GENOTYPE x ENVIRONMENT INTERACTIONS AND MERINO BREEDING PROGRAMMES FOR WOOL PRODUCTION

B.J.McGuirk

Strowan Lodge, 1c Albert Terrace, Edinburgh EH10 5EA, UK

SUMMARY

This paper and earlier reviews on genotype*environments interactions, many conducted on research stations and often restricted to fleece weights, indicate that whether the “genotypes” are breeds, Merinos bloodlines, selection flocks or, but rarely, sires, statistically significant genotype*environment interactions are of the scale type, and do not indicate changes of rank. These findings are discussed in the context of national across-flock genetic evaluations, and this paper suggests that greater use should be made of planned sire and bloodline evaluations if we are to better both quantify and explain such interactions, at least to the Merino industry.

INTRODUCTION

The term “genotype x environment interaction” (G x E) implies that the difference between genotypes is not equal in a set of environments (Woolaston, 1987). “Genotypes” might mean breeds, or, in the context of Merinos, strains, bloodlines, or even smaller genetic differences, such as between sire progeny groups. However, even this rather simple description hides some perhaps unexpected difficulties, such as the advisability of data transformations, or whether genotypes should be treated as fixed or random effects. But to my mind the critical issue is that our definition of interactions deals in outcomes, and not necessarily with causes, and that to “understand” such interactions, we feel the need to go beyond outcomes, and deal with causes. It seems to me fairly self-evident that as we move away from definable “causes”, we are less likely to find consistent results across “experiments”.

The possible existence of G x E interactions for fleece traits has long been discussed among practical Merino breeders and research workers. The focus here is on studies that were conducted in what might I term the “research station era” (see Morley, 1980, for a brief history), when comparisons were made of breeds, strains of Merino and of groups of smaller genetic differences (eg selection flocks), under the relatively well-defined environmental challenges possible on research stations. In other words, “causes” were usually well-controlled, as when samples of genetic groups were fed different diets (intakes) under pen conditions, with all other factors kept constant. As this topic was thoroughly reviewed by Woolaston (1987), the focus here will generally be on more recent studies.

The overriding impression from the older studies, aside from this control over the environmental factors generating the production differences, was the wish to show statistically significant G x E interactions for fleece weight. In consequence, populations of very different genetic merit were compared under environmental (generally nutritional) treatments that were also very different in degree. Thus these trials fall into what Dunlop (1962) termed Type 4 interactions. While this might be a good design to demonstrate a statistically significant G x E interaction, with an ensuing paper almost guaranteed, the relevance of this work to present-day industry concerns is sometimes hard to see. This review will attempt to bridge that gap.

COMPARISON OF WIDELY DIFFERENT GENETIC POPULATIONS - STRAINS AND SELECTION FLOCKS

This topic has had a fairly long history. For example, Dunlop (1962) makes reference to unpublished work of Marston, Pierce and Carter, who compared Strong and Fine-wool sheep fed at two nutritional levels. As Dunlop (op. cit.) states: "The figures quoted make it almost certain that there were appreciable interactions of strain and plane of nutrition, in that the strong-wool strain had a greater capacity to respond in body size and wool production to good nutrition then the fine wool strain".

Saville and Robards (1972) published a notable paper on the same theme, when they compared samples of various Merino strains at Trangie, with ewes being fed to maintain weight, and then *ad libitum*. Again one might conclude that differences in wool growth between the strains are more apparent on unlimited feed (see Table 1) than at maintenance. But from the evidence (the strain means), the differences between the populations are in scale, not in rank.

Table 1. Wool growth per head (g/day) for Merino Strains at different intake levels (from Saville and Robards, 1972).

	Peppin Merinos (Trangie Random Flock)	Bungarees	Collinsvilles
Maintenance Feed	7.1	8.1	8.4
Unlimited Feed	11.4	14.5	15.9

BIG GENETIC DIFFERENCES, SMALL ENVIRONMENTAL DIFFERENCES

In his classic study, Atkins (1980) compared hogget wool production of five breeds (Peppin and South Australian Merinos, Corriedale, Polwarth and a fixed Border Leicester x Merino halfbred) at Temora in different years, in which average production varied between years, but in a manner similar to that on any commercial property.

Using the joint regression method of Freeman (1973) and Hill (1975), Atkins (1980) showed that the breed x year of measurement interaction was significant, and that the "advantage of the breeds with the highest (average) fleece weights", which in this case were the South Australian Merino and the Corriedale, "increased as the mean yearly fleece weight increased".

In a small scale study, Williams (unpublished data) compared the mid-side greasy wool production of Fleece Plus, Random and Fleece Minus rams at pasture over a period in 1962-1963, when the rams were generally grazed together. This was after about five generations of selection. If the "environments" (in this case six week periods of wool growth) are ranked in ascending order, the wool production (measured as g wool produced on a mid-side patch) in the Fleece Plus rams increases by 1.27 g/g increase in average production, compared with only 0.75 g/g increase in the Fleece Minus flock. Again differences between the flocks were greatest at the highest levels of production, which occurred in the Spring.

My own introduction to G x E interactions started with the Fleece Plus Relaxed flock at Trangie (see McGuirk 1980 for a brief mention). The flock was formed in the late 1960s, in response to concerns that response in the Fleece Plus flock had "plateaued" (see Pattie and Barlow 1974 – but see also Hatcher and Atkins 1998). I thought that this suggestion could be tested by splitting the Fleece Plus ewe flock, using extreme selected rams in the Fleece Plus flock, while average rams from the Fleece Plus flock were used to create a Relaxed line. After it was created, all replacements were selected at random from within the Relaxed flock. Unfortunately no formal comparison of the Fleece Plus and Relaxed flocks has yet been made.

Here the Fleece Plus Relaxed flock has been compared with its unselected Random control flock over five groups of hoggets, when we might expect that the relative genetic merit of the flocks would be the same. However if we compare the hoggets in five different years, it is clear that the difference between the flocks in average clean fleece weight increases as the average clean fleece weight of the Random flock increases (Table 2). In short, the better the environment, the bigger the response.

Table 2. Average Clean Fleece Weights (kg) for Fleece Plus Relaxed and Random hoggets.

Drop	Fleece Plus Relaxed	Random	Difference
1970	2.39	1.97	0.42
1971	2.56	2.03	0.53
1973	2.96	2.42	0.54
1968	3.57	2.98	0.59
1969	3.99	3.11	0.88

This is not a radical or even a new suggestion. Turner, Dolling and Kennedy (1968) discussed the possibility that drought conditions affected responses in the latter years of the CSIRO flock selected primarily for clean fleece weight. When summarising 41 years of selection for increased fleece weight at Trangie, Hatcher and Atkins (1998) concluded that the "pattern of response suggests that the expression of genetic superiority may be sensitive to the level of nutrition", and went on to illustrate this point with reference to responses in drought years. The Fleece Plus Relaxed flock at Trangie simply provides another example of this phenomenon.

SMALL GENETIC DIFFERENCES, LARGE ENVIRONMENTAL CHALLENGE

Recently I re-examined the data used by Morley (1956), in what was the first study of G x E interactions with Merino sheep. While it is often cited, the results are rarely discussed. In my re-presentation of this material, I came to very different conclusions from Morley, although my conclusions are apparent in the raw data (McGuirk 2009).

In general terms, Morley (1956) described two groups of daughters of 23 sires, drawn from various selection flocks in 1953. This puts the study at the very beginning of the selection phase at Trangie. All sires were in fact selectively drawn from a common line of rams, and these in turn were mated to ewes which were selected from previously unselected groups of base ewes. But while there was some degree of assortative mating, the genetic differences between the "flocks" were still quite small, and hence similar to that which might be observed in "the industry". Sire groups were split, and, between weaning and 12 months of age, young ewes were managed at pasture to achieve two very different levels of production. In all there were 93 ewes in the Low Plane group, and 80 in the High Plane, so that the number of daughters per sire was low, on average 4 or fewer in each treatment. This obviously has implications for the "power" of any sire x nutrition effect.

Production differences between the two nutritional treatments were considerable. At 17 months of age, average bodyweights were 29.6 kg (Low Plane) and 38.0kg (High Plane), with corresponding clean fleece weights of 2.64 and 3.67 kg. While differences between the treatments for these traits were thus respectively 25 and 35 percent, the effects on clean scoured yield and crimp frequency were small.

When the untransformed data are analysed using a model which includes the effects of Flocks, Nutritional Treatment and their interaction, the interaction effect was significant for clean fleece weight. However if we apply Tukey's test for non-additivity (see Tukey 1949), the Flock x Nutrition interaction for clean fleece weight was reduced to non-significance, as it was if we used log-transformed data. In other words, the statistical significance of the Flock x nutrition interaction for clean fleece weight was simply a scale effect.

If we look at the results of this trial in detail, we find that:

- The flock rankings were similar on the two Nutritional treatments.
- The variation between flocks was less on the Low Nutritional treatment, and

- The difference in performance between the two Nutritional treatments, the responsiveness of the flocks, was positively related to their performance or rankings on the Low treatment.

When the raw data were analysed using a model which includes the effects of Sires, Nutritional Treatment and their interaction, the interaction effect was not significant for any trait. This is hardly surprising for, with very small numbers of offspring per sire, the power of any test for statistical significance is very low. However, the variation (standard deviation) was greater among sires for offspring on the High Nutritional treatment, and the difference between the two nutritional treatments was positively related to sire means on the Low Nutritional treatment.

GENERAL ASSESSMENT OF RESEARCH STATION STUDIES

If we look at fleece weight comparisons involving breeds, strains, selection flocks and the limited information on sires (McGuirk 2009), we consistently find statistically significant G x E interactions, with the differences in performance greater as average production increases This holds both in studies where differences in production were solely due to intake or dietary differences (ie where the “causes” of the differences in production were determined experimentally) or where we are dealing with time effects associated with differences in production. But we see interactions of the scale type, not of rank (see James 2008). However these studies focussed almost exclusively on fleece weight, and we really know little if anything about other economically important traits that are of concern in national evaluations.

But to “understand” the nature of sire x environment interactions, I suggest we need to look at evidence from planned experiments, which is not the same as imposing an experimental “design” on a national dataset, where the performance defines the environmental treatments. We should look at various levels of environmental effects, and whether differences in performance can adequately be explained by scale differences. Ideally, our “experiment”, should be seen as a vehicle to “explain” G x Es to a sceptical audience, as well as establishing statistical significance. Ideally then, our experiment should:

- Include a sufficient (and probably large) number of industry rams.
- Record as many traits as are economically important, including disease traits.
- Environmental differences should wherever possible be imposed, and not simply observed, with three or more “environments” represented.
- Sufficient offspring should be represented in each environment, so that all sire “proofs” have an agreed and minimum level of reliability.
- The results should be presented in a way that they are understood by producers, which will probably mean using regression methods, plotted with a pencil and ruler, rather than the sole reliance on estimates of genetic correlations.

As such an experiment is probably going to be seen as excessively large and expensive, we probably need to look at pre-existing data sets that meet at least some of these conditions. Thus the data set described by Dominik, Crook and Kinghorn (1999) should be revisited, especially as it deals with causes, and not simply with consequences, and includes data on a wide range of economically important traits. The data set used in “Merinos to Match” also needs a more rigorous assessment, even though the genetic groups included there are bloodlines, not sires. And data generated from Central Test Sire Evaluation programmes warrants continued monitoring. None of these studies meets all of the suggested requirements set out above, but I believe that the results might add to our overall understanding of G x E interactions.

My own view is that any G x Es observed for individual traits will be primarily of the scale type. However we might find rank changes if we look at indexes, as traits might show differential responses to changing environmental conditions.

IN WHICH ENVIRONMENT SHOULD WE SELECT?

As James (2008) pointed out, interest in the importance of G x E interactions was ignited by Hammond's (1947) proposal that selection should be in the best conditions available. Falconer (1952) responded that there was no general answer to this issue, which should be addressed in the general framework of correlated traits.

What then of the situation for fleece weight? My reading of the research station information, with only scale type interactions evident, is that the genetic correlation between "environments" is one or very close to that value. Hatcher and Atkins (1987) suggested that the heritability of clean fleece weight is essentially the same in Good and Poor years at Trangie, a conclusion supported by the work of Dominik ,Crook and Kinghorn (1999) in good and poor environments at Katanning. So if we want to improve clean fleece weights in "commercial" situations, unexpected differences in selection differentials aside, I suggest it does not matter whether the business of ram breeding is undertaken in "Stud" or "Commercial" conditions.

GENOTYPE x ENVIRONMENT INTERACTIONS IN THE MERINO INDUSTRY AND THE IMPLICATIONS FOR NATIONAL GENETIC EVALUATIONS

Are G x E interactions likely to be important in national genetic evaluations, such as those published by Sheep Genetics (SG)? Strictly speaking, from the evidence so far to hand, we do not know. Does it matter if sire x environment interactions were important in the Merino industry? The answer to that question is clearly "yes", as such interactions would make across flock sire evaluations rather meaningless, and complicate sire evaluations, even if they were only of the scale type. Yet the industry view is even more scathing. Carrick (2005) reported that a "recent survey among wool producers by the Victorian DPI" found "that over 90% of them believed that sires would **rank differently in different environments**" (my emphasis). Sadly I am unaware of how the producers were sampled for this survey, the precise wording of the questions posed, or if there was any empirical evidence for their views.

After numerous conversations, I am still not sure whether

- Producers do not accept that performance records for rams from the same flock/drop accurately reflect the ranking of these rams when they have offspring in other environments
- Whether the concern rather is to do with the suitability of various bloodlines in different environments
- Whether the concern is with the joint effect of these factors, and hence with cross-flock evaluations.

Each of these questions suggests analyses that might be undertaken using performance and progeny records.

I endorse the current SG practice of publishing information on the statistical importance of sire x flock/year effects (see for example, Brown, Swan, Johnston and Grasser 2009), as is commonly done with other species (see for example Mrode and Swanson 1994). However, perhaps SG should also routinely publish a more detailed description of such data sets. Specifically, for a group of sires, what numbers are used across flocks, and what proportion are sires used in different years? According to data collected Sheep Genetics, 105 sires had offspring born in different flocks in 2007, out of the 58,846 Merinos born in that year, for whom sire identities (1168 rams altogether) were also known. I suggest we need to know more about these 105 rams, whether they were used in other years, or other flocks, as part of Central Test Sire Evaluations, and whether they themselves had performance data. And what of the 228 rams used in flocks other than the one in which they were born?

I would advocate a more direct challenge to voices that are sceptical of cross-flock evaluations. But rather then doing this in a rather generalised manner, phrased in terms of genetic parameters, I would approach the matter rather more forensically, and be led by the sort of data presentations outlined above. This could be done "in-house" by Sheep Genetics, as a routine but integral part of its educational programme, accumulating information around an agreed set of questions.

ACKNOWLEDGEMENTS

I am grateful to the Director, Agricultural Research Centre, Trangie, for continuing access to historical data at Trangie, to Professor John James for statistical advice and general counselling, to numerous other people for stimulating discussions and access to their data, and to Sheep Genetics, especially Dan Brown and Sam Gill, for the data provided by that organisation.

REFERENCES

Atkins, K.D. (1980). *Aust. J. Exp. Agric.Anim.Husb*, **20**:280.

Brown, D.J., Swan, A.A., Johnston, D.J. and Graser, H-U. (2009) *Proc. Aust. Assoc. Anim. Breed. Genet.* **18:** these proceedings.

Carrick, M. (2005). Final report on Genotype by Environment Interaction in Wool Sheep. Meat and Livestock Australia, 58 pages.

Dominik, S., Crook, B.J, and Kinghorn, B.P.(1999). *Proc. Aust. Assoc. Anim. Breed. Genet.* **13:**98.

Dunlop, A.A., 1962. *Aust. J. Agric.Res.* **13**:503.

Dunlop, A.A. and Young, S.S.Y., 1966. *Aust. J. Agric. Res.* **17**: 227.

Falconer, D.S., 1952. *Amer. Nat.* **86**:293.

Freeman, G.H. (1973). *Heredity*, **31**:339.

Hammond, J. 1947. *Biol. Rev.* **22**: 195.

Hatcher, S. and Atkins, K.D. (1998). *Proc. Sixth Wld. Congr. Genet. Appl. Livest. Prod.* **24**:47.

Hill, J. (1975). *J. Agric. Sci.* **85**:477.

James, J.W. (2008). In "Adaptation and fitness in animal populations : evolutionary and breeding perspectives on genetic resource management". Eds J van der Werf, R. Frankham, H-U Graser and A.Gondro. Springer Verlag Berlin. Pp. 151 – 167.

McGuirk, B.J. (1980). *Proc Aust Soc Anim Prod* **13**:171.

McGuirk, B.J. (2009). *Anim. Prod. Sci*. (in press).

Morley, F.H.W. (1953). *Sci. Bull. 73, Dep. Agric. N.S.W.*

Morley, F.H.W. (1956*). Aust. J. Agric Res.*, **7**:140.

Morley, F.H.W. (1980). *Proc Aust Soc Anim Prod* **13**:168.

Mrode, R. and Swanson, G.J.T. (1994) *Proc. Fifth Wld. Congr. Genet. Appl. Livest. Prod.* **17**:19.

Pattie, W.A. and Barlow, R. (1974) *Aust. J. Agric Res.*, **29**:1077.

Saville, D.G. and Robards, G.E. (1972) *Aust. J. Agric Res.*, **23**:117.

Tukey, J.W. (1949) *Biometrics*, **5**: 232.

Turner, Helen Newton, Dolling, C.H.S. and Kennedy, (1968) *Aust. J. Agric Res.*, **19**: 79.

Williams, A.J. (1979). In "Physiological and Environmental Limitations to Wool Growth", ed. J. L. Black and P. J. Reis (The University of New England Publishing Unit, Armidale), pp. 337-354.

Woolaston, R.R. (1987) In "Merino Improvement Programs in Australia". (B.J.McGuirk, ed.), Australian Wool Corporation, Melbourne, Australia. pp. 421 – 435.