Genomic application in sheep and goat breeding

Rachel Rupp, Sebastian Mucha, Helene Larroque, John McEwan, Joanne Conington

*INRA, GenPhySE, Toulouse, France
†SRUC, Edinburgh, UK
‡AgResearch, New Zealand

Implications

- Genomic studies in small ruminants were first possible in 2009 with the development of the 50K ovine SNP chip.
- Genomic evaluation has now been implemented in sheep in New Zealand and Australia, dairy sheep in France, and in goats in France and the UK.
- Specific issues of genomic selection for these species include: small reference population sizes, low linkage disequilibrium, multi-breed evaluations, lack of phenotype recording in many countries, and marginal cost-benefit at historic genotyping costs.
- Rapidly reducing genotyping cost coupled with a better understanding of how to maximize benefits of genomic selection mean adoption is poised to rise dramatically.

Key words: genomic selection, goat, sheep

Sheep and goat breeds are selected worldwide for meat, wool, and dairy production, and breeding objectives also include other functional traits such as reproductive performance and disease resistance. In early 2007, the development of next-generation sequencing (NGS) allowed de novo sequencing of sheep (Jiang et al., 2014) and goat genomes (Dong et al., 2013). In turn, it offered an opportunity to create high-density SNP chips. The Illumina OvineSNP50 BeadChip (www.illumina.com) is a 54K SNP microarray that was developed as part of the International Sheep Genomics Consortium (ISGC; www.sheephapmap.org; Kijas et al., 2009). Similarly, the International Goat Genome Consortium (IGGC; www.goatgenome.org) was created in 2010 and promoted international effort toward the development of a 52K SNP chip for goats (Tosser-Klopp et al., 2014) commercialized by Illumina (SNP50 BeadChip; www.illumina.com). The availability of such high throughput DNA methods and tools in recent years has opened up the use of genome-wide information for sheep and goat breeding.

Genomic Evaluation in Large Commercial Breeds: State of the Art Using 50K

Genomic selection (GS) based on phenotypic, genotypic, and pedigree data opens new perspectives for breeding programs in ruminants. This is especially true for dairy cattle species where selection of sires for dairy traits is hampered by a progeny testing period and genomic schemes have been immediately profitable. Genomic selection in ruminants can also be useful for meat production breeds, especially for traits that are measured later in the life of the reproductive females, such as reproductive ability, breeding seasonality, and longevity as well as for invasive or destructive evaluation measures of carcass composition and meat quality, which are typically recorded on the relatives of selection candidates and require animals to be sacrificed (Daetwyler et al., 2012b). Compared with the use of genomic information for cattle, the higher cost of genotyping relative to the value of the animal is still a strong economic barrier to the uptake of such new technology in sheep and goat breeding. Similarly, many valuable traits can
be measured in both sexes before reproductive maturity (growth, ultrasound carcass measures, and some disease resistance measurements) so that the potential to accelerate genetic progress is also less compelling.

The feasibility of genomic selection in small ruminants has been evaluated recently in meat sheep in Australia (Daetwyler et al., 2010, 2012a, 2012b) and New Zealand (Auvray et al., 2014), in dairy sheep in France, and in dairy goats in France (Carillier et al., 2013, 2014) and the UK (Mucha et al., 2015). One of the key underpinning features of genomic selection is that a reference population should be created, whereby dense phenotyping occurs for animals that are genetically related to the wider population to link the genotypic information with the phenotype. Except for New Zealand, which has 13,420 pure (mostly Romney) and crossbred sheep, the reference population sizes are still rather limited when compared with cattle, with around 1,900 Western Pyrenees dairy sheep breeds (Legarra et al., 2014); around 2,400 and 2,700 UK and French goat populations, respectively; 4,800 Lacaune dairy sheep (Larroque et al., 2014); and up to 8,000 multi-breed Australian meat sheep (Daetwyler et al., 2010, 2012a, 2012b). Within-country reference populations are generally composed of various breeds and crossbreeds. Purebred populations reached at maximum about 5,300 for New Zealand Romney, 4,000 for Australian Merino, and 4,800 for French Lacaune, with all other populations being in the range of a few hundred to 2,000. Despite small reference populations, genomic best linear unbiased prediction (GBLUP) resulted in greater accuracies of EBV than pedigree-based BLUP although for some traits and population, the increase in accuracy was small. Gains in GEBV accuracy in the French and UK dairy goat populations amounted to 0.06 for milk yield and 0.14 for fat and protein content (Carillier et al., 2014). Daetwyler et al. (2012b) showed how the gain in accuracy was well correlated with the reference population size and the genomic heritability of the trait, thus suggesting that accuracy and expected genetic gain can increase in the future if reference populations increase in size.

The gains of reliability provided by molecular information were lower than for cattle with respect to reference population size, which is probably due to lower linkage disequilibrium (LD) due to higher effective population size and inclusion of crossbreds in sheep and goats. The extent of LD estimated by average $r^2$ values between adjacent markers (50kb) ranged from 0.10 to 0.18 for Saanen and Alpine goat populations (Carillier et al., 2013; Brito et al., 2015; Mucha et al., 2015) and were mostly between 0.08 and 0.12 in sheep (Baloche et al., 2014; Kijas et al., 2014). Soay sheep (Kijas et al., 2014) and boar goat (Brito et al., 2015) were an exception with higher LD (0.28 $< r^2 < 0.30$), which is probably due to low primary effective population size. The extent of LD was therefore lower than comparable estimates in Holstein dairy cattle ranging from 0.18 to 0.3 (de Roos et al., 2008; Habier et al., 2010) and in Landrace, Duroc, Hampshire, and Yorkshire pigs (from 0.46 to 0.36; Badke et al., 2012). The LD results indicate that, for some breeds, the addition of new genotypes is mandatory and that a denser SNP panel than the current 50K Beadchip could be beneficial.

Also, genomic evaluation methods can substantially improve the accuracies of GEBV estimation when applied to small ruminants and therefore accelerate response to selection. Indeed, the accuracy of methods that use only phenotypes of the genotyped animals and ignore records of the non-genotyped part of the population (e.g., GBLUP and BLUP-SNP) is limited when the reference population is small. Therefore, a single-step approach is the recommended method for such small reference populations. It integrates all of the available phenotypic, pedigree, and genomic information in a single-step procedure to calculate genomic breeding values (Legarra
et al., 2009; Misztal et al., 2009). It also avoid bias in the estimation of GEBV due to the preselection of candidates. The method is easy to implement as it can use raw phenotypic records without the need to calculate deregressed proofs, whereby records are adjusted to reflect the fact that there are different amounts of information between animals, coming from relatives. It also allows all animals to be evaluated (with and without genotypes) simultaneously. The single-step approach improved prediction accuracy of candidates from 22 to 37% for both Alpine and Saanen goat breeds compared with the two-step method (Carillier et al., 2014). The gain in accuracy when comparing traditional pedigree-based genetic evaluations and single-step genomic evaluations for milk production traits was also significant, i.e., from 5% up to 30%, in Western Pyrenees dairy sheep breeds despite very low reference population sizes (Legarra et al., 2014).

Given the diversity of meat and dairy sheep and goat breeds, and small population size for most of those breeds, multi-breed genomic evaluations have been preferred. The benefits from blending different breeds with similar breeding objectives and recorded traits were highly variable but generally limited. Auvray et al. (2014) concluded that training datasets with Romney, Coopworth, and Perendale animals all together usually predicted better than using just a pure breed training dataset for all traits except for a few traits in Perendales. In goats, Carillier et al. (2014) compared several models: a multi-breed model blending the two breeds together, a per-breed model, and a multi-trait model considering each trait in a breed correlated to a similar one in the other breed. They found the best regression coefficients were obtained with the per-breed model. Further, Daetwyler et al. (2012) found that accounting for the structure of their large multi-breed and crossbreed sheep population generally decreased the accuracy of across-breed genomic predictions. Accordingly, and because of limited persistence of LD phases between breeds (Baloche et al., 2014; Carillier et al., 2014), a denser SNP panel than the current 50K Beadchip, or imputation from sequence data in key ancestors, might be beneficial if one expects substantial gain in accuracy for multi-breed genomic evaluations. For some breeds that are bred in several countries with similar breeding objectives, blending populations on an international basis could be highly profitable, but this depends on the level of genetic connection (by commercial exchanges) between populations. This might be the case for Texel meat sheep (Ireland, UK, France, and New Zealand), Saanen goats and crossbreds (France, UK, Italy, and Canada), and Boer goats (Canada, Australia, and France) and has already been positively evaluated for some of a set of Western Pyrenees dairy sheep breeds in Spain and France (Legarra et al., 2014).

Impact for Breeding Schemes and Expected Benefits

Based on modeling genomic versus traditional French sheep and goat breeding schemes, Shumbusho et al. (2013) showed that annual genetic gains were up to 17.9% greater with genomic information when a reference population of 2,000 individuals was used. Authors (Shumbusho et al., 2015) further showed that the overall efficiency, including monetary inputs and outputs, was moderate but increased in some meat sheep genomic scenarios, when considering the reference population available (at no costs). Because of historic genotyping costs, these simulations were based only on genotyping male selection candidates. Alternative strategies, combining cheaper low-density chips or genotype by sequencing together with imputation, is an interesting perspective to maximize benefits of genomic selection.

In addition, as mentioned by Daetwyler et al. (2012b), some sheep breeders use juvenile in-vitro fertilized embryo transfer (JIVET), which consists of harvesting immature oocytes from 6- to 8-wk-old ewe lambs and implanting these into sexually mature individuals after in vitro fertilization. The combination of genomic selection and JIVET has considerable potential to increase genetic gain for novel traits.

Including Major Genes

It is noteworthy that a large number of major genes have been identified in sheep and goat populations. Such major genes were associated with various reproductive, disease, or production traits of interest to the breeders. They include as examples: GDF8 for muscling (Clop et al., 2006), BCO2 for yellow fat (Väge and Boman 2010), Bmp15 Gdf9 and FecL genes for hyperprolificacy in sheep (Demars et al., 2013; Martin et al., 2014), Prp for scrapie resistance in sheep (Elson et al., 1999) and goat (Barillet et al., 2009), Tmem154 for resistance to MAEDI-VISNA (Heaton et al., 2012), Socs2 for mastitis susceptibility (Rupp et al., 2015) in sheep, casein genes for protein content in goat milk (Leroux et al., 1990), and a 11.7-kb deletion for polledness in goat (Pailhoux et al., 2001). A few were identified from candidate gene studies (Bmp15, Prp, and caseins) and also more recently from GWAS using the 50K SNP chip (Socs2 and Tmem154). With the decrease of sequencing and genotyping costs, and the increase of genomic studies in small ruminants, it is expected that many more major genes and causal mutations will be available in the near future. Some of those genes are already used in breeding programs such as PrP (sheep worldwide), FecL (Martin et al., 2014), or the α-s1 casein gene (French goats), essentially to pre-select candidates for progeny testing. The availability of small sets of parentage SNP or low-density chips including major gene information could allow extending such a pre-selection approach for other genes and populations. For major genes and QTLs with large effects that are correlated with a selected trait, however, their inclusion in genetic evaluation models should be promoted to avoid bias. Indeed, Martin et al. (2014) confirmed such bias in estimated breeding values for prolificacy when ignoring the existence of the FecL major gene segregating in their Lacaune meat sheep population. Altogether, these reinforce the opportunity and need for more studies on the inclusion of major gene and large QTLs into genetic and/or genomic evaluation in small ruminants. Issues include predicting ungenotyped animals, especially in the single-step GBLUP approach, modeling multi-allelic variants or QTL haplotypes, and combining results from many genes and QTLs. Carillier et al. (2015) recently proposed four approaches to include the α-s1 casein genotype in genetic and genomic evaluations. The best model, called Gene Content Multiple trait BLUP and proposed initially by Legarra and Vitezica (2015) in a biallelic loci case, showed an increase in predictive ability from 1 to 16% for protein content. While there are opportunities offered by the addition of major genes in genetic/genomic evaluations, it raises the question of defining breeding goals when those major genes have a pleiotropic effect, such as the socs2 mutation, which is unfavorably associated with high milk somatic cell counts while also being positively correlated with higher milk production (Rupp et al., 2015).

Using SNP Chips for Parentage and Assignation

In most sheep and goat breeding schemes, parentage is an issue because the use of AI is unequal and limited (on average, 50% in goats and 23% in meat sheep in France, for example, and much lower elsewhere). Pedigree
information is typically unknown in systems that use natural matings based on multiple-sire natural mating groups of rams for breeding ewes or in extensive systems. The accuracy and completeness of pedigree is, however, an essential feature increasing the rate of genetic gain. Markers of DNA, first microsatellites and now SNPs, have proved useful to infer pedigree information. These markers can be used both to detect mis-identification of parents and to assign true parents among candidates using a likelihood-based approach. The latter approach is widely used in New Zealand in extensive multiple mating systems (Dodds et al., 2005). Similar to cattle, no less than six ovine sets of parentage SNPs derived from the ovine SNP50k BeadArray have been proposed to date. They include: 88 SNPs from the ISGC (http://www.sheep-hapmap.org/), 84 and 300 SNPs from AgResearch in New Zealand (Clarke et al., 2014), 109 SNPs from the Meat Animal Research Center (USMARC) in the US (Heaton et al., 2014), 382 SNP from CSIRO in Australia (Heaton et al., 2014), and 192 SNP from INRA in France (Tortereau et al., 2015). However, overlap between these sets is limited. As an example, the US and French sets had only 44 and 0 SNPs in common with the ISGC set. Because of the numerous commercial or local breeds of meat and dairy sheep, it is probably not realistic to target a unique limited set of SNPS that can be used worldwide although Heaton et al. (2014) showed good performance of their sets in globally diverse breeds of sheep. In goats, due to the introduction of the Illumina Caprine 50K BeadChip in 2011, similar initiatives are in progress at INRA (French Saanen and Alpine and local breeds), in Italy, and at SRUC (UK crossbred goats). The challenges are now to produce those tools at low cost and encourage large-scale use by breeders. Currently, genotyping of SNP parentage sets are based on the MassARRAY platform (Sequenom) technology, but a multi-species parentage BeadArray chip could also be an option that could considerably reduce costs. Preliminary survey and simulation suggest that a cost competitive price would be less than 10 euros (including DNA extraction and reporting). An alternative approach to parentage plus key loci is to reduce the cost of low-density genotyping sufficiently, perhaps less than 15 euros, so that all progeny are genotyped and included in the genomic evaluation. This approach is particularly attractive in extensive sheep systems already using multi-sire mating and DNA parentage assignment. Developments in other animal species suggest that such an approach is now nearing commercial reality (Dodds et al., 2015). Such initiatives are currently promoted by both the International Sheep and Goat Genomics Consortia.

Management of Diversity

Genomics offers ample opportunities, not only for genetic improvement of animals, but it can be also used as an important tool for the assessment of genetic diversity of local sheep and goat breeds. Both SNP chips and whole-genome sequence data can provide much more accurate estimates of relationship between animals than pedigree records (Eynard et al., 2015). This is particularly important in cases when accurate pedigree records are unavailable or incomplete. In such cases, decisions about which animals to conserve can be very biased and may lead to loss of genetic diversity. This has wide implications both for in vivo and in vitro (genebank) conservation (Mucha and Windig 2009). More ordinarily accurate estimation of co-ancestry can allow better mate assignment and thereby avoid accidental inbreeding. Moreover, sequence data can provide access to information on rare variants carried by individuals, which is unavailable when using pedigree or SNP chips (which contain mainly the common variants). This is of particular interest from the conservation point of view to preserve the widest amount of genetic diversity.

Genomic Tools for Improving New Traits Related to Resilience of Small Ruminants

Genomic selection also has potential to increase the resilience of small ruminant enterprises. These include breeding for disease resistance such as facial eczema (Phua et al., 2014) and parasite and fly-strike resistance (Pickering et al., 2015). Genomic selection in these cases also has ethical benefits...
in reducing the number of animals that need to be exposed to the disease as well as reducing suffering in future generations. Similarly, work is now under way to allow genomic prediction for traits such as feed efficiency and methane emissions. In both cases, measurement of these traits involves considerable cost and facilities unlikely to be broadly deployed on farms, making genomic selection an obvious alternative (Pickering et al., 2015).

**Literature Cited**


Rachel Rupp is a senior scientist at the Animal Genetics Department of the French National Institute for Agricultural Research (INRA). She specializes in quantitative genetics applied to small ruminants. Her main area of study has been in disease resistance and genomics applied to dairy sheep and goats. She has developed multidisciplinary research and implemented selection for mastitis resistance within several national and European projects. She has been involved in the recent advances in goat genomics, including the development of the 50K SNP chip, QTL detection, and first evaluation of genomic selection. She has lead EU Seventh Framework Programme 3SR project “Sustainable Solutions for Small Ruminants.”

Correspondence: rachel.rupp@toulouse.inra.fr

Sebastian Mucha is a post-doctoral researcher at the Scotland’s Rural College (SRUC). He undertakes genomic research for applied livestock programs in sheep and goats. He mainly focuses on implementation of genomic selection for production and functional fitness traits in dairy goats. Other areas of interest involve conservation of genetic diversity and QTL detection in sheep and goats.

Correspondence: sebastian.mucha@sruc.ac.uk

Hélène Larroque is a quantitative geneticist with 20 years experience in genetic evaluation of dairy ruminants for dairy and type traits at the Animal Genetics Department of the French National Institute for Agriculture Research (INRA). She has focused more recently on new aspects of genomic selection in dairy sheep and goats. She conducts research on genetic variability of traits related to once-daily milking ability of animals, on cheese making ability of milk, on fine milk composition, and on persistency. She’s involved in the EU iSAGE (Innovation for Sustainable Animal Production in Europe) project, for an innovative management of genetic resources and breeding solutions to cope with future environmental, climatic, and socio-economic challenges.

Correspondence: helene.larroque@toulouse.inra.fr

John McEwan is a principal scientist in the New Zealand AgResearch Animal Genomics team. He has specialized in quantitative genetics and genomics in sheep and deer. His main area of study has been in meat traits and disease resistance and genomics applied to sheep. He was involved in the international consortia physically mapping and then sequencing the cattle, sheep, and deer genomes. Subsequently, he used the results to design a variety of low- and high-density SNP chips in sheep in order to research and then implement genomic selection in New Zealand maternal and terminal sire sheep breeds. Current research is investigating the host genetics and genomics of methane emissions in sheep and developing new low cost genotyping by sequencing strategies in a variety of plant and animals species.

Correspondence: john.mcewan@agresearch.co.nz

Joanne Conington is a Reader in Applied Livestock Genetics in the Animal and Veterinary Sciences research group, Scotland’s Rural College (SRUC, previously known as SAC). Her research is focused on the development of new, broader breeding goals for multi-trait breeding programs incorporating aspects of disease resistance and maternal efficiency. Her current research is mostly undertaken in collaboration with industry partners investigating genetic and genomic approaches to breed for resistance to mastitis in sheep and goats, footrot and internal parasites in sheep, lamb survival, and ewe longevity. She lectures to under- and post-graduate students at SRUC and the University of Edinburgh and also advises the UK government on animal breeding issues relating to animal welfare.

Correspondence: Joanne.Conington@sruc.ac.uk