

ANIMAL BREEDING FROM INFINITESIMAL MODEL TO MAS: THE CASE OF A BACKCROSS DESIGN IN DAIRY SHEEP (SARDA X LACAUNE) AND ITS POSSIBLE IMPACT ON SELECTION

G. Pagnacco and A. Carta

Giulio.pagnacco@unimi.it

izcszoo@tin.it

Summary

Despite the great advancements achieved by quantitative techniques, selection in some species and for some traits, as in dairy sheep, can be very expensive and requires great efforts. In this paper we review the options offered by quantitative methods and by marker-assisted selection and we discuss the possible impact of a QTL detection design based on a backcross experimental population (Sarda X Lacaune) on sheep breeding.

Keywords

Infinitesimal model, Quantitative Trait Loci, Marker-assisted selection, dairy sheep, Backcross Design

Abstract

Selection under the infinitesimal model

Standard animal breeding techniques have been largely successful in improving the performance of domestic animals in the last century. The success of these techniques finds its roots in the theoretical work of Fisher and Wright who showed how selection, migration and mutation could be derived from the simple Mendelian laws. All that theory was then transferred into the operational application in animal breeding by J. L. Lush (*Animal Breeding Plans*, 1945) and by his many “sons” and “grandsons”. In more recent times another landmark book, *Introduction to Quantitative Genetics*, (Falconer, 1960) probably represents the official birth of the quantitative genetics in plant and especially animal breeding. Plant and animal breeding have based their tools of investigation on this “central paradigm for the analysis of phenotypic variation” (Lynch and Walsh, 1998) even if their methods have radically diverged probably because of the different population structure in crop and domestic animals.

Since the 1950s new techniques based on the mixed model methodology (C. Henderson) had a large impact in predicting the individual breeding value of the animals. But it was not before the 1970s that the development of high-speed computers ensured a widespread application to cattle populations and later to other species.

The great achievements in cattle and pig selection during the last 50 years largely relied on quantitative genetic theory and infinitesimal genetic model. Most of the traits selected show an additive variation which justifies the effort to capitalize on the additive effects summing up, generation after generation, their infinitesimal small contributions. Under infinitesimal model, selection is therefore a sort of blind process where genotypes are progressively modified without any real knowledge about gene number, location, effect and frequencies of the favorable alleles.

The whole process of predicting the breeding value of the animals to subsequently select the genetically superior parents of the next generation is entirely based on sophisticated computations. Complex statistical mixed models can adjust for all sort of environmental factors, maternal effects or repeated records and simultaneously release the requested individual additive genetic merit.

Standard quantitative selection requires therefore field data collection on:

- Individual phenotypic performance
- Expected covariance among animals due to blood relationship between them

Recorded traits so far have been mainly limited to production performances. But even for “simple” traits like milk yield, there may be regional or national differences which make comparisons difficult to manage. Phenotypes may also be sex-limited and recorded late in the life of an animal.

Recording is also quite expensive and in many countries costs are largely covered by National Governments. In Italy, for example, the total budget for selection granted by the Ministry of Agriculture is around 93 million euro, 77 of which is field phenotype recording. Costs of phenotype recording are also quite variable in different species and the value of a phenotyped animal may not justify the expense. As an example, fat and protein content in sheep milk requires the collection of a milk sample which may represent more than 10% of the daily released milk. Recording requires also efficient extension services covering a relevant part of the farms, a structure which cannot be created quickly and cheaply.

Parentage ascertaining is also extremely expensive. In cattle, where AI is largely used, the paternity of an animal is easily predictable. In species where natural mating is predominant, as in sheep, parentage definition is a crucial step in selection, but typing costs may be a relevant part of the certified animal.

In recent years infinitesimal model has been also applied to other functional traits such as longevity, calving ease, milkability or udder health. In perspective, we can foresee an increasing interest in traits related to safe animal products and to food quality for human consumption. Many of these traits have a limited genetic variability and a high cost of recording, for example, pork meat quality for dry-cured ham, CLA, parasite resistance, immune response and so on, which justify the use of different genetic models.

Selection under a finite locus model

In the last two decades the development of molecular genetic techniques has allowed the mapping of some loci affecting the quantitative variation of traits of economic importance. Today their subsequent exploitation in selection is in some case a concrete reality. The primary interest of animal breeding and selection in molecular genetics is in finding all the genes (QTLs) which control a trait, being able to distinguish among their alleles and consequently exploit the variability for faster genetic gain in that trait. In such a scenario we would not need to phenotype the animals, nor wait until the phenotype is expressed. We presently are at the very beginning of this. In very few cases we have reached the target, hitting, in a single gene, the mutation responsible for the protein modification which causes the quantitative variation of some trait. In these cases prediction of breeding value takes advantage of both the polygenic (infinitesimal) component and of the discovered QTL (the finite locus) by fitting *ad hoc* models. More frequently we have some evidence that in a chromosomal region of some cM of length (where several hundreds of genes may be located), a certain animal is probably heterozygous at one or more genes (QTLs) involved in a quantitative trait.

Traditionally two approaches are available to locate a QTL:

- Genome scan
- Candidate gene approach

Genome scan

A genome scan is based, first of all, on the *a priori* knowledge of the chromosomal location of a number of anonymous genetic markers (M). With approximately 150 markers an average of one marker each 20 cM is expected, covering the majority of the genome of a mammal, spanning roughly 30 Morgan. In a sparse map like this, a QTL (Q) may be located at an average distance of 5 cM from the nearest marker.

A second requirement is linkage *disequilibrium* (LD) between M and Q which in an outcrossing population is generally noticeable only for M in very close linkage to Q. LD must therefore be created *ad hoc* through specific experimental designs or investigated within single families. In livestock, the creation of LD is achieved in pigs by crossing distant breeds expected to have different genotypes at Q and M loci. The F1 progeny is then *inter se* mated to produce a F2, or backcrossed to the parental genotypes. This design requires an experimental station where all possible phenotypes can be recorded so that a single experiment allows the mapping of a very large number of traits. In cattle LD can be found within families using the commercial structure of the population (Weller and Soller, 1990) through a Daughter or a Granddaughter design. In this case only QTLs acting on the traits routinely recorded for the standard selection can be located.

By these approaches the position of a QTL can be located, associated to a marker haplotype and cautiously exploited in the marker-assisted selection (MAS) as, for example, in pre-selecting the young bulls entering a progeny test program. However it must be kept in mind that the linkage phase M-Q is subject to recombine, due to the weak linkage in a sparse map. At population level M and Q are in fact in linkage equilibrium (LE) and according to Dekkers (2003) markers of this type can be defined LE-markers.

Once a QTL is spotted in a sparse map, the region can be saturated by more markers creating a dense map and therefore shortening the M-Q distance to 1 to 3 cM. The problem in this case is that in most experiments there are not enough meioses to observe very rare recombination events. Techniques like LD analysis and identity by descent mapping (IBD) try to overcome the problem. The first focuses on the lack of M-Q recombinations, the latter on the conserved DNA region in individuals sharing a common ancestor. Markers of this type can be defined LD-markers.

Candidate gene approach

A candidate gene is one whose involvement in a trait of interest is known from work in other species (basically man or mouse). Through comparative mapping, which describes the pattern of the genomic location of known genes in different species, we may locate the candidate gene in a region previously identified by a genome scan program as likely to host a QTL, where the *Q* and *q* alleles are segregating. At this point the candidate gene becomes a positional candidate gene. It can be cloned and sequenced and hopefully one can find the causative mutation (functional marker) or polymorphic markers inside the gene (direct markers). We now need to fit phenotypic data to these polymorphisms and validate the association of specific marker haplotypes to quantitative trait differences. Here selection can take direct advantage from the gene itself and gene-assisted selection (GAS), rather than MAS, takes place.

Compared to plant breeding, the process of discovering genes affecting biology and metabolism in animals is much slower due to the generally longer generation interval. In addition, costs to raise, maintain and measure animals for a specific experimental design are also extremely high. Apparently, LE-markers are easier to find but less effective in their impact on selection. On the contrary, LD-, direct and functional markers are extremely difficult to obtain, but their effect can be evaluated at the level of change in gene frequency.

However, it seems that the structure of a genome scan project can play a central role both for mapping LE-markers and for field validation of direct and functional markers. In this respect we hereafter present a backcross design in dairy sheep, a species so far left aside by the main stream of these molecular investigations.

Dairy Sheep

At first glance, the economic importance of dairy sheep seems to be low. Milk sheep production is 1.3 % of the total milk production in Italy. The Mediterranean basin, with 60% of total world production, is the most important area. The dairy sheep industry is usually based on local breeds which are very well adapted to their production areas, systems and environments. Traditionally this activity has occupied less favoured areas, using natural resources of low interest for other species, and helping to maintain the ecological equilibrium and the natural landscape. In addition it has contributed to sustaining economic activity and the population in rural areas. However, in certain regions and for certain breeds, the management system varies from semi-extensive to intensive according to the economic relevance of the production chain and the specific environment. The differences in management conditions and the good adaptability of the local sheep to their specific environments leads to remarkable genetic diversity in the breeds. Local breeds are also often linked to “high quality” products protected by quality labels (cheese and meat). The above factors have resulted in the establishment of breeding programs which are designed to improve local breeds because of the difficulties that foreign breeds often encounter when trying to adapt to specific environments.

When compared to other species, dairy sheep have some features that determine the potential results of a breeding programme. Firstly, dairy ewes usually show marked reproductive anoestrous and consequently have a seasonal production pattern. Secondly, dairy sheep are dual purpose, with part of the income (around 25 %) deriving from lamb’s meat, and milking does not start immediately after lambing but usually after an average suckling period of 30 days. Finally, due to the physiological and biological characteristics of the sheep, the use of artificial insemination is limited and natural mating still plays an important role in reproduction. The other crucial point is that each individual recording of any trait is much more costly compared to dairy cattle due to the low individual income from each ewe. Milk production is the principal trait affecting the profitability of dairy sheep flocks and so breeding programmes focus on milk traits. The main objective is always selecting for milk yield and sometimes, in addition, milk composition. However, due to the evolution of the EU agricultural policy and consumer demand, more attention has been now given to traits related to the reduction of production costs (milkability, functional traits, longevity), health (resistance to mastitis or parasitic diseases), food safety (scrapie resistance) and nutritional value (milk content for fatty acids related to human health). In most situations, only a part of these new traits is extensively recorded in the nucleus flocks of the purebred breeding schemes. Therefore research combining classical quantitative approaches and QTL detection is needed, either on-farm by implementing experimental recording schemes, or in experimental flocks, especially for traits difficult and costly to record. Crosses between breeds to produce F2 or backcross experimental population allow the exploitation linkage disequilibrium for genes differing between breeds and to detect genome regions controlling the traits of interest. In light of this, an experimental Sarda x Lacaune backcross resource population was created in 1999 by INRA (France) and IZCS (Italy) in order to detect QTL for both milk production and functional traits by means of an intermediate density genome map. The Sarda and the Lacaune are the two most numerous French and Italian dairy sheep breeds and differ phenotypically for many traits as body size, growth rate, wool, prolificacy, milk yield level and milkability and for their dairy selective history. Furthermore,

they show well structured selection schemes based on a wide use of AI in Lacaune and the combination of AI and controlled natural mating in the Sardinian breed. Results from the experimental backcross population should be of use in planning QTL detection designs in pure breeds with a limited number of investigated genome regions and traits, and in developing tools for on-farm recording.

Overview of the backcross Sarda x Lacaune project

In 1998, 14 elite Lacaune rams from France were mated in Italy to 100 Sarda ewes to produce F1 rams. Ten sons of different Lacaune sires were then mated to 3,000 Sarda ewes to procreate 967 backcross females born in 1999 corresponding to a classical daughter design for QTL detection with ten families of approximately one hundred daughters each. These ewes have been raised in an experimental farm in Sardinia and in 2003 they ended their fourth lactation. The farm is located in the South of Sardinia with a semi-arid Mediterranean climate. The ewes are fed with the same regimen based on 4-5 hours of grazing irrigated mixed pasture of ryegrass and berseem clover with supplementation of lucerne hay, maize silage and concentrates, particularly in winter and late spring. The ewes are machine milked twice a day. A description of the traits recorded in the experiment and the measured phenotypes is reported in table 1.

Table 1
Description of the main phenotypic measures

Traits	Periodicity/Tools
Production traits:	
Daily milk yield	Three times per month/INRA electronic jar
Protein and fat content	Bi-monthly/INRA electronic jar
Growth (as ewe lamb)	Bi-monthly body weight/electronic balance
Body Weight (adult ewe)	Monthly body weight/electronic balance
Body condition score (adult ewe)	Monthly subjective score
Wool weight	Once a year/electronic balance
Wool quality	As ewe lamb/objective measurements & score
Milkability and udder morphology	
Kinetics of milk flow	Three times per month/INRA electronic jar
Udder morphology type traits	Monthly/subjective score
Udder measurements	Once a year/numeric picture analysis
Health traits	
Udder health : Somatic cell count	Bi-monthly/INRA electronic jar
Udder health: clinical mastitis detection	daily/visual detection +microbiological assay
Nematode parasite resistance	2- 3 times per year /fecal eggs count
Oestrus ovis resistance	Once a year/antigen ELISA test
Reproductive traits	
Ovulation rate	Once a year/laparoscopy
Fertility (after synchronization)	Date of lambing on the controlled cycle
Prolificacy	Litter size at lambing
Ability of out-of season mating	Once a year before mating/Progesterone assay
Anti PMSG antibody	3 ELISA tests around AI
Embryo mortality	Radio-immuno assay of PSPB
Milk quality traits	
Milk fatty acids composition	Once a year /gaschromatography

Genome scan

INRA has made a panel of around 155 microsatellite markers to enable a complete scan of the autosomic sheep genome. Most of the genotypings has been carried out by IZCS in multiplex of 5 – 9 microsatellites using an ABI377-96 automatic sequencer.

The results of the genome scan are summarized in table 2. For each chromosome the total length in cM (TL) the length of the analyzed segment (ASL) the number of analyzed markers (MN) and the average percentage of informative meioses (INF) is reported. INF is calculated as the ratio of the number of daughters for which it was possible to identify the marker allele received from the sire to the total number of daughters. It is evident that the sheep genome was not completely covered, particularly for some chromosomes where the length of the analyzed linkage group or % of informative meioses were not yet sufficient.

Table 2
Statistics of the genome scan

OAR	TL	ASL	MN	INF	OAR	TL	ASL	MN	INF
1	346	307	16	0.60	14	118	93	5	0.81
2	306	300	12	0.50	15	124	51	2	0.65
3	315	247	10	0.52	16	87	71	4	0.62
4	134	126	6	0.52	17	122	98	4	0.59
5	150	95	3	0.64	18	122	89	6	0.51
6	157	123	6	0.58	19	72	57	4	0.58
7	145	107	6	0.38	20	87	32	3	0.62
8	126	61	3	0.59	21	75	60	5	0.66
9	126	82	5	0.61	22	83	62	3	0.51
10	100	86	4	0.67	23	73	33	2	0.66
11	127	62	4	0.44	24	97	14	2	0.53
12	102	96	6	0.58	25	69	-	1	0.46
13	137	102	5	0.48	26	70	63	4	0.55

Preliminary QTL findings

Genome regions affecting traits were found for the majority of the measured traits. The next step will be the fine mapping of some regions in order to more precisely localise the QTL and try to find a candidate gene to study. Simultaneously, for the most interesting traits and regions QTL detection designs will be carried out in the pure breed in order to verify the segregation of the found QTL also in the original populations. With this aim, 20 Sarda breeding sires were mated to the backcross females in order to produce families of around 40 daughters. This experimental population is currently being raised in the same experimental farm as the backcross population. Moreover, the sampling of biological material for DNA extraction has started on the young females of the nucleus flock. However, the specific design to be applied to the pure breed population will be optimised either for routinely measured traits or traits difficult to record following the simulation results of Casu *et al.*, 2003.

References

- A. Carta, F. Barillet, D. Allain, Y. Amigues, B. Bibé, L. Bodin, S. Casu, E. Crihiu, J. M. Elsen, A. Fraghi, L. Gruner, J. Jacquiet, S. Ligios, C. Marie-Etancelin, L. Mura, G. Piredda, R. Rupp, S. R. Sanna, A. Scala, L. Schibler, S. Casu. 2002. QTL detection with genomic markers in a dairy sheep backcross Sarda x Lacacune resource population. *Proceeding of the 7th Congress on Genetic Applied to Livestock Production. Montpellier*. Vol 29: 211-214.

- A. Carta, G. Piredda, M. Fiori, C. Leroux, F. Barillet. 2003. A genome scan to detect QTL for CLA content in the milk fat of dairy sheep. *54th EAAP Annual Meeting, 31 August-2 September, Rome, Italy. Book of Abstract, 89.*
- S. Casu, A. Carta, J. M. Elsen. 2003. Strategies to optimize QTL detection designs in dairy sheep populations: the example of the Sarda breed. *Options Méditerranéennes, Vol A-55: 19-24.*
- Jack C. M. Dekkers. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. *EAAP Rome 2003.*
- Grisart, Wouter Coppieters, Frédéric Farnir, Latifa Karim, Christine Ford, Paulette Berzi, Nadine Cambisano, Myriam Mni, Suzanne Reid, Patricia Simon, Richard Spelman, Michel Georges, Russel Snell. Positional Candidate Cloning of a QTL in Dairy Cattle: Identification of a Missense Mutation in the Bovine DGAT1 Gene with Major Effect on Milk Yield and Composition. *Genome Research 2001, 12:221-231.*
- Morris Soller. Genetic mapping of the bovine genome using DNA-level Markers to identify Loci Affecting Quantitative Traits of Economic Importance. *J. Dairy Sci. 1990, 73:2628-2646.*