International Genomic Co-operation; Who, what, when, where, why and how?

A. R. Cromie[1], D. P. Berry[2], B. Wickham[1], J.F. Kearney[1], J. Pena[3], J.B.C.H. van Kaam[4], N. Gengler[5], J. Szyda[6], U. Schnyder[7], M. Coffey[8], B. Moster[9] , K. Hagiya[10], , J. I. Weller[11], D. Abernethy[12] and R. Spelman[13]

[1]Irish Cattle Breeding Federation, Ireland, [2]Teagasc, Moorepark, Ireland, [3]CONAFE, Spain, [4]Anafi, Cremona, Italy, [5]University of Liegge, Belguim, [6]Department of Animal Genetics and Institute of Natural Sciences Wrocław University of Environmental and Life Sciences, Wrocław, Poland, [7]Swiss herdbook, Zollikofen, Switzerland, [8]SAC, Bush Estate, Penicuik, Midlothian, Scotland, EH26 0PH, [9]ARC-Animal Production Institute, P/bag X2, Irene, 1675, South Africa, [10]National Livestock Breeding Center, Fukushima, Japan, [11]Institute of Animal Sciences, ARO, The Volcani Center, Bet Dagan, Israel, [12]Austalian Dairy Herd Improvement Association, Melbourne, Australia[13]Livestock improvement, Hamilton, New Zealand.

Introduction

In 1995 Interbull undertook its first international MACE genetic evaluation in Holsteins across 9 countries. The concept of such a service began as far back as 1975 but its urgency intensified with increased international trade of frozen dairy bull semen. The objective of Interbull was to utilize all available data on individual sires within each member country to generate more accurate predictions of genetic merit on the individual scales of the different countries.

Whilst there was sharing of national proofs before the introduction of MACE, the handicap was that it involved a lot of extra work, including the exchange of relevant files and development of conversion equations between participating countries. Despite this extra work requirement, there were still major concerns when MACE was first introduced, primarily around the accuracy of these new proofs and the more direct involvement of Interbull in the process. Today, these concerns have all been allayed, with 30 countries now participating in the service and each considering MACE evaluations as the standard tool for selecting bulls on both a within and an across country basis. Compared to the previous conversion-based approach, MACE is better statistically, it saves time and money for individual members and it guarantees the same method and rules for all participating countries.

Today there is some concern that a new type of data, invaluable in the generation of genomic breeding values, should be shared between participating countries and/or at Interbull. The consequence is that many countries are now participating in bi-lateral and multi-lateral sharing arrangements, with consequential effects on time, effort and money. Clearly, given our collective experiences with MACE and the benefits of open collaboration in this area there is much scope for more discussion on this subject.

The objective of this paper is therefore to review the current state of the art in national genomic evaluations and international cooperation in genomic evaluations, and to put forward some points for discussion. The paper will be structured on the five “W”s; who, what, when, where, why and will conclude with “how”.

Who is collaborating?

Although still early days in the implementation of genomic selection, the results emanating from the countries that implemented genomic selection indicate that genomic selection, or some derivative of such, will be the new system for identifying genetically elite animals. Therefore, any country with a national genetic evaluation system and a mission statement of providing the most accurate genetic evaluations possible to its stakeholders should be interested in cooperating on international genomic evaluation.

Some adversaries to genomic cooperation may think that cooperation in genomics is only beneficial to smaller countries that import semen. However, access to the best software and training population to undertake genomic

evaluations is not only vital for importing countries but is also crucial for the accurate evaluation of bulls being marketed by exporting countries.

In December 2008 the Interbull Center undertook a survey on genomic selection across all member organisations. All 31 organisations answered. At that time, 5 groups of countries were co-operating in genomic evaluations: 1) Denmark-Finland-Sweden, 2) Canada and the USA, 3) The Netherlands and New Zealand, 4) Ireland and New Zealand, and 5) Austria and Germany. An additional 8 other countries indicated that they would probably collaborate in the future, totalling 17 countries in all. Since this survey, it is interesting to note that some of the countries that did not indicate, at that time, that they were willing to cooperate are in fact now swapping bull genotypes. Therefore, it is clear that the appetite for cooperating in genomic selection breeding programs is present and appears to be increasing. It is obvious that all countries should be interested in collaboration on genomic evaluations to ensure each country has access to the best expertise and resources available and ultimately the best genetics.

What are they collaborating on?

Collaboration comes under many guises: 1) expertise and knowledge, 2) software, 3) phenotypes, and 4) genotypes. All 4 types of collaboration are underway in dairy cattle breeding. Expertise and knowledge are being shared at Interbull workshops and other international conferences. Software is being shared, for example through genomicseletion.net (http:///www.genomicselection.net ; Coffey and Mrode, 2009). Phenotypes, in the form of national proofs are being shared through Interbull, although only on traits for which international genetic evaluations are undertaken. A more recent development has seen the sharing of sire-dam pedigree files, as part of each country's efforts to maximise the accuracy of MACE evaluations. Genotypes are being shared by some countries through bilateral agreements, or through consortia such as the North American Consortium, EuroGenomics or Intergenomics; expertise and computer code is also shared through these agreements.

The question is what ideally should be swapped? When asked in the Interbull survey on what data each country would be willing to supply to Interbull, only Ireland was willing to supply any relevant data although many countries did not appear to respond to this question, probably because it was not already defined as to who would own the genotype information in those countries. Given the likely change in perception of genomic selection in the past 18 month, would a re-run of the survey obtain a different set of results?

Using the simulation study of VanRaden (2009) based on 8,193 Brown Swiss bulls from 9 countries, Sullivan and VanRaden (2009) evaluated alternative methods of international genomic evaluations. Although the differences were relatively small, in the majority of countries multi-trait genomic evaluations achieved higher accuracy than G-MACE at that time. However, the G-MACE methodology has since been refined (Sullivan and VanRaden, 2010). Nonetheless, theoretically at least, sharing of all genotypic and phenotypic data should be optimal when deriving international genomic evaluations.

When did the collaborations commence?

In early 2008, the US and Canada joined forces to create a large training population for Holstein-Friesians. Clubware, the collaboration on software development began in October 2008. In December 2008, Ireland and New Zealand shared 3,000 genotypes of Holstein-Friesian bulls. In early 2009, discussions between CRV, DHV/VIT, UNCEIA and VikingGenetics began on the development of an agreement which was later called EuroGenomics (David *et al.*, 2010); exchange of genotypes began in the fall of 2009. Intergenomics, an international collaboration on genomic selection for Brown Swiss began in 2009.

March 2009 saw the beginning of an initiative to generate a file that contained all the Holstein-Friesian bulls genotyped in the world with the Illumina Bovine50 Beadchip, as

well as those with DNA available for genotyping. This was an attempt by a number of countries to avoid further duplication in genotyping and to facilitate the process of sharing genotypes across countries, for those countries interested in pursuing this objective. This file to-date contains 36,898 bulls from 13 different countries (US and Canada are treated as one "country"), and will be reported on at the Interbull business meeting in Riga by Berry *et al.*

Where is the collaboration happening?

To-date most collaboration has been undertaken through bilateral agreements between national evaluations centers, herdbooks, and AI companies as well as through multi-lateral agreements between a small number of countries. In the December 2008 survey of member countries undertaken by Interbull, a question was asked on the expected role of Interbull in genomic evaluations. Of the 18 countries that made suggestions, 8 countries mentioned that Interbull should either store the genotypes or facilitate/coordinate the exchange of genotypes. Three countries, Australia, Ireland and New Zealand suggested that Interbull should store all genotypes presumably to undertake international genomic evaluations using all available genotypes.

Why is the collaboration happening?

There are many reasons as to why collaboration in sharing of genotypes is underway. These include:

Increasing the reference population. The motivation for international genomic evaluations now is identical to the motivation for the establishment of Interbull over 35 years ago. The most accurate proofs, be they genetic or genomic, are fundamental to success for both importing and exporting countries alike. Between Ireland, North America, Switzerland, Poland, Israel, Italy, Japan, New Zealand and Australia, 30,643 Holstein bulls have been genotyped of which 20,739 have a MACE evaluation for milk production. Canada reported an increase in accuracy of 0.16 using a training population size of 9,300 genotypes with domestic (~2000 bulls) and MACE proofs (Muir *et al.,* 2010). Using a very simplistic calculation, assuming the marginal benefit in accuracy of increasing the training population size from 9,300 to 20,739 is the same as from zero to 9,300 genotypes this equates to an overall increased accuracy of ~0.32 from using a training population size of 20,739 genotypes. VanRaden *et al.* (2009) suggested that reliability of genomic evaluations could increase to over 80% if the reference population was >40,000 proven bulls.

Table 1. The benefit in reliability of DGVs of including genotyped bulls with MACE evaluations in the Irish national genomic evaluations.

	Rel.
Parent Average Only	0.302
Genotypes + Domestic Proofs	0.418
Genotypes + (Domestic & MACE proofs)	0.523

Incorporating genotyped bulls with MACE evaluations into national genomic evaluations has been shown in many populations to increase the accuracy of selection (Muir *et al.,* 2010; Schrooten and van der Linde, 2010). Table 1 summarises the benefit of including genotyped bulls with MACE evaluations for milk protein yield over and above only including genotyped bulls with domestic evaluations in Ireland. The majority of genotypes used in genomic selection in Ireland have been attained through swapping with New Zealand, Switzerland, Poland, the UK, and Italy. In the most recent evaluation run (May 2010), there were 3,660 genotypes included in the training population for milk production and 4,561 genotypes included in the full run. It is prudent to note that some 70% of the training population in the Irish evaluation was obtained from bi-lateral sharing arrangements.

Avoid repeated genotyping. Genotyping of animals is now relatively cheap and the technology platforms for genotyping are available to all. Therefore, there is little competitive advantage other than who has the most money. However, as previously alluded, money can be saved by sharing of information on animals genotyped and subsequent

swapping of these genotypes. Between Ireland, Poland, Switzerland, Japan, Israel, Italy, Australia and New Zealand, 257 Holstein-Friesian bulls have been genotyped more than once. Including the US and Canada the number of bulls genotyped more than once increases to 522. Among these 522 bulls, 401 have been genotyped twice, 84 have been genotyped three times, 27 have been genotyped four times, 8 bulls (Silky Gibson, Silky Cousteau, Bosside Ruben, Sandy Valley Forbidden, Roylane Jordan, Pursuit September Storm, Ricecrest Brett, and Sikkema-Star Air Magna) have been genotyped 5 times and 2 bulls (Jocko Besne and O-Bee Manfred Justice) have been genotyped 6 times. Assuming a cost of €160 per genotype (including procurement of semen and DNA extraction) this amounts to a squandering of over €111,000 or US$138,000. However, this analysis does not include the genotypes of EuroGenomics which are likely to have considerable overlap. Therefore the squandering of funds of all genotypes analysis is likely to be at least double.

Smaller breeds: Cooperation in genomic evaluation to achieve high accuracy of selection is likely to be greatest in smaller breeds because their global population will be considerably bigger than the respective populations in each individual country. This is the rationale behind the highly successful Intergenomics project for the Brown Swiss breed (Jorjani *et al.*, 2010).

Algorithms and software. Algorithms used in genomic selection differ among countries but details on those used are, in general, in the public domain. Irrespective, access to the best algorithms for genomic evaluation are key for each country to ensure that the GEBVs entering any international genomic evaluation are of the highest quality and rigorously tested. Some software for undertaking genomic selection is on the Clubware website (http://www.genomicselection.net).

Traits difficult or expensive to measure. One advantage of collaboration in genomics often ignored is the ability to generate large populations for traits difficult or expensive to measure, such as animal health, product quality and greenhouse gas emissions. Surely dairy cattle breeders have learnt from past experiences where aggressive selection for a particular breeding goal had serious repercussions for other traits that were not measured at that time. The impact of such actions cost the dairy industry untold amounts and its effects are still being felt internationally. Pooling of resources across countries to measure such traits will lessen the possibility of reducing genetic merit for traits known to affect profitability, but not yet measured in sufficient numbers. The increase in accuracy of selection through genomic selection will indeed result in faster genetic gain for the traits under selection, but may also result in faster deterioration in traits not routinely measured. How this could work may be illustrated using the example of Bovine Tuberculosis (TB). Research in Ireland has shown that Bovine TB is heritable (Bermingham *et al.*, 2009) and breeding values have been estimated for TB susceptibility. Several hundred Holstein-Friesian sires with breeding values for TB have been genotyped with the Illumina Bovine50 Beadchip. Knowledge of the selection intensity placed on each SNP in international breeding goals can be used to estimate the expected response to selection in genetic merit for susceptibility to TB.

In theory, phenotypic files could also be made available to other participating countries in the future, to enable those countries start selecting for traits of interest. This would require the country to make some assumption regarding G*E interactions, but they would, at least, be in a position to start selecting for (or against) these important new traits. This is an example of where co-operation in genomics can and will move far beyond the routine sharing of genotypes, which is the focus of current discussions.

Higher density chip and/or the sequencing of selected animals. Going forward, decisions will have to be made regarding the re-genotyping of animals on a higher density chip and/or the sequencing of selected animals. Clearly having visibility of all animals in one common file would be advantageous for this process. Who pays for these ongoing costs is a relevant point that must be considered; individual countries that are prepared to share, co-operation through bi-lateral agreements, or is a more defined multi-country system required?

Why is the collaboration not happening?

There are also many reasons why genomic cooperation should not be pursued. Arguably the main reason relates to investment. Some companies/countries have made considerable investments in the genotyping of animals and acquirement of resources to undertake accurate genomic evaluations. Obviously these countries require a return on investment. This can be achieved through having (almost) exclusive access to the most accurate international genomic evaluations. By simply making a "pot" of genotypes other organisations can access these resources without any prior investment. However, these genotypes will be used by the "smaller" countries to generate the most accurate genomic evaluation for their respective countries, although this will also include a greater accuracy of selection for local bulls and dams which will subsequently compete with the importing market. Nevertheless, the most accurate genomic evaluations possible in importing countries can only be beneficial for the large exporting countries.

How will we collaborate?

There are many different options for collaboration including: 1) sharing of all genotypes and expertise amongst everyone, 2) sharing of genotypes and expertise among members of a consortium with or without similar resource input into the initiative, 3) swapping of genotypes in bilateral agreements, 4) swapping of information on genotyped animals and animals where DNA or biological material is available. The latter three types of collaboration are already underway. Several consortia have been developed to share both genotypic information and skills in genomic evaluations. These include the North American consortium, EuroGenomics, and Intergenomics. In February 2010 the North American consortium had access to 37,409 genotypes on Holstein males and females of which 8,991 were of proven Holstein bulls (Muir *et al.*, 2010). In March 2010, EuroGenomics had over 16,000 Holstein bulls in their reference population (David *et al.*, 2010).

Also, bilateral agreements on swapping of genotypes or semen samples have been undertaken by many pairs of countries including Ireland, Poland, Italy, Switzerland, New Zealand and the UK. More importantly, the list of animals genotyped will be extremely helpful to those countries that have not yet genotyped (e.g., Spain, the UK, Belgium, South Africa) as they decide which bulls to genotype to maximise their training population through swapping. It is also useful information for all countries when deciding on which bulls to genotype in the future. Funding is always tight so the simple availability of such information is key to obtaining the best return on investment. The list of genotyped animals, and partnerships built, will also be useful in deciding which bulls should be genotyped/sequenced at greater densities.

Arguably the best method to maximise the achievable accuracy of genomic evaluation in all countries is to openly share all genotypes. One such proposition has been called IGenoP (International Genomic Evaluation Partnership).

IGenoP. Arising from the creation of the list of genotyped animals and various bi-lateral swapping agreements, it was obvious that some countries favoured an approach of more open collaboration. The motivation for IGenoP is therefore to facilitate national cattle animal evaluation units in the provision of genomic evaluations on their national base and scale using freely available genotypes. The initiative is open to all national cattle animal evaluation units that are members of Interbull.

Although the IGenoP concept is still in the development phase, a number of important principles have been identified by the partnering countries. These include:

- Open sharing of knowledge, tools and expertise.
- That GEBVs should be made available to all breeders availing of the service.
- Bull genotypes would be hosted at the Interbull Centre.
- Each partner contributes all owned genotypes to the pool.

- Partners can use the genotypes in the shared pool for all relevant research, development and implementation.
- A sharing of knowledge regarding software development, e.g., Clubware.
- Partners will not provide genotypes to third parties.
- A definite commitment to only publish GEBVs on their country's base and scale.

To-date 12 countries have expressed interest in becoming involved in the IGenoP initiative. These are; Ireland, UK, Poland, Italy, South Africa, Spain, Japan, Switzerland, Israel, Belgium, New Zealand and Australia – all of which are listed as co-authors in this paper. Collectively these countries represent some 13,285 bulls with genotypes, of which 11,801 have MACE evaluations (another 6,871 proven bulls are due to be genotyped by these countries in the coming months). Other countries that are interested in becoming involved in the IGenoP initiative should make contact with any of the authors from this paper.

Although still at an early stage, some issues have been identified by the partnering countries for further discussion. For example, will all participants be able to undertake required genomic evaluations for their respective country? If not, what are their options? What about sharing of phenotypes and is the sharing of software necessary? Also, should there be some minimum contribution of genotypes by a participating country to ensure participation? Clearly these are all issues that merit discussion.

Some people have also asked as to whether IGenoP is a "competitor" to other initiatives that involve the sharing of genotypes. The simple answer is No. Instead, IGenoP should be viewed as simply another step in the process of ensuring greater international genomic co-operation and hence more accurate genetic evaluations for participating countries.

Avoiding duplication is a strong motivational desire within the current partners. Whilst this relates to the cost of genotyping, it also relates to co-operation agreements between the various countries (i.e., bi and multi-lateral sharing arrangements), all of which take up time, effort and money. It is for these reasons that the IGenoP partners are committed to seeing these functions (i.e., lists, genotypes and phenotypes) placed within the Interbull centre and within the various steering and technical groups, of which we are participating members. By doing so, each country is making best use of its national resources, whilst also remaining committed to its overall objective of providing the most accurate genetic evaluations for the country that it represents.

Conclusions

There is no denying the benefits in accuracy of selection achievable through collaboration. Several types of collaboration in genomics are currently underway among different bodies each requiring their own resource input from already strained national genetic evaluation bodies. There is an opportunity now to develop a streamline international genomic evaluation process to maximise the gain achievable by all, for the benefit of all. Finally, advancements in both the methodology of genomic selection but more importantly the perspective of genomic selection and collaboration has changed. Therefore, it might be worthwhile re-evaluating through another Interbull survey, individual countries' thoughts on genomic selection and international collaboration. This is especially prudent as we move towards the introduction of a new Interbull service, based on GMACE.

References

Bermingham, M.L., More, S.J., Good, M., Cromie, A.R., Higgins, I.M., Brotherstone, S. & Berry, D.P. 2009. Genetics of tuberculosis in Irish Holstein-Friesian dairy herds. *J. Dairy Sci. 92,* 3447-3456.

Coffey, M. & Mrode, R. 2009. Clubware – A collaborative genomic breeding value estimation software development platform. *Interbull Bulletin 39,* 73-76. UPPSALA, SWEDEN. January 26-29, 2009.

David, X., de Vries, A., Feddersen, E. & Borchersen, S. 2010. International genomic cooperation - European perspective.

Interbull Bulletin 41. Paris, France. March 4-5, 2010.

Jorjani, H., Zumbach, B., Dürr, J. & Santus, E. 2010. Joint genomic evaluation of BSW populations. *Interbull Bulletin 41*. Paris, France. March 4-5, 2010.

Muir, B., Van Doormaal, B. & Kistemaker, G. 2010. International genomic cooperation – North American perspective. *Interbull Bulletin 41*. Paris, France. March 4-5, 2010.

Schrooten, C. & van der Linde, C. 2010. Multi-country genomic selection. *Interbull Bulletin 41*. Paris, France. March 4-5, 2010.

Sullivan, P.G. & VanRaden, P.M. 2009. Development of Genomic MACE. *Interbull Bulletin 40,* 157-161. Barcelona, Spain, August 21-24, 2009.

Sullivan, P.G. & VanRaden, P.M. 2010. GMACE implementation. *Interbull Bulletin 41*. Paris, France. March 4-5, 2010.

VanRaden, P.M., Wiggans, G.R., Van Tassell, C.P., Sonstegard, T.S. & Schenkel, F. 2009. Benefits from Cooperation in genomics. *Interbull Bulletin 39,* 67-72. UPPSALA, SWEDEN. January 26-29, 2009.