutores entre os rebanhos onde as observações são feitas e se todas as caselas reprodutor-rebanho são preenchidas, ou seja, as caselas que contenham observações puderem ser ligadas por uma linha continua formada por segmentos na horizontal e vertical somente, que mudam de direção apenas em caselas preenchidas, então este conjunto de dados está conectado, sendo estimáveis as comparações entre pares de reprodutores e rebanhos. Este quadro de distribuição dos dados representa na verdade a matriz $X'Z$ (X e Z são matrizes de incidência dos efeitos de rebanho e reprodutor, respectivamente), que pode ser obtida usando um algoritmo computacional simples

para conectividade geométrica, proposto por Fernando et al. (1983). Este método, como os demais métodos para modelos fixos, apenas determina se os dados são conectados, não medindo o grau de conectividade para um conjunto de dados. O número de progênies por classe reprodutor-rebanho também não afeta a conectividade, uma vez que uma única observação por casela é suficiente para gerar a ligação entre rebanhos.

Foulley et al. (1992) propuseram duas formas alternativas para medir a conectividade. Na primeira, o grau de conectividade é dado pela razão direta entre as PEV dos valores genéticos, utilizando modelos com e sem efeitos fixos. Na segunda, o grau de conectividade é dado pela raiz de ordem n ($n =$ número de indivíduos), da razão direta entre os determinantes das matrizes das PEV dos valores genéticos, utilizando modelos com e sem efeitos fixos. Estes métodos apresentam correlação quase perfeita com a média das PEV das comparações entre efeitos aleatórios em diferentes níveis de efeitos fixos. Porém, eles não resolvem o problema computacional relacionado ao grande volume de dados, pois os dois métodos exigem a obtenção da PEV dos valores genéticos.

Kennedy e Trus (1993) também propuseram três outros métodos, quando a matriz de PEV não pode ser calculada. O primeiro, o método de fluxo de gene, marca a transferência de genes de um GC para outro por meio da formação de grupos com base em animais que originaram os GC. Este tipo de agrupamento é bastante viável em algumas situações. O método é bem adequado para medir, de forma rápida e simples, por exemplo, conexões entre as regiões de um país ou países diferentes. Uma medida de conexão pode ser fornecida pela multiplicação das matrizes $\mathbf{X'ZTQ}$, em que \mathbf{X} e \mathbf{Z} são matrizes de incidência para as unidades de manejo e dos animais, respectivamente, \mathbf{Q} identifica população base em relação à sua unidade de manejo de origem, e \mathbf{T} é uma matriz triangular inferior, que traça a fluxo de genes a partir de uma geração para a seguinte.

O segundo método de Kennedy e Trus (1993) é a variância da deriva genética, que é uma medida mais precisa e versátil de conexão, obtidas através de $\mathbf{X'ZAZ'X}$, mas a um custo computacional aumentado. Este método não requer a identificação da população base ou ancestrais com qualquer unidade de manejo particular. Quando as comparações são em todas as unidades de manejo a relação genética dentro da unidade de manejo aumenta a PEV e a relação genética entre as unidades de manejo diminui a PEV. Matrix $\mathbf{X'ZAZ'X}$ mede a soma das relações genéticas dentro e entre as unidades de manejo.

O terceiro método dos mesmos autores é a variação de estimativas dos efeitos da unidade de manejo e, como indicado anteriormente, a variação de estimativas de diferenças de efeitos de unidades de manejo afeta a PEV de comparação entre animais de diferentes unidades de manejo. Se as unidades de manejo não forem conectadas, os elementos fora da diagonal da matriz de (co)variância serão iguais a zero. Os elementos positivos fora da diagonal são o resultado de relações genéticas entre os animais nas diferentes unidades de manejo, ou conexão genética.

Importância da conexão de dados na avaliação genética

A baixa conexão genética é um problema para a estrutura dos dados. Este problema ocorre quando as populações são isoladas ou semi-isoladas geneticamente em razão de pequena troca de material genético entre populações. O uso de rebanhos fechados, a pequena troca de animais entre rebanhos, a baixa intensidade do uso de inseminação artificial e o uso de matrizes de parentesco esparsas são algumas das principais causas da baixa conexão de dados (Carneiro et al., 2001).

Kennedy e Trus (1993) afirmaram que a conexão de dados é motivo de preocupação para os criadores na tomada de decisões na seleção entre os animais em diferentes ambientes. Algumas medidas relativamente simples de conexão que foram propostas e testadas que podem ser bastante úteis, especialmente para medir a conexão entre as diferentes regiões de um país ou de países diferentes.

Os efeitos aleatórios, como os valores genéticos dos animais e os efeitos fixos, como GC, podem formar dados desconectados e definem ainda as diferenças entre os valores genéticos dos animais em um conjunto de dados. No entanto, a falta ou a existência fracas ligações genéticas afeta a predição dos erros das comparações. Ligações genéticas entre os GC podem reduzir a PEV de comparações entre animais de GC diferentes, porque a PEV de comparações entre animais relacionados são menores do que aquelas entre animais independentes (Kennedy e Trus, 1993).

O efeito materno também pode ser avaliado melhor se uma alta correlação genética com efeitos diretos existe. Como resultado destas ligações genéticas os criadores podem identificar reprodutores que são potencialmente melhores em outros rebanhos, especialmente para características com herdabilidades moderadas e altas. Os valores genéticos dos animais provenientes de rebanhos conectados devem ser publicados em um "nível racial", enquanto que os valores genéticos de animais provenientes de grupos desconectados devem ser usados somente dentro de rebanhos

ou dotados de um aviso que as comparações entre rebanhos mal conectados podem ser viesadas (Tarrés et al., 2010).

O grau de conexão dos dados influencia a acurácia das avaliações genéticas que envolvem diferentes rebanhos, regiões ou países. A baixa conexão reduz a acurácia das estimativas dos efeitos fixos, afetando as comparações entre valores genéticos de animais pertencentes a diferentes níveis de efeitos fixos, uma vez que estas comparações envolvem as estimativas das diferenças entre os efeitos fixos. Enfim, o problema básico da baixa conexão dos dados é que as funções lineares de efeitos fixos não são estimáveis, ou são viesadas, e a predição de efeitos aleatórios é de baixa acurácia (Carneiro et al., 2001).

MATERIAL E MÉTODOS

Foram utilizados 713.474 registros de pesos de animais Nelore com idades entre 165 à 255 dias, padronizados para 210 dias de idade (P210), coletados entre os anos 1993 a 2014, oriundos de 3066 rebanhos das regiões Centro-Oeste e Norte do Brasil, incluídos no Serviço de Registro Genealógico das Raças Zebuínas (SRGRZ) disponibilizados pela Associação Brasileira dos Criadores de Zebu – ABCZ. Os dados foram editados e analisados utilizando o pacote estatístico do R (Versão 3.1.3).

Os critérios utilizados para eliminar fenótipos do arquivo de dados foram: animais com identificação semelhante aos pais, animais em que os pais e as mães tinham idade inferior à 609 dias no nascimento, animais em que as mães tinham idade superior a 9126 dias ao nascimento dos filhos, irmãos nascidos no mesmo ano com intervalo de idade inferior a 315 dias, e animais provenientes de transferência de embrião (TE) e fertilização *in vitro* (FIV).

Os meses de nascimento foram agrupados em 4 estações, a fim de tornarem mais homogêneos os efeitos ambientais condicionados às diferenças climáticas no decorrer de um ano (estação 1, de novembro a janeiro, 2, de fevereiro a abril, 3, de maio a julho e 4, de agosto a outubro). Foram consideradas três condições de criação (CC) caracterizadas por animais lactentes, desmamados e participantes de provas de ganho de peso. Como regime alimentar (RA) foram considerados animais a pasto criados de forma extensiva, semiconfinados e confinados. A idade da vaca ao parto foi agrupada em 14 classes, tal que a classe 1 corresponde a vacas com idades ao parto inferiores a 30 meses de idade, as classes subsequentes foram formadas considerando

intervalos a cada 12 meses e a classe 14 com vacas de idade ao parto superior a 174 meses.

Os grupos de contemporâneos (GC) foram formados pelos efeitos de rebanho, condição de criação, regime alimentar, sexo, data de pesagem, ano e estação de nascimento. Os GC e as classes de idade da vaca ao parto concatenada com sexo da progênie foram definidos como efeitos sistemáticos e a idade do animal na data de pesagem foi considerada como covariável (linear e quadrática).

Para o presente estudo, foram aplicadas diferentes critérios para remoção de outliers dentro de cada GC, aplicando três níveis de desvios-padrão em relação à média (2, 2,5 e 3,5) e eliminando GC que continham menos de 3, 7 e 15 animais. Dessa forma, foi possível a combinação de nove diferentes arquivos de dados (Tabela 3). O número de observações nos arquivos de dados, de animais no arquivo de pedigree, de grupos de contemporâneos, de touros, de vacas e as médias dos P210 de cada estrutura de dados estão descritos na Tabela 4.

Tabela 3. Descrição das estruturas (E) dos dados analisados de acordo com os desvios-padrão para remoção de outliers em relação à média e ao número mínimo de animais por grupo de contemporâneos (GC)

Estrutura	Desvio-Padrão	Mínimo de animais por GC
E1	2	3
E2	2	7
E3	2	15
E4	2,5	3
E5	2,5	7
E6	2,5	15
E7	3,5	3
E8	3,5	7
E9	3,5	15

A conexão entre os dados foi mensurada para cada estrutura de dados definida anteriormente em função do número total de ligações genéticas diretas entre GC. As ligações genéticas diretas entre GC em função dos touros e vacas comuns entre os mesmos (Fries & Roso, 1997, Roso et al., 2004) foram calculadas utilizando o programa AMC proposto por Roso & Schenkel (2006). Foram considerados 5, 10 e 15 laços genéticos entre os GC (L5, L10 e L15, respectivamente) e foram excluídos os GC que não estavam conectados.

A avaliação genética foi realizada utilizando um modelo animal unicaracterístico para obtenção dos valores genéticos, descrito matricialmente como:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{m} + \mathbf{Z}_3\mathbf{pm} + \mathbf{e},$$

em que \mathbf{y} é o vetor de observações dos indivíduos, \mathbf{b} é o vetor de efeitos sistemáticos, \mathbf{a} é o vetor de efeitos aleatórios genéticos aditivos diretos, \mathbf{m} é o vetor de efeitos aleatórios genéticos aditivos maternos, \mathbf{pm} é o vetor de efeitos aleatórios de ambiente permanente materno e \mathbf{e} é o vetor de efeito residual; \mathbf{X} , \mathbf{Z}_1 , \mathbf{Z}_2 e \mathbf{Z}_3 são matrizes de incidências que relacionam as observações aos efeitos sistemáticos, genéticos aditivos diretos, genéticos aditivos maternos e de ambiente permanente materno, respectivamente.

A distribuição condicional de \mathbf{y} dados os parâmetros foi assumida como:

$$\mathbf{y} | \mathbf{b}, \mathbf{a}, \mathbf{m}, \mathbf{pm}, \mathbf{G}, \sigma_{pm}^2, \sigma_e^2 \sim \mathbf{N}(\mathbf{Xb} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{m} + \mathbf{Z}_3\mathbf{pm}, \mathbf{R}),$$

sendo a matriz de (co)variância residual dada por $\mathbf{R} = \mathbf{I}_n * \sigma_e^2$.

Para o efeito sistemático assumiu-se: $\mathbf{b} \sim \mathbf{N}(\mathbf{b}_0, \mathbf{V}_0)$, em que \mathbf{V}_0 é uma matriz diagonal da variância a priori de \mathbf{b} , assumindo $\mathbf{V}_0 \rightarrow \infty$.

Os efeitos genéticos aditivos diretos e maternos foram assumidos como:

$$\begin{bmatrix} \mathbf{a} \\ \mathbf{m} \end{bmatrix} | \mathbf{G}_0 \sim \mathbf{N}(\mathbf{0}, \mathbf{G});$$

$$\mathbf{G} = \mathbf{A} \otimes \mathbf{G}_0;$$

$$\mathbf{G}_0 = \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_m^2 \end{bmatrix}$$

Para o efeito de ambiente permanente materno assumiu-se:

$$\mathbf{pm} | \sigma_{pm}^2 \sim \mathbf{N}(\mathbf{0}, \mathbf{I}\sigma_{pm}^2)$$

Sendo, \mathbf{A} a matriz dos numeradores dos coeficientes de parentesco de Wright, σ_a^2 a variância genética aditiva direta, σ_m^2 a variância genética aditiva materna e σ_{pm}^2 a variância de ambiente permanente materno.

A distribuição Qui-quadrado escalada invertida foi assumida e com distribuição a priori para σ_{pm}^2 e σ_e^2 com as respectivas densidades.

$$\sigma_e^2 \sim S_e^2 \chi_{v_e}^{-2} \text{ e}$$

$$\sigma_{pm}^2 \sim S_{pm}^2 \chi_{v_{pm}}^{-2}, \text{ em que } S_e^2 = 194 \text{ e } S_{pm}^2 = 60 \text{ são os valores a priori para}$$

as variâncias residuais e de ambiente permanente materno, respectivamente, e $v_e = v_{pm} = 5$ são os graus de confiança nesses valores.

Assumiu-se a distribuição Wishart Invertida escalada para \mathbf{G}_0 , com densidade

$$\mathbf{G}_0 \sim IW_2(\mathbf{\Sigma}, v_a), \text{ tal que } \mathbf{\Sigma} = \begin{bmatrix} 162,50 & 0 \\ 0 & 40 \end{bmatrix} \text{ é uma matriz positiva definida a priori de}$$

(co)variância entre os efeitos genéticos aditivos diretos e maternos e $v_a = 5$ é o grau de confiança nesses valores.

A avaliação genética foi realizada sob enfoque bayesiano utilizando Métodos de Monte Carlo via Cadeias de Markov (MCMC). Utilizou-se o programa GIBBS2F90 (Misztal, 2014), com um total de 250.000 iterações com descarte inicial de 20.000 iterações para o período de aquecimento. As amostras foram armazenadas a cada 4 ciclos ($\text{thin} = 4$) gerando um total de 57.500 amostras. As análises de convergência dos componentes de variância e parâmetros genéticos foram realizadas por meio da avaliação gráfica (ANEXO A).

O ganho genético (ΔG) para as diferentes estruturas de dados foi estimado a partir do produto entre a herdabilidade (h^2) e o diferencial de seleção (ΔS), sendo que ΔS é a diferença entre a média dos animais selecionados e a média da população, considerando a seleção de 1% dos melhores machos e 20% das melhores fêmeas (Top 1_20) e seleção de 5% dos melhores machos e 40% das melhores fêmeas (Top 5_40).

Para avaliação das estruturas de dados, foi considerada a porcentagem de animais em comum (coincidência na classificação) em função dos valores genéticos, as acurácias dos valores genéticos, os parâmetros genéticos estimados e os ganhos genéticos.

RESULTADOS

A avaliação de diferentes estruturas de dados baseadas no número mínimo de animais por grupo de contemporâneos (GC) e no desvio-padrão do fenótipo proporcionou arquivos de dados de tamanhos variados. As análises dentro de cada nível de desvio-padrão permite avaliar os efeitos do número mínimo de animais para formação dos grupos de contemporâneos. As análises entre os níveis de desvios-padrão, proporciona entender os efeitos de cada desvio-padrão utilizado sobre as estruturas dos dados. Para inferências estatísticas sobre os resultados foram consideradas as sobreposições dos intervalos de credibilidade (IC) entre as estruturas de dados obtidas nas análises. Assim, a sobreposição dos IC permite inferir se houve semelhança entre as estruturas avaliadas.

As estruturas de dados com 3,5 desvios-padrão (E7, E8 e E9) apresentaram maior número de observações, quando comparadas ao uso de 2 e 2,5 desvios padrão,

com redução média de 3,81% e 1,05% no número de observações, respectivamente. Houve alteração no número de animais nos arquivos de pedigree, com redução média de 2,7% e 0,74%, ao se comparar estruturas de 3,5 com 2 e 2,5 desvios-padrão, respectivamente. Em relação ao número de GC, não houve mudanças entre os desvios-padrão considerados para remoção de outliers.

Ao considerar GC com o mínimo de 15 animais (E3, E6 e E9), verifica-se redução no número de animais em média 25,9% e 37,11% em relação aos GC com mínimo 7 e 3 animais, respectivamente. Verifica-se também uma redução expressiva no número de animais nos arquivos de pedigree e no número de grupos de contemporâneos (Tabela 3) quando se avalia dentro de cada nível de desvio-padrão.

Tabela 4. Número de observações (N), de grupos de contemporâneos (GC), de animais no arquivo de pedigree (Ped), estimativas de médias e desvios-padrão para peso aos 210 dias nas diferentes estruturas de dados

Estrutura*	N	Ped	Touros	Vacas	GC	Média	DP
E1 (2 – 3)	607658	1079821	29686	508738	49248	184,100	30,9850
E2 (2 – 7)	513894	944924	27086	455230	26479	183,600	28,8714
E3 (2 – 15)	380120	739149	23225	369193	12390	183,800	28,1548
E4 (2,5 – 3)	622915	1098412	29789	514226	49248	183,800	31,2111
E5 (2,5 – 7)	529133	964501	27228	461535	26479	183,300	29,8284
E6 (2,5 – 15)	392031	755646	23376	375017	12390	183,500	29,1947
E7 (3,5 – 3)	628328	1105063	29813	516219	49248	183,600	31,6591
E8 (3,5 – 7)	534546	971442	27262	463760	26479	183,000	30,3874
E9 (3,5 – 15)	397066	762644	23425	377532	12390	183,100	29,9182

* Entre parênteses (desvio-padrão – N° mínimo de animais por GC); E_i = Diferentes estruturas de dados; Touros = número de touros em cada estrutura de dados; Vacas = número de vacas em cada estrutura de dados; DP = desvio padrão; MIN = peso mínimo; MAX = peso máximo;

Análises dentro de desvio-padrão

Nota-se que não houve diferença estatística dentro de desvio-padrão ao comparar a sobreposição dos intervalos de credibilidade (IC), entre as variâncias genética aditiva direta (Figura 1A) e materna (Figura 1B).

Não se verifica diferença estatística (pela sobreposição dos IC) entre as estruturas dentro dos desvios-padrão 2,5 (E4, E5 e E6) e 3,5 (E7, E8, e E9) para variância de ambiente permanente materno (Figura 1C). Entretanto, dentro do desvio-padrão 2, as estruturas E3 e E1 foram semelhantes a E2, mas diferiram entre si.

Houve redução acentuada (com diferença estatística ao avaliar a sobreposição dos IC) nas variâncias residuais (Figura 1D) dentro de cada nível de desvio-padrão com o aumento do número mínimo de animais por GC. Para as estruturas com 2

desvios-padrão, as reduções nas variâncias residuais foram de 6,24% e 10,41% ao considerar os aumentos de 3 para 7 e de 3 para 15 animais nos GC, respectivamente. Ao considerar as estruturas com 2,5 desvios-padrão, as reduções nas variâncias residuais foram de 4,34% e 8,48%, respectivamente. Para as estruturas com 3,5 desvios-padrão, as reduções nas variâncias residuais foram de 3% e 5,49%, respectivamente.

Houve redução significativa nas variâncias fenotípicas (Figura 1E) entre as estruturas dentro dos desvios-padrão 2 e 2,5 com o aumento do número mínimo de animais por GC. Porém, as estruturas com 3 desvios-padrão foram semelhantes.

Não se verifica diferença estatística (pela sobreposição dos IC) para os valores de herdabilidade (h^2) (Figura 1F), herdabilidade materna (h_m^2) (Figura 1G) e para coeficiente da relação da variância de ambiente permanente materno com a variância fenotípica (pm^2) (Figura 1H).

Análises entre os níveis de desvios-padrão

Ao avaliar a variância genética aditiva direta entre os desvios-padrão considerados, valores de maior magnitude (com diferença significativa pela sobreposição dos IC) são observados ao considerar maiores desvios-padrão (2,5 e 3,5) em relação à 2 desvios-padrão. As variâncias genética aditiva direta e materna (Figuras 1A e 1B) apresentaram maior magnitude ao aplicar 3,5 desvios-padrão e GC com no mínimo 15 animais (E9).

Constata-se aumento acentuado (com diferença significativa pela sobreposição dos IC) da variância do efeito de ambiente permanente materno entre as estruturas de dados com aumento do nível de desvio-padrão, ou seja, maiores valores de variância do efeito de ambiente permanente materno foram obtidos com 3,5 desvio-padrão (Figura 1C).

Houve aumento significativo (pela sobreposição dos IC) nas variâncias residuais (Figura 1D) quando foram utilizados diferentes desvios-padrão para remoção de outliers. Com o aumento nos níveis de desvios-padrão verifica-se aumento significativo nas variâncias residuais. As variâncias residuais das estruturas com 3,5 desvios-padrão foram 8,7% e 18,85% maiores que as variâncias residuais das estruturas com 2,5 e 2 desvios-padrão, respectivamente. A estrutura E3 apresentou menor valor de variância residual ($206,86 \pm 1,4512$), enquanto que o maior valor foi verificado na estrutura E7 ($276,42 \pm 1,4902$).

Houve aumento significativo (pela sobreposição dos IC) nas variâncias fenotípicas (Figura 1E) quando foram utilizados diferentes desvios-padrão para remoção de outliers. Maiores variâncias fenotípicas foram obtidas com uso de 3,5 desvios-padrão para remoção de outliers em relação aos demais desvios-padrão.

Para os valores de herdabilidade (h^2) (Figura 1F) e herdabilidade materna (h_m^2) (Figura 1G) não se verifica diferença estatística (pela sobreposição dos IC).

Para coeficiente da relação da variância de ambiente permanente materno com a variância fenotípica (pm^2), verificam-se diferenças estatísticas (pela sobreposição dos IC) ao comparar 2 desvios-padrão com 2,5 e 3,5 desvios padrão, porém ao comparar 2,5 e 3,5 desvios-padrão não houve diferença entre as estruturas desses níveis de desvio padrão (Figura 1H).

Porcentagem de animais em comum, acurácias e ganhos genéticos

A porcentagem de animais em comum (coincidências na classificação) entre as diferentes estruturas de dados ao considerar 1% dos melhores machos e 20% das melhores fêmeas (Top 1_20) e 5% dos melhores machos e 40% das melhores fêmeas (Top 5_40) estão descritos na Tabela 5. Em ambas as situações são mantidas hipoteticamente a maior intensidade de seleção para os machos em relação às fêmeas para garantir uma representação das intensidades de seleção utilizadas na prática pelos melhoristas e criadores. Dentro de cada nível de desvio-padrão, os GC com o mínimo 3 e 7 animais tiveram maiores porcentagens de animais em comum com médias de 75,07 e 78,34% para os Top 1_20 e Top 5_40, respectivamente. Verificou-se maiores porcentagens de animais em comum entre as estruturas de dados quando se comparam os níveis de 2,5 e 3,5 desvios-padrão.

As médias das acurácias dos valores genéticos preditos estão expressas na Tabela 6. Os menores valores médios de acurácia foram encontrados para as estruturas de dados com 2 desvios-padrão. Entre as demais estruturas, menores diferenças podem ser notadas, demonstrando que as diferentes estruturas para remoção de outliers não provocam mudanças extremas nas acurácias dos valores genéticos, o que possibilita a escolha da melhor estrutura com base na maior estimativa de herdabilidade e, conseqüentemente, maior ganho genético.

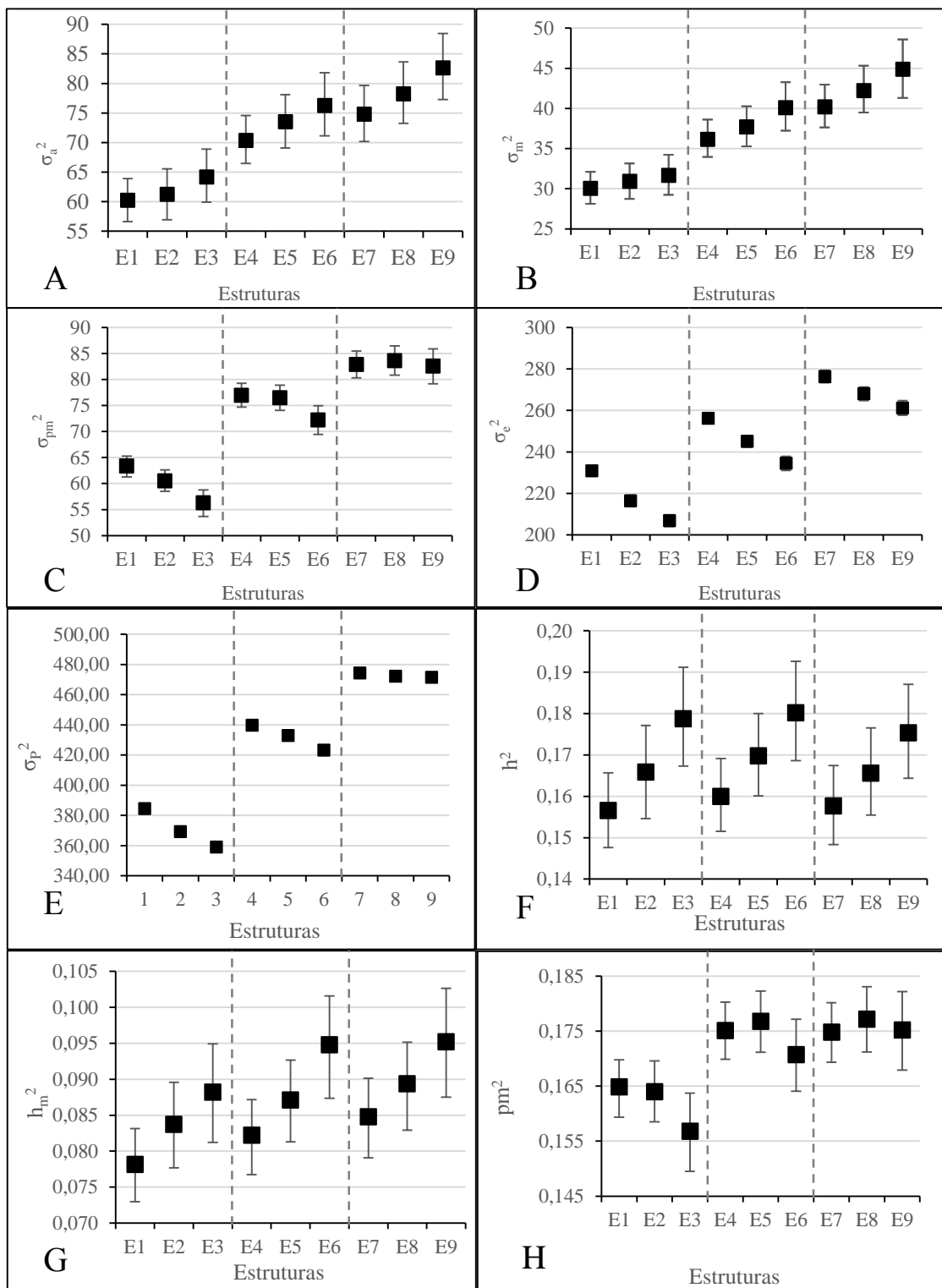


Figura 1. Estimativas a posteriori: A = variância genética aditiva direta (σ_a^2); B = variância genética aditiva materna (σ_m^2); C = variância de ambiente permanente materno (σ_{pm}^2); D = variância residual (σ_e^2); E = variância fenotípica (σ_p^2); F = herdabilidade (h^2); G = herdabilidade materna (h_m^2); H = proporção da variância de ambiente permanente materno em relação à variância fenotípica (pm^2) e os respectivos intervalos de credibilidade nas diferentes estruturas de dados

Tabela 5. Porcentagem de animais em comum entre as diferentes estruturas (E_i) de dados ao selecionar 1% dos melhores machos e 20% das melhores fêmeas (Top 1_20) acima da diagonal principal e 5% dos melhores machos e 40% das melhores fêmeas (Top 5_40) abaixo da diagonal principal

Estrutura*	E1	E2	E3	E4	E5	E6	E7	E8	E9
E1 (2 – 3)	100,00	76,61	56,72	87,55	73,37	55,67	85,25	71,70	54,46
E2 (2 – 7)	79,85	100,00	67,47	72,63	85,89	64,86	71,12	83,22	63,14
E3 (2 – 15)	60,24	68,32	100,00	54,49	63,62	85,08	53,79	62,27	81,62
E4 (2,5 – 3)	90,86	70,23	58,35	100,00	78,22	58,35	91,08	76,33	57,09
E5 (2,5 – 7)	77,60	89,51	67,40	81,32	100,00	68,46	76,26	90,15	66,46
E6 (2,5 – 15)	59,55	68,52	89,13	61,42	71,08	100,00	57,50	66,83	88,93
E7 (3,5 – 3)	89,25	75,73	57,74	93,52	79,74	60,64	100,00	78,67	58,79
E8 (3,5 – 7)	76,51	87,68	66,38	80,01	92,88	69,89	81,66	100,00	68,74
E9 (3,5 – 15)	58,66	67,31	86,57	60,54	69,62	91,90	61,60	71,23	100,00

* Entre parênteses (desvio-padrão – N° mínimo de animais por GC); E_i = Diferentes estruturas de dados; Top 1_20 = 1% dos melhores machos e 20% das melhores fêmeas; Top 5_40 = seleção de 5% dos melhores machos e 40% das melhores fêmeas;

Tabela 6. Valores médios de acurácia dos valores genéticos preditos ao considerar as diferentes estruturas de dados

Estrutura*	Média	DP	Min	Max
E1 (2 – 3)	0,37	0,11	0,01	0,92
E2 (2 – 7)	0,35	0,11	0,01	0,89
E3 (2 – 15)	0,35	0,11	0,01	0,88
E4 (2,5 – 3)	0,40	0,11	0,01	0,91
E5 (2,5 – 7)	0,39	0,11	0,01	0,90
E6 (2,5 – 15)	0,40	0,12	0,01	0,90
E7 (3,5 – 3)	0,41	0,11	0,01	0,92
E8 (3,5 – 3)	0,41	0,11	0,01	0,92
E9 (3,5 – 3)	0,38	0,11	0,01	0,90

* Entre parênteses (desvio-padrão – N° mínimo de animais por GC); E_i = Diferentes estruturas de dados; Média = média das acurácias para cada estrutura de dados; DP = desvios-padrão; Min = acurácia mínima; Max = acurácia máxima;

Na Figura 2, estão descritos os valores de ganhos genéticos para diferentes estruturas de dados para os Top 1_20 e Top 5_40 melhores animais, demonstrando que as estruturas E6 e E9 proporcionaram maiores ganhos genéticos, ambas com 5 e 3,34 kg para os Top 1_20 e Top 5_40, respectivamente. Quando se utilizam as estruturas E6 e E9 são obtidos maiores ganhos genéticos. Ao se comparar com a estrutura E1 (menor ganho genético) as diferenças foram de 840 e 660 g para os Top 1_20 e Top 5_40, respectivamente. Ao comparar as estruturas E6 e E9 com a E3 as diferenças em médias foram de 400 e 310 g para os Top 1_20 e Top 5_40, respectivamente. Quanto menor a proporção de indivíduos selecionados, maiores foram os ganhos genéticos,

em razão da maior intensidade de seleção quando se usa Top 1_20 em vez de Top 5_40.

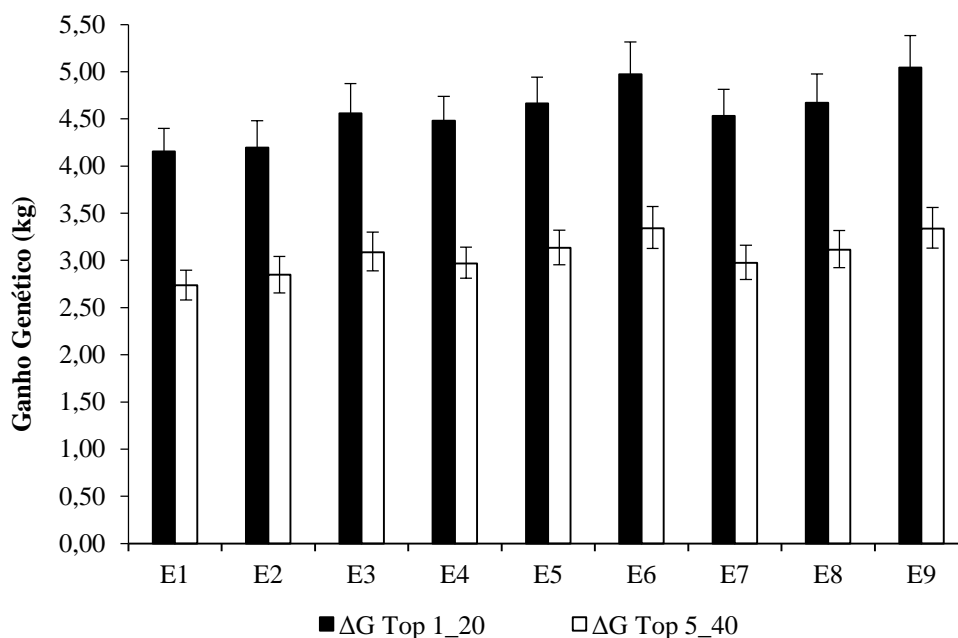


Figura 2. Ganhos genéticos (ΔG) e intervalo de credibilidade das diferentes estruturas (E_i) de dados, considerando 1% dos melhores machos e 20% das melhores fêmeas (Top 1_20) e 5% dos melhores machos e 40% das melhores fêmeas (Top 5_40)

Na Tabela 6 estão descritas as influências do número de ligações genéticas na estrutura dos dados no teste de conexidade de dados. Não se verificaram diferenças entre as estruturas de dados, independente das ligações genéticas avaliadas, sendo a conexidade acima de 99% em todas as estruturas.

Não foi observado influência da conexidade de dados sobre a avaliação genética (dados não apresentados), os valores encontrados para as estimativas dos parâmetros genéticos nas diferentes estruturas de dados foram iguais aos encontrados com a avaliação sem considerar a conexidade entre os dados.

Tabela 7. Número de animais (N), porcentagem de animais conectados (%Nc) e desconectados (%Nd), porcentagem de grupos de contemporâneos (GC), porcentagem de grupos de contemporâneos conectados (%GCc) e desconectados (%GCd), número de animais no arquivo de pedigree (Ped) e porcentagem de animais no arquivo de pedigree conectados (%Pedc) e desconectados (%Pedd) nas diferentes estrutura (E) de dados e ligações genéticas (L)

Estrutura*	Ligação	N	%Nc	%Nd	GC	%GCc	%GCd	Ped	%Pedc	%Pedd
E1 (2 – 3)	L5	607658	99,82	0,18	49248	99,42	0,58	1079821	99,73	0,27
	L10	607658	99,63	0,37	49248	98,82	1,18	1079821	99,50	0,50
	L15	607658	99,45	0,55	49248	98,32	1,68	1079821	99,28	0,72
E2 (2 – 7)	L5	513894	99,93	0,07	26479	99,84	0,16	944924	99,89	0,11
	L10	513894	99,87	0,13	26479	99,70	0,30	944924	99,81	0,19
	L15	513894	99,80	0,20	26479	99,53	0,47	944924	99,71	0,29
E3 (2 – 15)	L5	380120	99,93	0,07	12390	99,86	0,14	739149	99,90	0,10
	L10	380120	99,92	0,08	12390	99,85	0,15	739149	99,89	0,11
	L15	380120	99,91	0,09	12390	99,82	0,18	739149	99,87	0,13
E4 (2,5 – 3)	L5	622915	99,83	0,17	49248	99,42	0,58	1098412	99,74	0,26
	L10	622915	99,64	0,36	49248	98,85	1,15	1098412	99,52	0,48
	L15	622915	99,47	0,53	49248	98,34	1,66	1098412	99,30	0,70
E5 (2,5 – 7)	L5	529133	99,94	0,06	26479	99,85	0,15	964501	99,90	0,10
	L10	529133	99,89	0,11	26479	99,72	0,28	964501	99,83	0,17
	L15	529133	99,81	0,19	26479	99,57	0,43	964501	99,72	0,28
E6 (2,5 – 15)	L5	392031	99,93	0,07	12390	99,87	0,13	755646	99,90	0,10
	L10	392031	99,93	0,07	12390	99,86	0,14	755646	99,89	0,11
	L15	392031	99,91	0,09	12390	99,83	0,17	755646	99,87	0,13
E7 (3,5 – 3)	L5	628328	99,83	0,17	49248	99,43	0,57	1105063	99,74	0,26
	L10	628328	99,65	0,35	49248	98,86	1,14	1105063	99,53	0,47
	L15	628328	99,48	0,52	49248	98,35	1,65	1105063	99,31	0,69
E8 (3,5 – 7)	L5	534546	99,94	0,06	26479	99,85	0,15	971442	99,90	0,10
	L10	534546	99,89	0,11	26479	99,72	0,28	971442	99,83	0,17
	L15	534546	99,82	0,18	26479	99,57	0,43	971442	99,72	0,28
E9 (3,5 – 15)	L5	397066	99,93	0,07	12390	99,87	0,13	762644	99,90	0,10
	L10	397066	99,93	0,07	12390	99,86	0,14	762644	99,89	0,11
	L15	397066	99,91	0,09	12390	99,83	0,17	762644	99,87	0,13

* Entre parênteses (desvio-padrão – N° mínimo de animais por GC); E_i = Diferentes estruturas de dados; L5, L10 e L15 = 5, 10 e 15 laços genéticos diretos entre grupos de contemporâneos, respectivamente.

DISCUSSÃO

O uso desvio-padrão como critério de remoção de outliers nas avaliações genéticas permite identificar e remover registros discrepantes, como animais com desempenhos diferenciados devido a fatores não genéticos dentro dos grupos de contemporâneos (GC), evitando o uso de informações incorretas e ocasionar avaliações genéticas viesadas. Assim, ao considerar os intervalos para remoção de

outliers de 2, 2,5 e 3,5 desvios-padrão sob as condições de normalidade dos dados, tem-se 95,46, 98,76 e 99,96% dos dados, respectivamente. Isso justifica o fato das estruturas com 3,5 desvios-padrão apresentarem maiores informações em relação às demais estruturas de dados. Contudo, a remoção de dados pelos níveis de desvios-padrão é menos expressiva quando comparado com a remoção de dados baseada no mínimo de animais por GC, isso porque todos os animais que estão nos GC com mínimo de animais inferior ao estipulado são removidos.

Observa-se pouca interferência dos níveis de desvios-padrão adotados neste estudo sobre as médias de peso ao desmame (Tabela 4) das diferentes estruturas com diferença de 1,1 Kg. Isso se deve à particularidade de cada estrutura em que a média foi calculada em relação ao conjunto de dados original. As médias de peso ao desmame encontradas neste estudo correspondem às médias encontradas na literatura que variaram de $172,54 \pm 24,24$ à 187 ± 28 Kg (Cyrillo et al., 2004; Yooko et al., 2007; Boligon et al., 2008; Boligon et al., 2010; Yooko et al., 2010; Araújo et al., 2011; Laureano et al., 2011; Faria et al., 2011; Lopes et al., 2013; Araújo et al., 2014).

As estruturas com mínimo de 15 animais por GC dentro de cada nível de desvio-padrão, permite maior captação da variância genética aditiva direta e materna (Figuras 1A e 1B). Em contrapartida, proporciona menores variâncias de ambiente permanente materno e residuais (Figuras 1C e 1D), o que resulta em menores variâncias fenotípicas (Figura 1E) e, conseqüentemente, maiores valores das estimativas de parâmetros e ganhos genéticos. Por outro lado, o uso mínimo de 3 animais por GC proporciona menor captação da variância genética aditiva direta, aumentando a variância residual e reduzindo os valores das estimativas de parâmetros e ganhos genéticos. Portanto, não se recomenda o uso de apenas 3 animais por GC na estruturação dos dados para avaliação genética. O uso de GC com mínimo de 15 animais dentro de cada nível de desvio-padrão resulta em estruturas de dados com menor número de animais em relação às demais estruturas (com mínimos de 3 e 7 animais por GC) e proporciona maior captação da variância genética aditiva e menores efeitos de ambiente como comprovado pela relação da variância residual com a variância fenotípica ($e^2 = \sigma_e^2 / \sigma_p^2$), como demonstrado no ANEXO B.

Ao considerar as estruturas entre os níveis de desvios-padrão (2, 2,5 e 3,5) verifica-se que as variâncias genéticas aditivas direta e materna aumentaram à medida que se aumentam os desvios-padrão. Porém, há também aumento nas variâncias de

ambiente permanente materno e residual, o que mantém os valores de herdabilidade constantes entre os diferentes níveis de desvios-padrão.

A variação no número total de animais permite uma redistribuição das variâncias entre as estruturas. Assim, as estruturas com maiores níveis de desvios-padrão (2,5 e 3,5) proporcionam maiores variâncias genéticas aditivas direta e materna e menores variâncias para os efeitos de ambiente permanente materno e residual. Baseado nessa redistribuição, pode-se inferir que as estruturas com 2 desvios-padrão ou com mínimo de 3 animais por GC não seriam recomendadas para avaliação genética dos animais dessa população.

Maiores valores de herdabilidade (h^2) (Figura 1F) foram estimados nas estruturas de dados E3, E6 e E9 devido principalmente à combinação de maiores variâncias genéticas aditiva e menores variâncias residuais (Figura 1A e 1D) quando comparadas com as demais estruturas. Isso demonstra que ao se utilizar o mínimo de 15 animais por GC nas avaliações genéticas maiores ganhos genéticos são obtidos (Figura 2).

Boligon et al., (2009) e Boligon et al., (2010), ao utilizarem 3 desvios-padrão para remoção de outliers e GC com o mínimo de 4 animais nas avaliações genéticas para peso ao desmame em bovinos da raça Nelore, encontraram valor de h^2 (0,33) superiores aos encontrados em todas as estruturas de dados avaliadas neste estudo. Porém, para os valores de variâncias genéticas aditivas direta e materna, Boligon et al. (2009) encontraram valores superiores ($\sigma_a^2 = 200,00$ e $\sigma_m^2 = 50,00$) aos encontrados neste estudo, mas Boligon et al., (2010) encontraram valor da $\sigma_a^2 = 89,02$, próximo ao valor encontrado na estrutura E9, e $\sigma_m^2 = 29,98$, inferior aos valores encontrados em todas estruturas avaliadas neste estudo.

Yooko et al., (2010), ao utilizarem GC com mínimo de 8 animais e 3 desvios-padrão para remoção de outliers, encontraram valores de h^2 de 0,22 e Silva et al. (2013), ao utilizarem GC com mínimo de 9 animais, encontraram valores de h^2 de 0,13. Os valores de h^2 de ambos os trabalhos, bem como os valores de h^2 encontrados neste estudo são inferiores aos de outros trabalhos encontrados na literatura com diferentes critérios para remoção de outliers, com os valores de h^2 variando de 0,23 à 0,61 (Boligon et al., 2008; Boligon et al., 2009; Boligon et al., 2010; Araújo et al., 2011; Laureano et al., 2011; Lopes et al., 2013; Araújo et al., 2014).

Faria et al. (2011), ao utilizar enfoque Bayesiano e eliminar GC com menos de 4 animais encontraram maiores valores de h^2 (0,33) e de h_m^2 (0,26) quando

comparados aos obtidos neste estudo. Maiores estimativas de h^2 (0,21) e de h_m^2 (0,27) também foram obtidas por Matos et al., (2013) quando comparadas aos obtidos neste estudo, os autores utilizaram 3 desvios-padrão para remoção de outliers e GC com mínimo de 7 animais.

Os valores de h_m^2 encontrados neste estudo foram semelhantes aos valores encontrados por Boligon et al., (2010) que utilizaram 3 desvios-padrão para remoção dos outliers e GC com mínimo de 4 animais. Os valores de h_m^2 encontrados neste estudo, também foram semelhantes aos valores encontrados por Laureano et al., (2011) e Lopes et al. (2013), ambos utilizaram GC com mínimo de 3 animais. Cyrillo et al., (2004), ao utilizar 3 desvios-padrão para remoção de outliers, encontraram $h_m^2=0,04$, entretanto, Silva et al., (2013), ao utilizar o mínimo de 9 animais por GC, encontraram $h_m^2=0,03$, valores inferiores aos encontrados em todas estruturas analisadas neste estudo.

Mesmo que os critérios para remoção de outliers e formação de GC sejam padronizados para as avaliações genéticas, os parâmetros genéticos a serem estimados podem ser diferentes. Isso ocorre em razão dos parâmetros serem inerentes a cada população e as circunstâncias de ambientes às quais os indivíduos avaliados estão inseridos (Falconer e Mackay, 1996). Assim, estimativas de parâmetros genéticos com maiores acurácias poderiam ser encontradas nas avaliações genéticas ao se utilizar critérios para remoção de outliers e formação de GC adequados para cada estrutura de dados.

Ao considerar a seleção 1% dos melhores machos e 20% das melhores fêmeas (Top 1_20) e a seleção de 5% dos melhores machos e 40% das melhores fêmeas (Top 5_40), verificou-se um acréscimo de, em média, 33,27% no ganho genético para os Top 1_20 em relação ao Top 5_40, devido a menor proporção de indivíduos selecionados e, conseqüentemente, maior intensidade de seleção empregada. As limitações no número mínimo de animais por GC influenciam no ganho genético e as estruturas de dados com o mínimo de 15 animais por GC são as mais recomendadas na avaliação genética por proporcionarem maiores ganhos genéticos. Em contrapartida, as estruturas de dados com o mínimo de 3 animais por GC apresentam menores ganhos genéticos. Isso está relacionado ao conceito estatístico de amostragem da população, ou seja, ao considerar maior número de animais por GC, maior será a amostra tomada ao acaso de uma população e assim maior a representatividade dos parâmetros estimados para tal população (Oliveira, 1995).

As estruturas de dados com 2,5 desvios-padrão apresentaram maiores porcentagens de animais em comum com as demais estruturas (2 e 3,5) por serem intermediárias. As estruturas com o mínimo de 7 animais por GC também apresentaram maiores porcentagens de animais em comum com as demais estruturas; conseqüentemente, menores porcentagens de animais em comum foram verificadas entre as estruturas com mínimo de 3 e 15 animais por GC. A porcentagem de animais em comum permite compreender a importância do estudo dos outliers e as formas adotadas para remoção destes, mostrando ainda, que a remoção de outliers utilizando de forma empírica valores de desvios-padrão e número mínimos de animais por GC, sem a investigação profunda, como tem sido empregada, pode influenciar nas avaliações genéticas.

A alta conectividade entre os GC é reflexo da boa qualidade na estrutura de dados. Sendo assim observa-se baixa porcentagem de GC isolados. Pode ser verificado que mesmo entre os laços genéticos considerados o nível de conectividade continua muito alto (Tabela 6). Devido à alta qualidade dos dados, não houve influência da conectividade de dados sobre as avaliações genéticas em relação às estruturas de dados analisadas sem avaliação da conectividade de dados. A justificativa para a boa estrutura dos dados é a comercialização intensa de sêmen de touros provados. Assim, a troca de material genético entre os criadores e a aquisição em comum de um mesmo reprodutor em diversas propriedades favorece essa boa estrutura nos dados e maior conectividade.

CONCLUSÕES

Os diferentes critérios para formação de grupos de contemporâneos influenciaram na estrutura dos dados e na avaliação genética. O uso de critérios para remoção de outliers com 2 desvios-padrão e grupos de contemporâneos com mínimo de 3 animais proporcionaram menores estimativas de parâmetros genéticos. O uso de critérios para remoção de outliers com 2,5 e 3,5 desvios-padrão e grupos de contemporâneos com mínimo de 15 animais proporcionaram maiores estimativas de parâmetros genéticos, porém resulta em maior perda de registros. Os dados apresentam boa qualidade na estrutura em função da alta conectividade existente entre os grupos de contemporâneos.

REFERÊNCIAS BIBLIOGRÁFICAS

- AMBROSINI, D. P.; CARNEIRO, P. L. S.; NETO, J. B.; MARTINS FILHO, R.; AMARAL, R. S.; CARDOSO, F. F.; CARLOS, H. M. M. Reaction norms models in the adjusted weight at 550 days of age for polled Nelore cattle in Northeast Brazil. **Revista Brasileira de Zootecnia**, v. 43, n. 7, p. 351-357, 2014.
- ARAÚJO, C. V.; LÔBO, R. B.; FIGUEREIDO, L.G.G; MOUSQUER, C.J.; LAUREANO, M.M.M.; BITTENCOUR, T. C. B. S. C.; ARAÚJO, S. I. Estimates of genetics parameters of growth traits of Nelore cattle in the Midwest region of Brazil. **Revista Brasileira Saúde Produção Animal**, v. 15, n. 4, p. 846-853, 2014.
- ARAÚJO, C. V.; BITTENCOUR, T. C. B. S. C.; ARAÚJO, S. I.; LÔBO, R. B.; BEZERRA, L. A. F. Estudo da heterogeneidade de variâncias na avaliação genética de bovinos de corte da raça Nelore. **Revista Brasileira de Zootecnia**, v. 40, n. 9, p. 1902-1908, 2011.
- BANOS, G.; CADY, R. A. Genetic relationship between the United States and Canadian Holstein bull populations. **Journal of Dairy Science**, v. 71, n. 5, p. 1346-1354, 1998.
- BIGNARDI, A. B.; GORDO, D. G. M.; ALBUQUERQUE, L. G.; SESANA, J. C. Parâmetros genéticos de escore visual do umbigo em bovinos da raça Nelore. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 63, n. 4, p. 941-947, 2011.
- BOLIGON, A.A.; CARVALHEIRO, R.; ALBUQUERQUE, L.G. Evaluation of mature cow weight: Genetic correlations with traits used in selection indices, correlated responses, and genetic trends in Nelore cattle. **Journal of Animal Science**, v. 91, n. 1, p. 20-28, 2013.
- BOLIGON, A.A.; SILVA, J.A.V.; SESANA, R.C.; JUNQUEIRA, J.B.; ALBUQUERQUE, L.G. Estimation of genetic parameters for body weights, scrotal circumference, and testicular volume measured at diferente ages in Nelore cattle. **Journal of Animal Science**, v. 88, n. 4, p. 1215-1219, 2010.
- BOLIGON, A. A.; ALBUQUERQUE, L. G.; MERCADANTE, M. E. Z.; LOBO, R. B. Herdabilidades e correlações entre pesos do nascimento à idade adulta em rebanhos da raça Nelore. **Revista Brasileira de Zootecnia**, v. 38, n. 12, p. 2320-2326, 2009.
- BOLIGON, A. A.; ALBUQUERQUE, L. G.; RORATO, P. R. N.; Associação genética entre peso e características reprodutivas em rebanhos da raça Nelore. **Revista Brasileira de Zootecnia**, v. 37, n. 4, p. 596-601, 2008.
- BOLIGON, A. A.; RORATO, P. R. N.; WEBER, T.; EVERLING, D. M.; LOPES, J. S. Herdabilidade para ganho de peso da desmama ao sobreano e perímetro escrotal ao sobreano e tendências genética e fenotípica para ganho de peso da desmama ao sobreano em bovinos Nelore-Angus. **Revista Brasileira de Zootecnia**, v.35, n. 4, p. 1323-1328, 2006.
- CARNEIRO, A. P. S.; TORRES, R. A.; EUCLYDES, R. F.; SILVA, M. A.; LOPES, P. S.; CARNEIRO, P. L. S.; TORRES FILHO, R. A. Efeito da conexidade de

- dados sobre a acurácia dos testes de progênie e performance. **Revista Brasileira de Zootecnia**, v.30, n. 2, p. 342-347, 2001.
- COBUCCI, J. A.; ABREU, U. G. P.; TORRES, R. A. **Formação de Grupos de contemporâneos em Bovinos de Corte**. Corumbá, MS: Embrapa Pantanal, 2006. 27p.
- CYRILLO, J. N. S. G.; ALENCAR, M. M.; RAZOOK, A. G.; MERCADANTE, M. E. Z.; FIGUEIREDO, L. A. Modelagem e estimação de parâmetros genéticos e fenotípicos para pesos do nascimento à seleção (378 dias) de machos Nelore. **Revista Brasileira de Zootecnia**, v. 33, n. 6, p. 1405-1415, 2004.
- FALCONER, D.S.; MACKAY, T.F.C. **Introduction to quantitative genetics**. 4.ed. Edinburgh : Longman Group Limited, 1996. 464p.
- FARIA, C.U.; TERRA, J.P.; YOOKO, M.J.I.; MAGNABOSCO, C.U.; ALBUQUERQUE, L.G.; LÔBO, R.B. Interação genótipo-ambiente na análise genética do peso o desmame bovinos Nelore sob enfoque bayesiano. **Acta Scientiarum. Animal Science**, v. 33, n. 2, p. 213-218, 2011.
- FERNANDO, R., GIANOLA D., GROSSMAN, M. Identifying all connected subsets in a two-way classification without interaction. **Journal of Dairy Science**. v. 66, n. 6, p. 1399-1402, 1983.
- FERRAZ, J. B. S.; ELER, J. P. Qualidade dos dados coletados. In: SIMPÓSIO NACIONAL DA SOCIEDADE BRASILEIRA DE MELHORAMENTO ANIMAL, 1998. Uberaba, MG, **Anais...** Uberaba: Sociedade Brasileira de Melhoramento Animal, 1998., p. 265-269.
- FERRIANI, L.; ALBUQUERQUE, L. G.; BALDI, F. S. B.; VENTURINI, G. C.; BIGNARDI, A. B.; SILVA, J. A. II. V.; CHUD, T. C. S.; MUNARI, D. P.; OLIVEIRA, J. A. Parâmetros genéticos de características de carcaça e de crescimento de bovinos da raça Nelore. **Archivos de Zootecnia**, v. 62, n. 237, p. 123-129, 2013.
- FOUILLOUX, M. N.; CLÉMENT, V.; LALOE, D. Measuring connectedness among herds in mixed linear models: From theory to practice in large-sized genetic evaluations. **Genetics Selection Evolution**, v. 40, p. 145-159, 2008.
- FOUILLOUX, M. N., MINERY S.; MATTALIA S.; LALOE D. Assessment of connectedness in the international genetic evaluation of Simmental and Montbeliard breeds. **Interbull Bulletin**, n. 35, p. 129-135, 2006.
- FOULLEY, J. L., HANOCQ, E., BOICHARD, D. A criterion for measuring the degree of connectedness in linear models of genetic evaluation. **Genetics Selection Evolution**, v. 24, n. 4, p. 315-330, 1992.
- FRIES, L. A.; ROSO, V. M. Conectabilidade em avaliações genéticas de gado de corte: uma proposta heurística. In: REUNIÃO DA SOCIEDADE BRASILEIRA ZOOTECNIA, 34., 1997, Juiz de Fora. **Anais...** Juiz de Fora: SBZ, 1997. p 159-161.
- KENNEDY, B. W., TRUS, D. Considerations on genetic connectedness between management units under an animal model. **Journal of Animal Science**, v. 71, n. 9, p. 2341-2352, 1993.

- KUEHN, L. A.; LEWIS, R. M.; NOTTER, D. R. Connectedness in Targhee and Suffolk flocks participating in the United States national sheep improvement program. **Journal of Animal Science**, v. 87, n. 2, p. 507-515, 2009.
- KUEHN, L. A.; NOTTER, D. R.; LEWIS, R. M. Assessing genetic gain, inbreeding, and bias attributable to different flock genetic means in alternative sheep sire referencing schemes. **Journal of Animal Science**, v. 86, n. 3, p. 526-535, 2008.
- LALOE, D. Precision and information in linear models of genetic evaluation. **Genetics Selection Evolution**, v. 25, p. 557-576, 1993.
- LAUREANO, M. M. M.; BOLIGON, A. A.; COSTA, R. B.; FORNI, S.; SEVERO, J. L. P.; ALBUQUERQUE, L. G. Estimativas de herdabilidade e tendências genéticas para as características de crescimento e reprodutivas em bovinos da raça Nelore. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 63, n. 1, p. 143-152, 2011.
- LEWIS, R. M.; SIMM, G. Selection strategies in sire referencing schemes in sheep. **Livestock Science**, v. 67, p. 129-141, 2000.
- LOPES, F. B.; MAGNABOSCO, C. U.; PAULINI, F.; SILVA, M. C.; MIYAGI, E. S.; LÔBO, R. B. Genetic analysis of growth traits in polled Nelore cattle raised on pasture in tropical region using Bayesian approaches. **PLoS ONE**, v.8, n. 9, p. 1-6, 2013.
- MARTINS, E. N. Uso de modelos mistos no melhoramento animal. In: SIMPÓSIO INTERNACIONAL DE PRODUÇÃO de NÃO-RUMINANTES, 31., 1994, Maringá. **Anais...** Maringá: SBZ, 1994. p.41-46.
- MATHUR, P. K.; SULLIVAN, B.; CHESNAIS, J. Measuring connectedness: concept and application to a large industry breeding program. In: World Congress on Genetics Applied Livestock Production, 7., 2002, Montpellier. **Anais eletrônicos...** Montpellier: INRA, 2002. Disponível em: <http://www.ccsi.ca/include/docs/connectedness/connectedness_wcgalp_2002.pdf>. Acesso em 03 de Agosto de 2014.
- MATHUR, P. K., SULLIVAN, B., CHESNAIS, J. **Estimation of the degree connectedness between herds or management groups in the canadian swine population**. Ottawa: Canadian Center for Swine Improvement, 1999. 19 p. Disponível em: <http://www.ccsi.ca/include/docs/connectedness/connectedness_asds_1998.pdf>. Acesso em 25 de julho de 2014.
- MATOS, A. S.; SENA, J.S.S.; MARCONDES, C.R.; BEZERRA, L.A.F.; LÔBO, R.B.; RORATO, P.R.N.; CUCCO, D.C.; ARAÚJO, R.O. Interação genótipo-ambiente em rebanhos Nelore da Amazônia Legal. **Revista Brasileira Saúde Produção Animal**, v. 14, n. 3, p. 599-608, 2013.
- MISZTAL, I; TSURUTA, S.; LOURENCO, D.; AGUILAR, I; LEGANA, A.; VITEZICA, Z. Manual for BLUPF90 family of programs, 2014. Disponível em: <http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all1.pdf>. Acesso em 05 de Agosto de 2014.
- OLIVEIRA JUNIOR, G. A.; ELER, J. P.; FERRAZ, J. B. S.; PETRINI, J.; MATTOS, E. C.; MOURÃO, G. B. Definição de grupos genéticos aditivos visando melhor

- predição de valores genéticos em bovinos de corte. **Revista Brasileira Saúde Produção Animal**, v. 14, n. 2, p. 277-286, 2013.
- OLIVEIRA, H. N. Grupos de contemporâneos e conectabilidade. In: CURSO SOBRE A AVALIAÇÃO GENÉTICA EM BOVINOS DE CORTE, 1995, Ribeirão Preto. **Anais...** Ribeirão Preto: FPCRC, 1995. p. 1-13.
- PEDROSA, V.B.; ELER, J.P.; FERRAZ, J.B.S.; SILVA, J.A.V.; RIBEIRO, S.; SILVA, M.R.; PINTO, L.F.B. Parâmetros genéticos do peso adulto e características de desenvolvimento ponderal na raça Nelore. **Revista Brasileira Saúde Produção Animal**, v.11, n.1, p. 104-113, 2010.
- PETERSEN, P. H. A test for connectedness fitted for the two-way BLUP sire evaluation. **Acta Agriculturae Scandinavica**, v. 28, n. 4, p. 360-362, 1978.
- ROSO, V. M.; SCHENKEL, F. S. AMC – A computer program to assess the Degree of connectedness among contemporary groups. In: World Congresson Genetics Applied to livestock production, 8., 2006, Belo Horizonte. **Proceedings...** Belo Horizonte, 2006. p.26. (Communication, 27).
- ROSO, V. M.; SCHENKEL F. S.; MILLER S. P. Degree of connectedness among groups of centrally tested beef bulls. **Canadian Journal Animal Science**, v. 84, p. 37-47, 2004.
- SANTOS, N. P. S.; FIGUEIREDO FILHO, L. A. S.; SARMENTO, J. L. R.; MARTINS FILHO, R.; BIAGIOTTI, D.; REGO NETO, A. A. Estimação de parâmetros genéticos de pesos em diferentes idades de bovinos da raça Nelore criados no Meio-Norte do Brasil usando amostragem de Gibbs. **Acta Tecnológica**. v. 7, n. 2, p. 1-7, 2012.
- SEARLE, S. R. **Linear models for unbalanced data**. New York: John Wiley & Sons, 1987.
- SEARLE, S. R. **Linear models**. New York: John Wiley & Sons. 1971. 532p.
- SHIOTSUKI, L.; CARDOSO, F. F.; SILVA, J. A. II. V.; ROSA, G. J. M.; ALBUQUERQUE, L. G. Evaluation of an average numerator relationship matrix model and a Bayesian hierarchical model for growth traits in Nellore cattle with uncertain paternity. **Livestock Science**. v. 144, p. 89-95, 2012.
- SILVA, R. M.; SOUZA, J. C.; SILVA, L. O. C.; SILVEIRA, M. V.; FREITAS, J. A.; MARÇAL, M. F. Parâmetros e tendências genéticas para pesos de várias idades em bovinos Nelore. **Revista Brasileira Saúde Produção Animal**, v. 14, n. 1, p. 21-28, 2013.
- SIMM, G.; LEWIS, R. M.; COLLINGS, J. E.; NEIUWHOF, G. J. Use of sire referencing schemes to select for improved carcass composition in sheep. **Journal of Animal Science**, v. 79, (E. Suppl), p. 255-259, 2001.
- SOGA, N., M. L. SPANGLER, C. R. SCHWAB, P. J. BERGER, AND T. J. BAAS. **Comparison of connectedness measures and changes in connectedness of the U. S. Duroc population**. In: World Congresson Genetics Applied to livestock production, 9., 2010, Leipzig. **Anais eletrônicos...** Leipzig, 2010. 4 p. Disponível em: < <http://www.kongressband.de/wcgalp2010/assets/html/0909.htm>>. Acesso em 07 de Agosto de 2014.

- TARRÉS, J.; FINA, M.; PIEDRAFITA, J. Connectedness among herds of beef cattle bred under natural servisse. **Genetics Selection Evolution**, v. 42, n. 6, p. 1-9, 2010.
- WEEKS, D. L., WILLIAMS, D. R. A note on the determination of connectedness in a n-way classification. **Technometrics**, v. 6, n. 3, p. 319-324, 1964.
- YOKOO, M. J.; LOBO, R.B.; ARAUJO, F.R.C.; BEZERRA, L.A.F.; SAINZ, R.D. ABULQUERQUE, L.G. Genetic associations between carcass traits measured by real-time ultrasound and scrotal circumference and growth traits in Nelore cattle. **Journal of Animal Science**, v. 88, n. 1, p. 52-58, 2010.
- YOKOO, M.J.; ALBUQUERQUE, L.G.; LÔBO, R.B.; SAINZ, R.D.; JUNIOR, J.M.G.; BEZERRA, L. A. F.; ARAÚJO F.R.C. Estimativas de parâmetros genéticos para altura do posterior, peso e circunferência escrotal em bovinos da raça Nelore. **Revista Brasileira de Zootecnia**, v. 36, n. 6, p. 1761-1768, 2007.

ANEXOS

ANEXO A – Gráficos para análises de convergência dos parâmetros

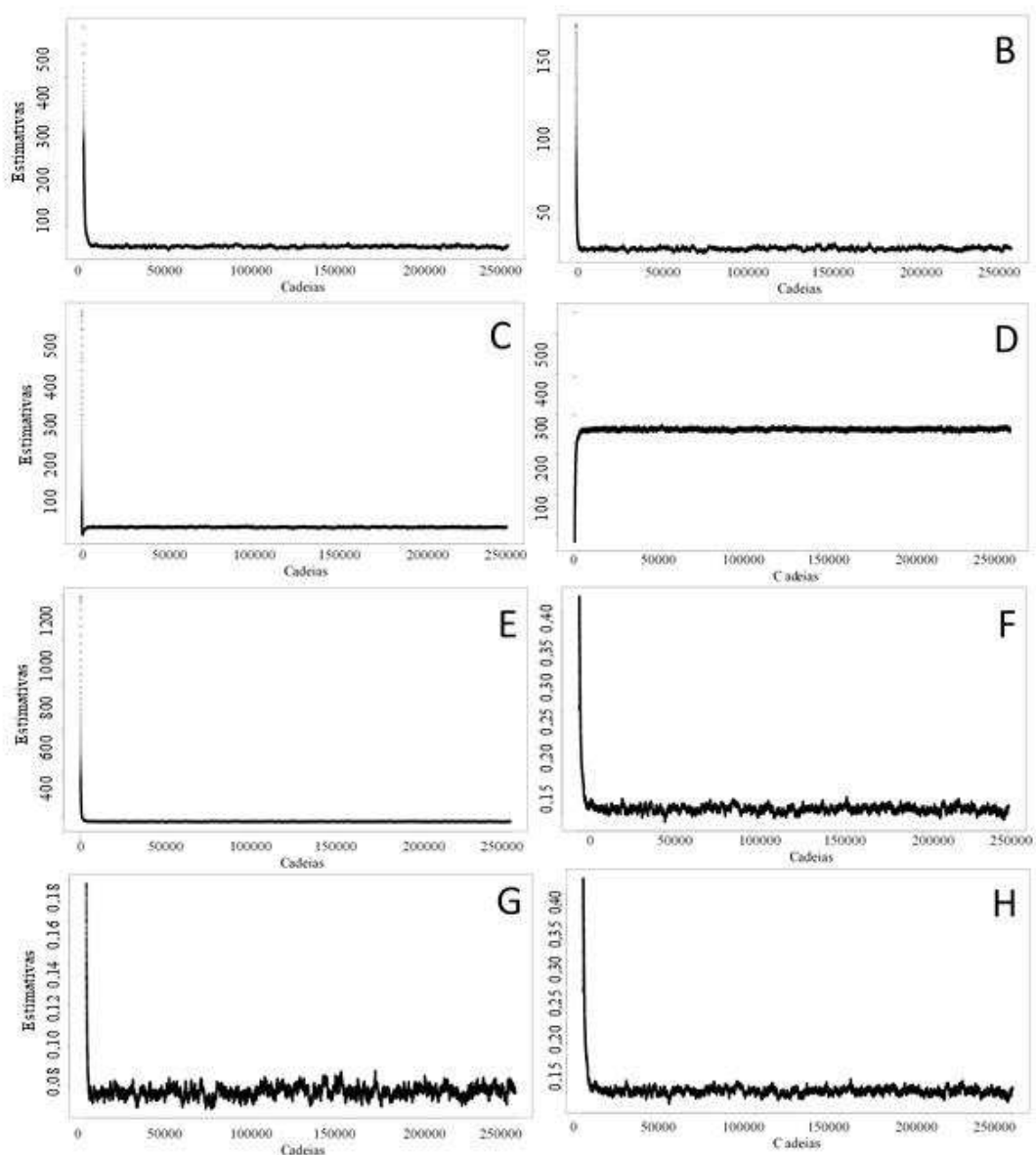


Figura 3. Estimativas de variâncias genética aditiva direta (A), genética aditiva materna (B), de ambiente permanente materno (C), residual (D), fenotípica (E) herdabilidade (F), herdabilidade materna (G) e a proporção da variância de ambiente permanente materno em relação à variância fenotípica (H) em relação as cadeias para estrutura de dados com 2 desvios-padrão e mínimo de 3 animais por grupos de contemporâneos (E1)

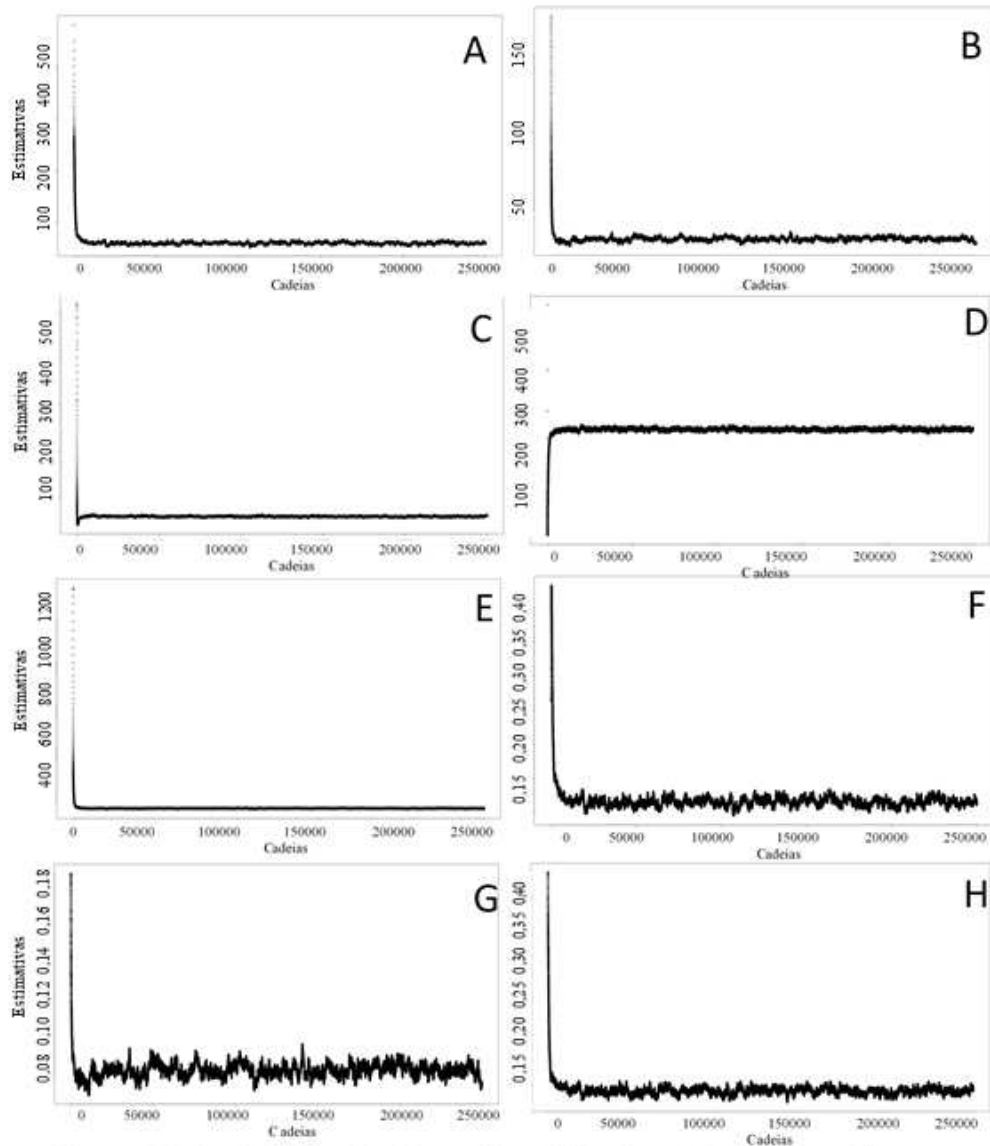


Figura 4. Estimativas de variâncias genética aditiva direta (A), genética aditiva materna (B), de ambiente permanente materno (C), residual (D), fenotípica (E) herdabilidade (F), herdabilidade materna (G) e a proporção da variância de ambiente permanente materno em relação à variância fenotípica (H) em relação as cadeias para estrutura de dados com 2 desvios-padrão e mínimo de 7 animais por grupos de contemporâneos (E2)

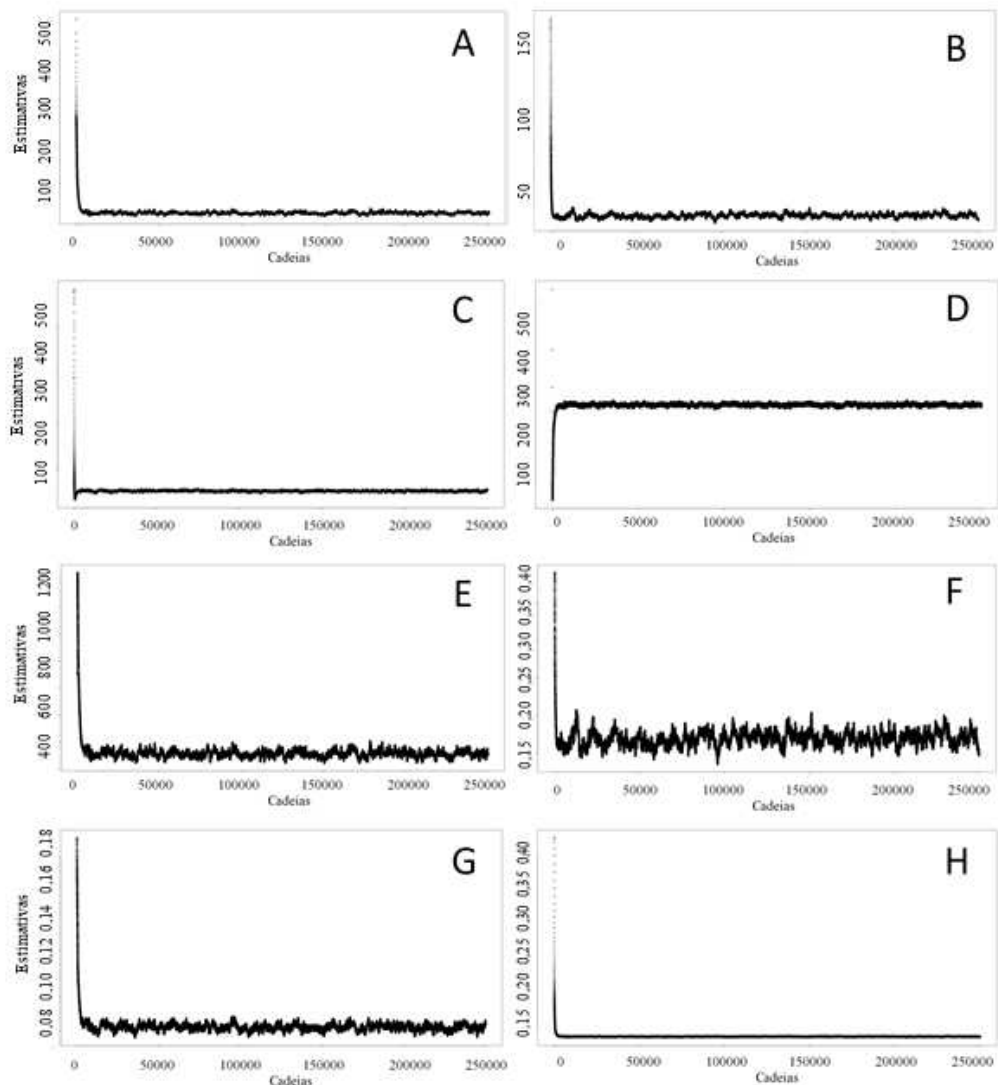


Figura 5. Estimativas de variâncias genética aditiva direta (A), genética aditiva materna (B), de ambiente permanente materno (C), residual (D), fenotípica (E) herdabilidade (F), herdabilidade materna (G) e a proporção da variância de ambiente permanente materno em relação à variância fenotípica (H) em relação as cadeias para estrutura de dados com 2 desvios-padrão e mínimo de 15 animais por grupos de contemporâneos (E3)

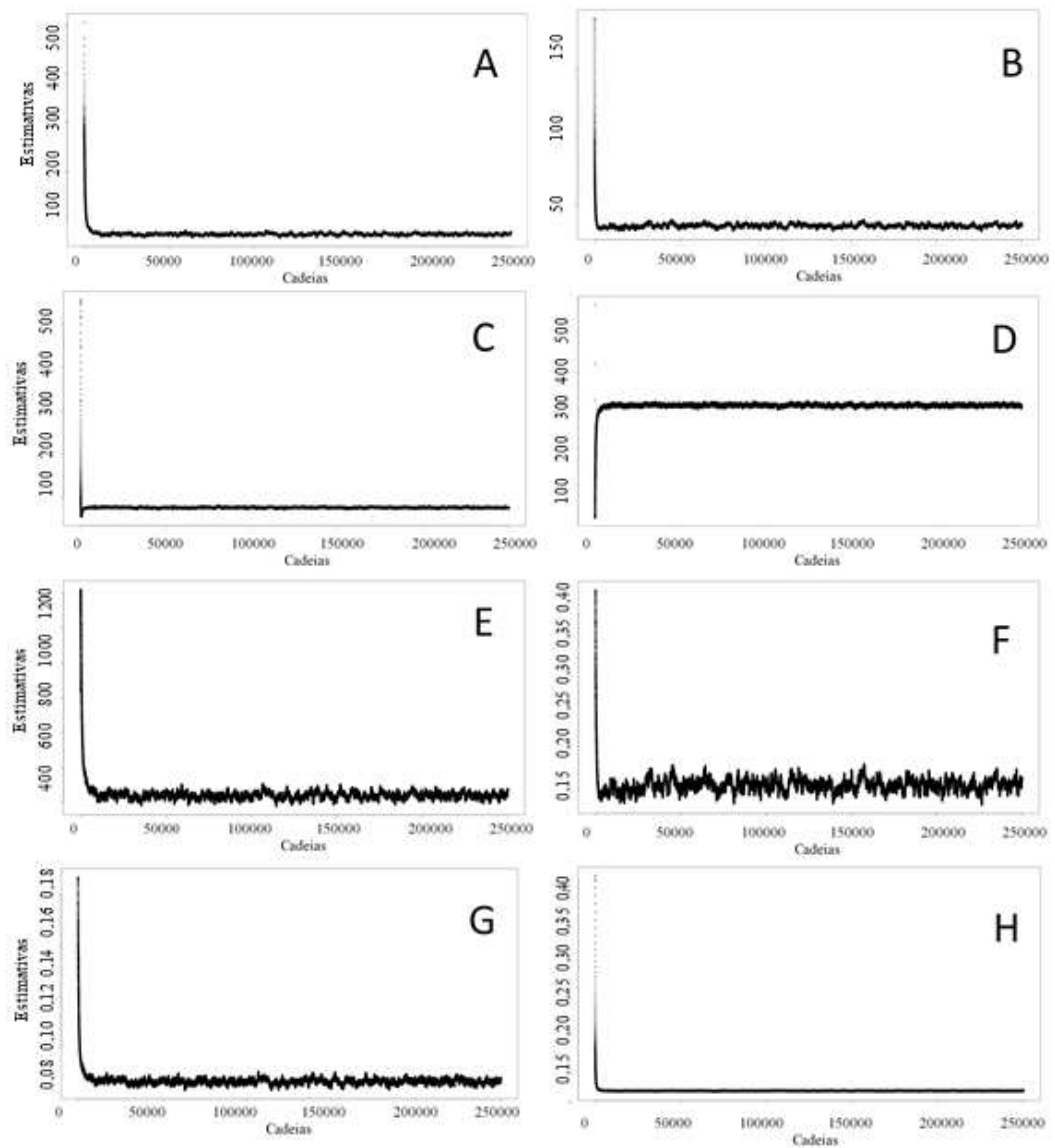


Figura 6. Estimativas de variâncias genética aditiva direta (A), genética aditiva materna (B), de ambiente permanente materno (C), residual (D), fenotípica (E) herdabilidade (F), herdabilidade materna (G) e a proporção da variância de ambiente permanente materno em relação à variância fenotípica (H) em relação as cadeias para estrutura de dados com 2,5 desvios-padrão e mínimo de 3 animais por grupos de contemporâneos (E4)

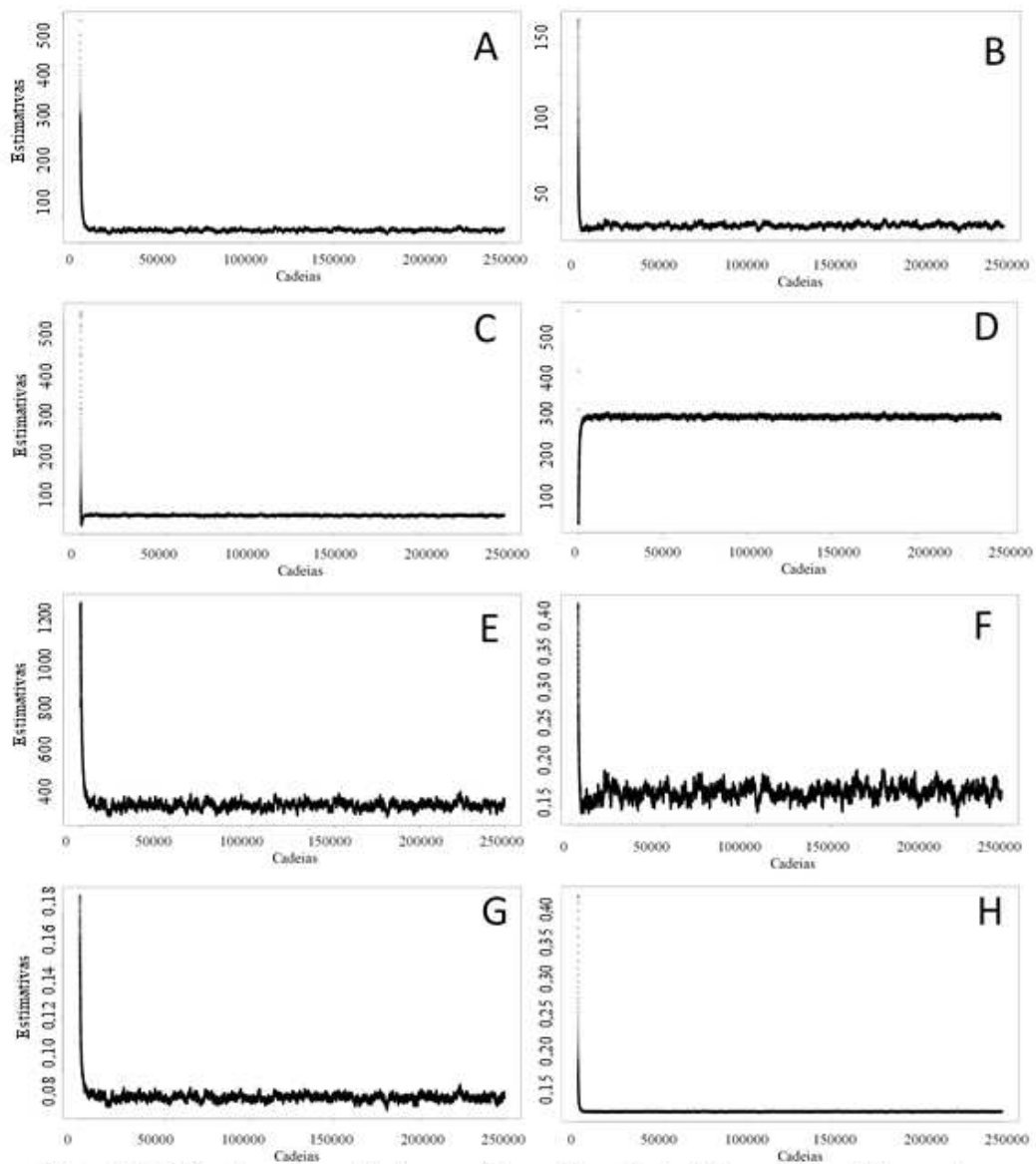


Figura 7. Estimativas de variâncias genética aditiva direta (A), genética aditiva materna (B), de ambiente permanente materno (C), residual (D), fenotípica (E) herdabilidade (F), herdabilidade materna (G) e a proporção da variância de ambiente permanente materno em relação à variância fenotípica (H) em relação as cadeias para estrutura de dados com 2,5 desvios-padrão e mínimo de 7 animais por grupos de contemporâneos (E5)

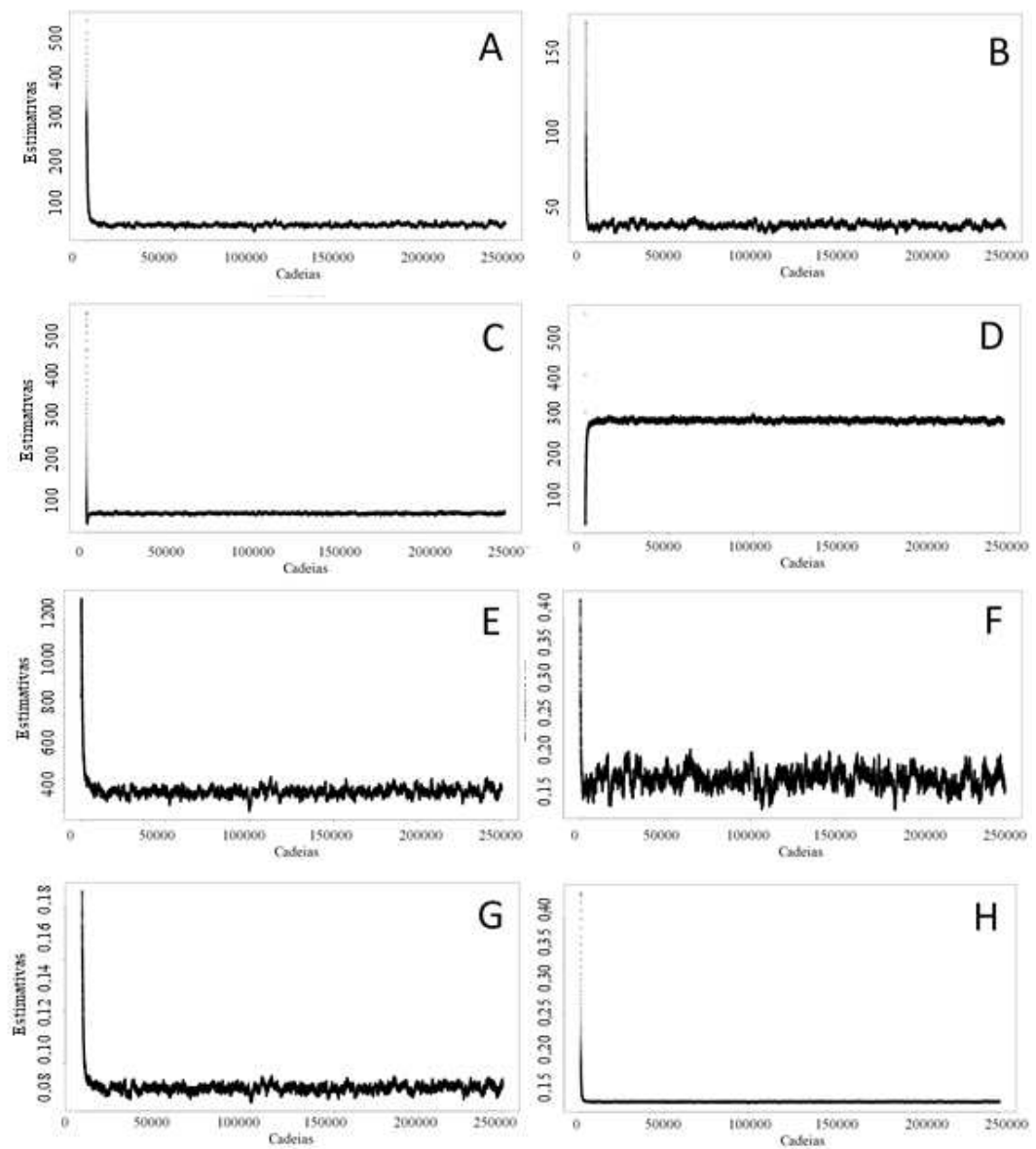


Figura 8. Estimativas de variâncias genética aditiva direta (A), genética aditiva materna (B), de ambiente permanente materno (C), residual (D), fenotípica (E) herdabilidade (F), herdabilidade materna (G) e a proporção da variância de ambiente permanente materno em relação à variância fenotípica (H) em relação as cadeias para estrutura de dados com 2,5 desvios-padrão e mínimo de 15 animais por grupos de contemporâneos (E6)

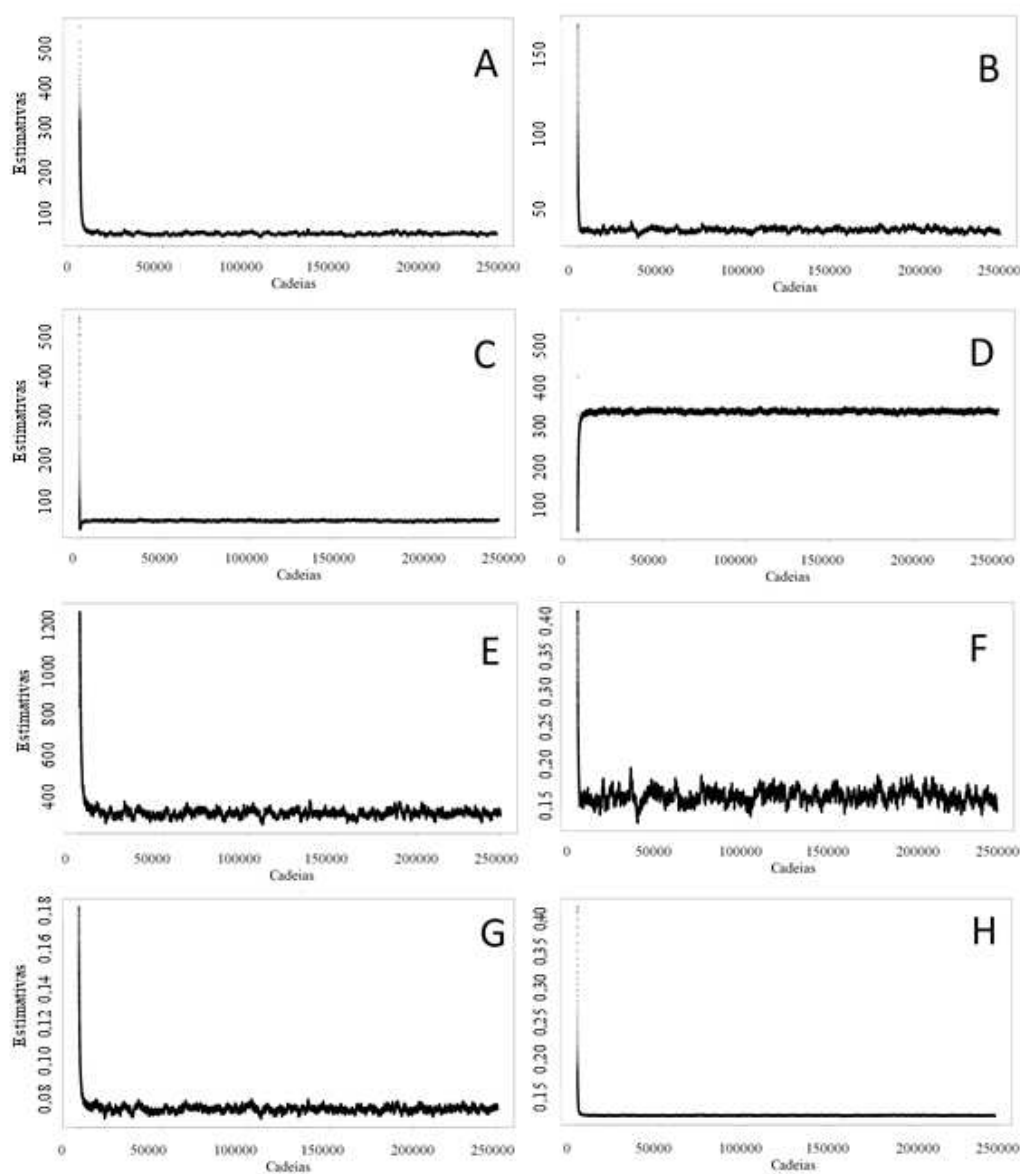


Figura 9. Estimativas de variâncias genética aditiva direta (A), genética aditiva materna (B), de ambiente permanente materno (C), residual (D), fenotípica (E) herdabilidade (F), herdabilidade materna (G) e a proporção da variância de ambiente permanente materno em relação à variância fenotípica (H) em relação as cadeias para estrutura de dados com 3,5 desvios-padrão e mínimo de 3 animais por grupos de contemporâneos (E7)

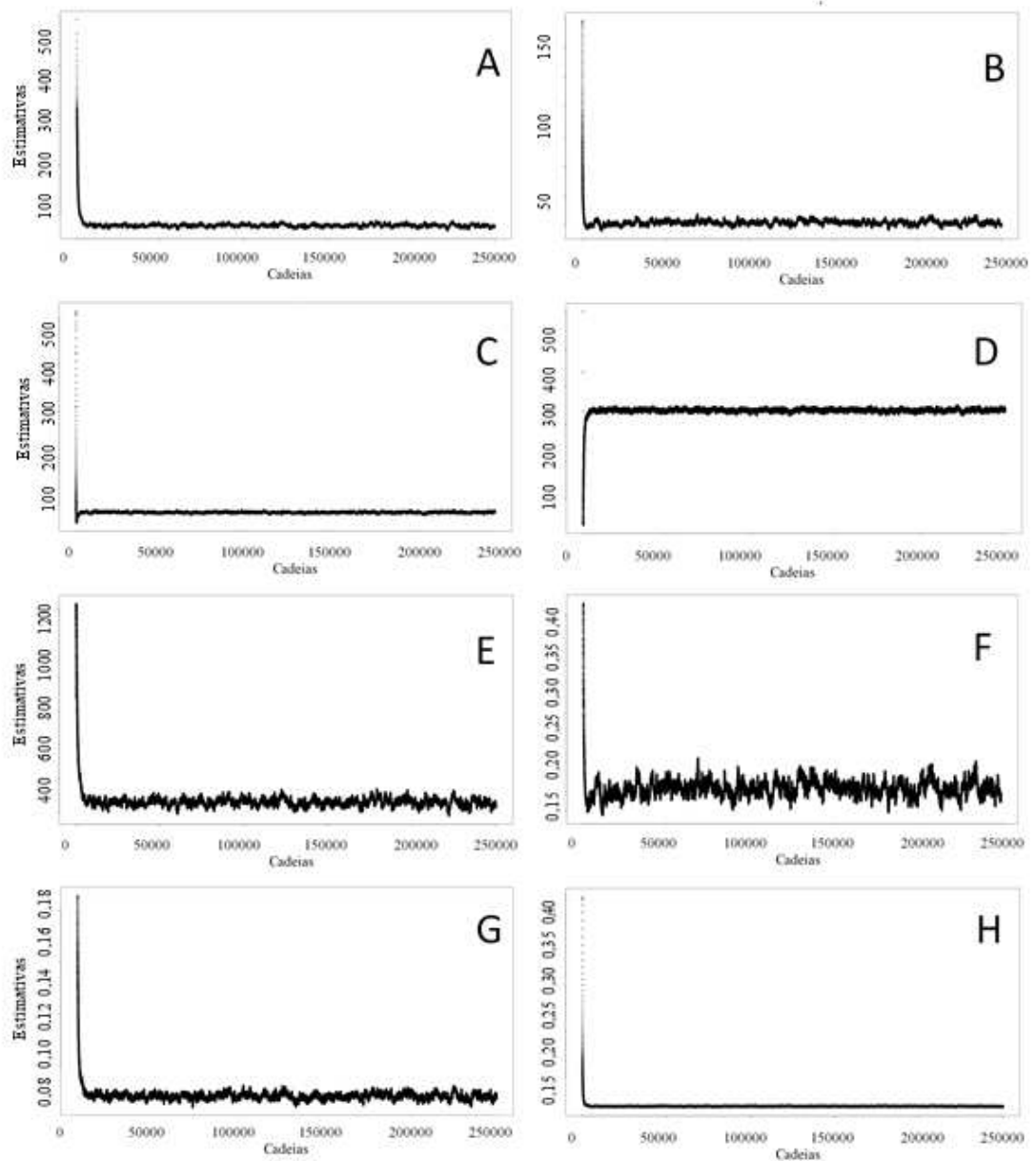


Figura 10. Estimativas de variâncias genética aditiva direta (A), genética aditiva materna (B), de ambiente permanente materno (C), residual (D), fenotípica (E) herdabilidade (F), herdabilidade materna (G) e a proporção da variância de ambiente permanente materno em relação à variância fenotípica (H) em relação as cadeias para estrutura de dados com 3,5 desvios-padrão e mínimo de 7 animais por grupos de contemporâneos (E8)

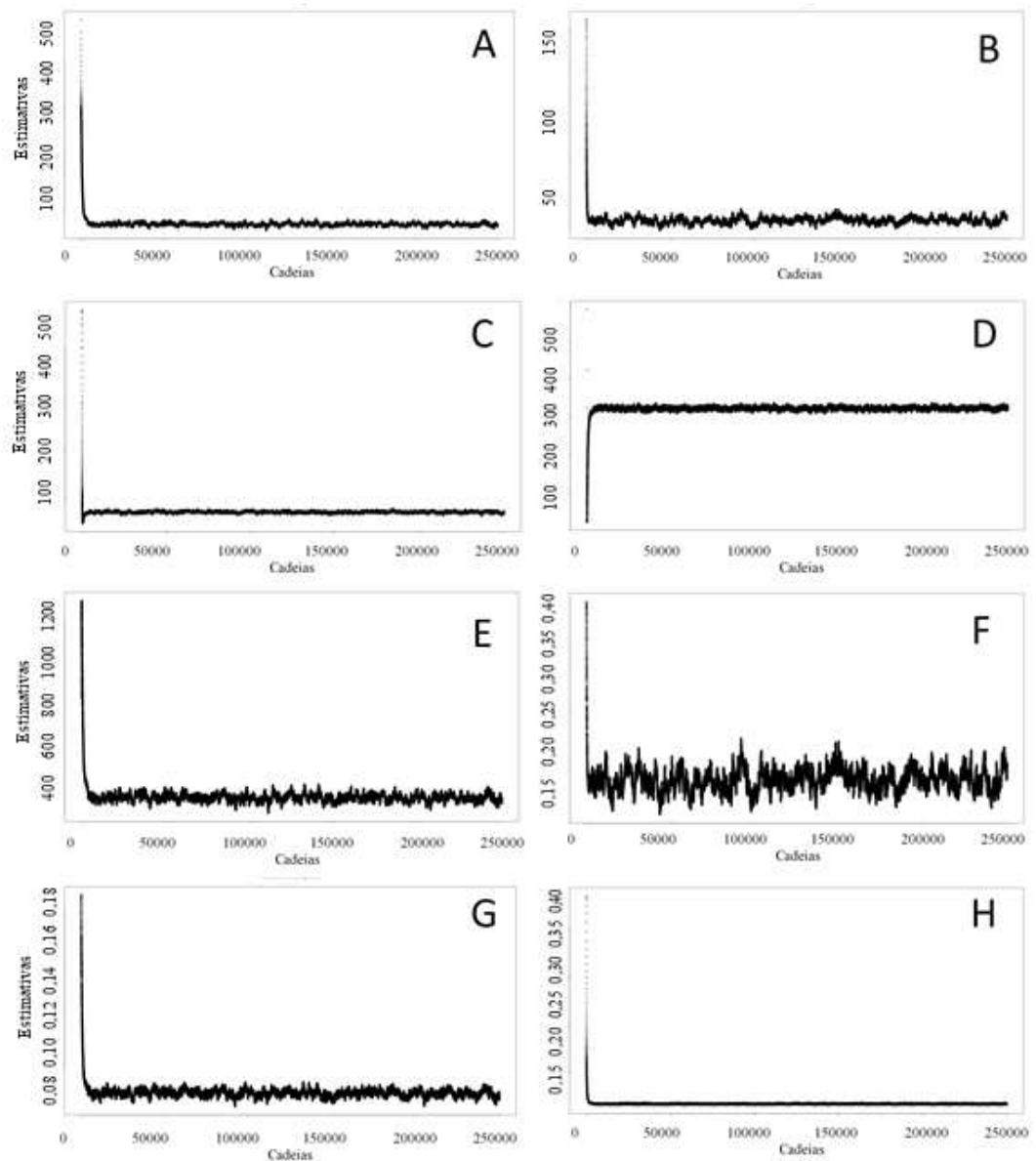


Figura 11. Estimativas de variâncias genética aditiva direta (A), genética aditiva materna (B), de ambiente permanente materno (C), residual (D), fenotípica (E) herdabilidade (F), herdabilidade materna (G) e a proporção da variância de ambiente permanente materno em relação à variância fenotípica (H) em relação as cadeias para estrutura de dados com 3,5 desvios-padrão e mínimo de 15 animais por grupos de contemporâneos (E9)

ANEXO B – Tabela com variâncias genéticas aditivas diretas e maternas, de ambiente permanente materno, residuais, fenotípicas, valores de herdabilidade, herdabilidade materna, proporção da variância de ambiente permanente materno em relação à variância fenotípica e proporção da variância residual em relação a variação fenotípica para as diferentes estruturas de dados

Tabela 8. Variâncias genéticas aditivas diretas (σ_a^2) e maternas (σ_m^2), de ambiente permanente materno (σ_{pm}^2), residuais (σ_e^2), fenotípicas (σ_p^2), valores de herdabilidade (h^2), herdabilidade materna (h_m^2), proporção da variância de ambiente permanente materno em relação à variação fenotípica (pm^2) e proporção residual em relação a variação fenotípica (e^2) para as diferentes estruturas de dados (E_i)

Estruturas	σ_a^2	σ_m^2	σ_{pm}^2	σ_e^2	σ_p^2	h^2	h_m^2	pm^2	e^2
E1	60,21	30,06	63,39	230,90	384,57	0,16	0,08	0,17	0,60
E2	61,25	30,91	60,53	216,49	369,18	0,17	0,08	0,16	0,59
E3	64,17	31,67	56,31	206,86	359,02	0,18	0,09	0,16	0,58
E4	70,36	36,16	77,00	256,25	439,77	0,16	0,08	0,17	0,58
E5	73,51	37,71	76,52	245,12	432,86	0,17	0,09	0,18	0,57
E6	76,27	40,11	72,23	234,53	423,15	0,18	0,09	0,17	0,55
E7	74,80	40,21	82,90	276,42	474,32	0,16	0,09	0,17	0,58
E8	78,22	42,22	83,66	268,13	472,24	0,17	0,09	0,18	0,57
E9	82,65	44,90	82,59	261,25	471,39	0,17	0,09	0,17	0,55