Proc. Assoc. Advmt. Anim. Breed. Genet. ***17****: 545-548*

THE EFFECT OF LINKAGE AND FLOCK SIZE ON THE ACCURACY OF ESTIMATED GENETIC GROUP EFFECTS

M. Khusro, D.J. Brown, H.-U. Graser and B. Tier

Animal Genetics and Breeding Unit[1], University of New England, Armidale, NSW 2351, Australia

SUMMARY

Simulation was used to examine the effect of linkage and flock size on the accuracy of estimation of genetic group effects and their associated variance in a single-trait across-flock evaluation. The accurate partitioning of genetic and environmental effects and the accuracy of estimation of genetic group variance is dependent on the level of linkage (at least 20%) among flocks, inclusion of genetic group effects in the evaluation model, size of genetic groups and the heritability of the trait.

INTRODUCTION

Field data in the Australian Merino industry are collected across a wide range of geographical regions wherein flocks have varying phenotypic means. Dam and sire identities are missing in varying amounts in most flocks. Therefore, to account for the missing pedigree information and for possible genetic differences in the means of base populations, genetic group (GG) effects are generally included in the evaluation model. However, the estimability of GG effects is an important issue that needs to be investigated prior to fitting GG and other fixed effects simultaneously. Depending on the number of flocks being evaluated and their respective means, GGs formed could vary in number. Additionally, if only few GGs are formed, they may be included in the model as fixed effects (Ashtiani and James, 1993). However, the Australian Merino industry comprises a large number of bloodlines which would make the assumption of treating their mean breeding values as random effects a reasonable one (Ashtiani and James, 1993). This study was undertaken to examine the effect of linkage and flock size on the accuracy of estimation of GG effects in a single-trait across-flock evaluation.

MATERIAL AND METHODS

We simulated two scenarios to estimate the level of linkage required for an accurate estimation of GG effects and their variance, respectively. Additionally, Scenario 2 also investigated the effect of flock size on the accuracy of estimated GG effects.

Scenario 1. Pedigree and performance data on a single trait were simulated for four flocks of 1000 animals each involving only one generation. The progeny descended from 3 types of sires that did not have pedigree information; Xsires: external sires used in all four flocks but derived from a fifth flock for which there is no information; WFsires: internal flock sires bred in one flock and used as sires in all flocks; WNsires: Within-flock sires of the four flocks and used only in their own flocks. Two genetic groups (flocks 1 and 2 in GG1; flocks 3 and 4 in GG2) were formed and a third genetic group (GG3) was for Xsires. Three different heritabilities (h^2) were tested; 0.43, 0.30, and 0.10. Each sire was mated to 10 dams and 10 progeny per sire were generated. Further testing was performed (with h^2=0.30 only) by changing the number of progeny per sire to 5 and 2. The TBVs (True breeding

[1] AGBU is a joint venture of NSW Department of Primary Industries and University of New England

values) of progeny in each flock were simulated (N~0, σ_a^2) as: $TBV_i = ½ (TBV_{sire} + TBV_{dam}) + \Phi$, where Φ is a factor for Mendelian sampling. The simulated TBVs of all sires and dams included the GG effect plus the additive genetic effect. All dams were assumed to have been bred in the same flock in which progeny were generated. For each animal, a phenotypic observation on a single trait was simulated around a global mean of 19 micron and phenotypic variance of 3.5 micron^2 at each level of h^2. The phenotype was simulated as the sum of the mean, TBV, flock effect, GG effect and residual (N~0, σ_e^2). A genetic difference of 0.8 micron (-0.4 and 0.4 micron for GG1 and GG2, respectively) in the mean performance was simulated between the two GGs with the flock effects of -0.5, -0.7, 0.5, and 0.3 micron for flocks 1, 2, 3, and 4, respectively. The mean effect of GG3 was simulated to be zero. The proportion of linkage (defined as percent progeny from link sires) due to either Xsires or WFsires between the four flocks was varied from 0 to 50%.

Scenario 2. Data and pedigree for a single trait were simulated for 50 flocks with 100 animals each per flock from a single generation. The progeny descended from WNsires and Xsires described above. Two flocks each were assigned to 25 GGs with all Xsires being allocated to the 26th GG. For each GG, a random effect was simulated (N~0, σ_{gg}^2), and σ_{gg}^2 was the same as σ_a^2, which was 1.5 micron^2 for a h^2 of 0.43. The flock effects were sampled from a uniform distribution with values ranging between 0 and 1 micron. Genetic linkage (defined as percent progeny from link sires) was established among 50 flocks through Xsires and was varied over 5 levels from 10, 20, 30, 40, to 50%. Flock sizes of 200, 300 and 400 were also examined. This scenario was further tested by generating varying progeny per sire from constant number of Xsires for flock size of 100, 200, 300 and 400. Additional testing was performed by treating each flock as a separate GG. Thus, 51 GGs were considered and the 51st GG was formed by Xsires.

Genetic analysis. For Scenario 1, GG effects and EBVs were estimated from each replicate in univariate analyses using an animal model in ASReml (Gilmour *et al.* 2002). The mixed models fitted without and with GG effects were $y = Xb + Zu + e$ and $y = Xb + ZQg + Zu + e$, respectively where y, b, u, g, and e are the vectors of records, fixed (environmental) flock effects, additive effects, GG effects, and residual, respectively; X, Z and Q are the incidence matrices associating records in **y** with the fixed flock, random animal genetic and GG effects in b, u and g, respectively. Each analysis was repeated without and with GG effects in the model and at varying levels of each type of sire linkage. GGs were defined on the basis of source of origin and were treated as fixed effects using the approach of Westell *et al.* (1988). For Scenario 2, data were analyzed in the same manner as described above. However, GGs were fitted in the model as random effects. The covariance among different GGs was assumed to be zero. All results shown are averages from 100 replicates.

RESULTS AND DISCUSSION

Scenario 1. The results presented in Figure 1 are for h^2 of 0.42 and are at varying levels of Xsire linkage. At 15 to 20% Xsire linkage, the GG effects are estimable with relatively high accuracy (low standard errors; Figure 1). Similar results were obtained for WFsire linkage except for the lower standard errors since at WFsire linkage only two GG effects were estimated whereas with Xsire linkage three GG effects were estimated. Since contrasts between the two GGs (2 flocks each in a GG) were available, their effects were estimable. However, if the data structure does not allow for such comparisons between different GGs, their effects would be confounded with other fixed effects included in the model such as flock. The estimated flock effects were the same as simulated when

GGs are included in the model, provided linkage was greater than zero. When GGs were not fitted in the model, GG and flock effects could not be separately estimated.

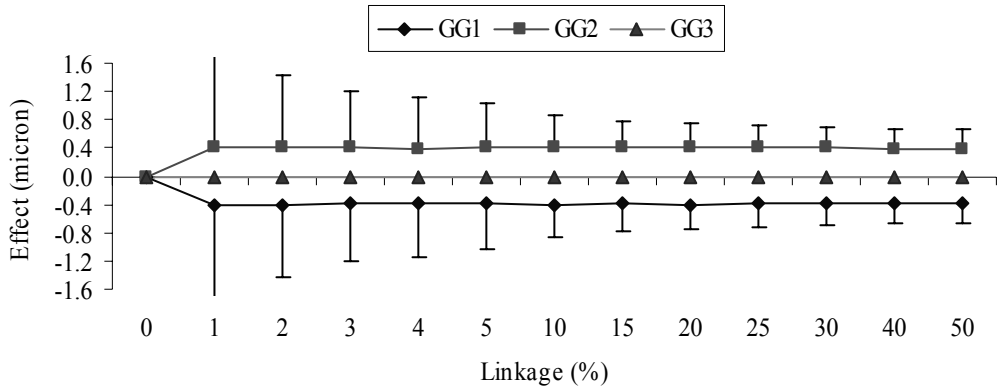

Figure 1. GG effects at varying levels of Xsire linkage (vertical lines indicate standard errors)

Similar results (not shown) were obtained for h^2 of 0.30 and 0.10 implying that the linkage required for estimating GG effects accurately remains the same irrespective of the h^2 of the trait. The effect of h^2 on the accuracy of estimated GG effects decreases gradually with increasing linkage between GGs (Table 1). The number of progeny per sire was then varied to 5 and 2 (h^2 of 0.30). The mean GG effects were estimable at all levels of linkage above zero. At the same number of link progeny, the average standard errors of estimated GG effects increased when more progeny per sire were generated (Table 2), for example, 2/10 = 0.64 versus 10/2 = 0.91.

Table 1. Standard errors (micron) of estimated GG effects at varying levels of heritability and linkage

Heritability	Linkage (%)									
	1	2	3	4	5	10	15	20	25	30
0.10	1.02	0.71	0.58	0.51	0.45	0.33	0.27	0.25	0.23	0.21
0.30	1.29	0.91	0.75	0.64	0.59	0.42	0.35	0.31	0.28	0.27
0.42	1.44	1.02	0.81	0.74	0.65	0.47	0.39	0.34	0.32	0.29

Table 2. Standard errors (micron) of estimated GG effects at varying number of link sires and progeny per sire (h^2 = 0.30)

Progeny	Number of link sires									
per sire	1	2	3	4	5	10	15	20	25	30
2	2.00	1.42	1.17	1.00	0.90	0.64	0.52	0.46	0.41	0.38
5	1.51	1.08	0.87	0.75	0.68	0.49	0.39	0.35	0.32	0.29
10	1.29	0.91	0.75	0.64	0.59	0.42	0.35	0.31	0.28	0.27

Scenario 2. At any given level of linkage and for varying flock sizes, the estimated GG variance is the same as simulated. The standard errors of estimated GG variance for large flocks are slightly but not significantly smaller than those associated with small flocks (results not presented). The accuracy of estimated GG effects was greater (higher correlation between simulated and estimated GG effects) when 26 effects were estimated in comparison to that with 51 effects (Table 3). This shows that the accuracy of estimated GG effects is higher when small numbers of GG effects are estimated.

Table 3. Correlation between simulated and estimated GG effects (26 versus 51 GGs) for varying flock sizes and levels of linkage

Linkage	Flock size							
(%)	100(51)	100(26)	200(51)	200(26)	300(51)	300(26)	400(51)	400(26)
10	0.70	0.78	0.81	0.88	0.86	0.91	0.88	0.93
20	0.80	0.87	0.87	0.92	0.89	0.94	0.92	0.96
30	0.82	0.87	0.89	0.93	0.91	0.95	0.92	0.97
40	0.83	0.89	0.88	0.93	0.90	0.96	0.92	0.97
50	0.82	0.89	0.88	0.93	0.90	0.96	0.91	0.96

CONCLUSION

The accurate partitioning of genetic and environmental effects, and the accurate estimation of genetic group variance in across-flock evaluations can be achieved at 20% linkage across flocks in conjunction with genetic grouping. The number of progeny per link sire has a relatively smaller effect on the mean standard errors of genetic group solutions which are more a function of the total number of progeny from link sires. The accuracy of estimated genetic group effects is dependent on the number of genetic group effects fitted in the model. The size of genetic groups formed should be sufficiently large to estimate their effects accurately.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge MLA for providing funding and support for this study.

REFERENCES

Ashtiani, S.R.M. and James, J.W. (1993) *Gen. Sel. Evol.* **25**: 339.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. and Thompson, R. (2002) VSN International Ltd, UK.

Westell, R.A., Quaas, R.L. and Van Vleck, L.D. (1988) *J. Dairy Sci.* **71**: 1310.