

Prospects and Challenges for the Conservation of Farm Animal Genomic Resources, 2015-2025

Michael W. Bruford^{1, 2*}, Catarina Ginja^{3, 4}, Irene Hoffmann^{5, 6}, Stéphane Joost⁷, Pablo Orozco-Wengel⁸, Florian Alberto⁹, Andreia Amaral¹⁰, Mario Barbato⁸, Filippo Biscarini¹¹, Licia Colli¹², Mafalda Costa⁸, Ino Curik¹³, Solange Duruz⁷, **Maja Ferencaković**¹³, Daniel Fischer^{14, 15}, Robert Fitak¹⁶, Linn Fenna Groeneveld¹⁷, Stephen Hall¹⁸, Olivier H. Hanotte¹⁹, Faiz-ul Hassan^{19, 20}, Philippe Helsen²¹, Laura Iacolina²², Juha Kantanen^{14, 15}, Kevin Leempoel⁷, Johannes A. Lenstra²³, Paolo Ajmone Marsan¹², Charles Masembe²⁴, Hendrik-Jan Megens²⁵, Mara J Miele²⁶, Markus Neuditschko²⁷, Ezequiel L Nicolazzi¹¹, Francois Pompanon⁹, Jutta Roosen²⁸, Natalia Sevane²⁹, Anamarija Smetko³⁰, Anamaria Štambuk³¹, Ian Streeter³², Sylvie Stucki⁷, China Supakorn^{19, 33}, Luis Telo da Gama³⁴, Michèle Tixier-Boichard³⁵, Daniel Wegmann³⁶, Xiangjiang Zhan^{8, 37}

¹Organisms and Environment Division, School of Biosciences, Cardiff University, United Kingdom,

²Sustainable Places Research Institute, Cardiff University, United Kingdom, ³Centro de Ecologia, Evolução e Alterações Ambientais (CE3C), Faculdade de Ciências, Universidade de Lisboa, Portugal,

⁴Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO-InBIO), Universidade do Porto, Portugal, ⁵Animal Genetic Resources Branch, Animal Production and Health Division, Food and

Agriculture Organization of the United Nations, Italy, ⁶The views expressed in this information product are those of the author and do not necessarily reflect the views or policies of FAO., Italy,

⁷School of Civil and Environmental Engineering (ENAC, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland, ⁸School of Biosciences, Cardiff University, United Kingdom, ⁹Laboratoire d'Ecologie Alpine and CNRS, Univ. Grenoble Alpes, France, ¹⁰Faculty of Sciences, University of Lisbon - BioISI- Biosystems and Integrative Sciences Institute, Portugal, ¹¹Parco Tecnologico Padano, Italy, ¹²Instituto di Zootecnica,

BioDNA Centro di Ricerca sulla Biodiversità e sul DNA Antico, Università Cattolica del S. Cuore di Piacenza, Italy, ¹³Faculty of Agriculture, University of Zagreb, Croatia, ¹⁴Department of Biology, University of Eastern Finland, Finland, ¹⁵Green Technology, Natural Resources Institute Finland (Luke), Finland, ¹⁶Institut für Populationsgenetik, Vetmeduni, Austria, ¹⁷NordGen -The Nordic Genetic Resource Center, Norway, ¹⁸Livestock Diversity Ltd, United Kingdom, ¹⁹School of Life Sciences, University of Nottingham, United Kingdom, ²⁰Department of Animal Breeding and Genetics, University of Agriculture Faisalabad (UAF), Pakistan, ²¹Centre for Research and Conservation, Royal Zoological Society of Antwerp, Belgium, ²²Department of Chemistry and Bioscience, Aalborg University, Denmark, ²³Faculty of Veterinary Medicine, Utrecht University, Netherlands, ²⁴Institute of the Environment and Natural Resources, Makerere University, Uganda, ²⁵Animal Breeding and Genomics Centre, Wageningen University, Netherlands, ²⁶School of Planning and Geography, Cardiff University, United Kingdom, ²⁷Agroscope, Swiss National Stud Farm, Switzerland, ²⁸TUM School of Management, Technische Universität München, Germany, ²⁹Dpto. Animal Production, Veterinary Faculty, Universidad Complutense de Madrid (UCM), Spain, ³⁰Croatian Agricultural Agency, Croatia, ³¹Department of Biology, Faculty of Science, University of Zagreb, Croatia, ³²European Molecular Biology Laboratory, European Bioinformatics Institute - Wellcome Trust Genome Campus, United Kingdom, ³³School of Agricultural Technology, Walailak University, Thailand, ³⁴CIISA - Faculty of Veterinary Medicine, University of Lisbon, Portugal, ³⁵INRA, AgroParisTech, UMR GABI, France, ³⁶Department of Biology, University of Fribourg, Switzerland, ³⁷Institute of Zoology, Chinese Academy of Sciences, China

Submitted to Journal:

Frontiers in Genetics

Specialty Section:

Livestock Genomics

Article type:

Hypothesis & Theory Article

Prospects and Challenges for the Conservation of Farm Animal Genomic Resources, 2015-2025.

Michael W Bruford^{1,*}, Catarina Ginja¹⁰, Irene Hoffmann^{16***}, Stéphane Joost⁷, Pablo Orozco-Wengel¹, Florian J Alberto², Andreia Amaral³, Mario Barbato¹, Filippo Biscarini⁴, Licia Colli⁵, Mafalda Costa¹, Ino Curik⁶, Solange Duruz⁷, Maja Ferenčaković⁶, Daniel Fischer⁸, Robert Fitak⁹, Linn Fenna Groeneveld¹¹, Stephen JG Hall¹², Olivier Hanotte¹³, Faiz-ul Hassan^{13,14}, Philippe Helsen¹⁵, Laura Iacolina¹⁷, Juha Kantanen^{8,\$}, Kevin Leempoel⁷, Johannes A Lenstra¹⁸, Paolo Ajmone-Marsan⁵, Charles Masembe¹⁹, Hendrik-Jan Megens²⁰, Mara J Miele²¹, Markus Neuditschko²², Ezequiel L Nicolazzi⁴, François Pompanon^{2,***}, Jutta Roosen²³, Natalia Sevane²⁴, Anamrija Smetko²⁵, Anamaria Štambuk²⁶, Ian Streeter²⁷, Sylvie Stucki⁷, China Supakorn^{13,32}, Luis Telo da Gama²⁸, Michèle Tixier-Boichard²⁹, Daniel Wegmann³⁰, Xiangjiang Zhan^{1,31}.

1 School of Biosciences, Cardiff University, Cathays Park, Cardiff CF10 3AX, UK; * Sustainable Places Research Institute, Cardiff University, 33 Park Place, Cardiff CF10 3BA; 2 Univ. Grenoble Alpes, Laboratoire d'Ecologie Alpine, F-38000 Grenoble, France; 3 University of Lisbon, Faculty of Sciences, BIOLSI- Biosystems and Integrative Sciences Institute, Campo Grande, Lisboa, Portugal; 4 Parco Tecnologico Padano, Via Einstein, 26900 Lodi, Italy; 5 Istituto di Zootecnica, BioDNA Centro di Ricerca sulla Biodiversità a sul DNA Antico, Università Cattolica del S. Cuore di Piacenza, Italy; 6 Faculty of Agriculture, University of Zagreb, Zagreb, Croatia; 7 Laboratory of Geographic Information Systems (LASIG), School of Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland; 8 Natural Resources Institute Finland (Luke), Green Technology, Jokioinen, Finland; \$ Department of Biology, University of Eastern Finland, Kuopio, Finland 9 Institut für Populationsgenetik, Vetmeduni, Vienna, Austria; 10 Centro de Ecologia, Evolução e Alterações Ambientais (CE3C), Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal and Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO-InBIO), Universidade do Porto, Campus Agrário de Vairão, Portugal; 11 NordGen -The Nordic Genetic Resource Center, P.O. Box 115, 1431 Ås, Norway; 12 Livestock Diversity Ltd. www.livestockdiversity.com, Lincoln, UK; 13 School of Life Sciences University of Nottingham, UK; 14 Department of Animal Breeding and Genetics, University of Agriculture, Faisalabad, Pakistan; 15 Centre for Research and Conservation, Royal Zoological Society of Antwerp, Antwerp, Belgium; 16 Food and Agriculture Organization of the United Nations, Animal Genetic Resources Branch, Animal Production and Health Division, Rome, Italy; 17 Department of Chemistry and Bioscience, Aalborg University, Aalborg, Denmark; 18 Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands; 19 Institute of the Environment and Natural Resources, Makerere University, Kampala, Uganda; 20 Animal Breeding and Genomics Centre, Wageningen University, De Elst 1, Wageningen, The Netherlands; 21 School of Planning and Geography, Cardiff University, Cardiff, UK; 22 Agroscope, Swiss National Stud Farm, Avenches, Switzerland; 23 TUM School of Management, Technische Universität München, Munich, Germany; 24 Dpto. Animal Production, Veterinary Faculty, Universidad Complutense de Madrid (UCM), Madrid, Spain; 25 Croatian Agricultural Agency; 26 Department of Biology, Faculty of Science, University of Zagreb, Rooseveltov trg 6, 10000 Zagreb, Croatia; 27 European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, UK; 28 CIISA – Faculty of Veterinary Medicine, University of Lisbon, Portugal; 29, INRA, AgroParisTech, UMR GABI, Jouy-en-Josas, France; 30 Department of Biology, University of Fribourg, Fribourg, Switzerland; 31 Institute of Zoology, Chinese Academy of Sciences, Beijing, China; 32 School of Agricultural Technology, Walailak University, Thailand. ** Author for correspondence.

Footnote *** The views expressed in this information product are those of the author and do not necessarily reflect the views or policies of FAO.

Abstract

Livestock conservation practice is changing rapidly in light of policy developments, climate change and diversifying market demands. The last decade has seen a step change in technology and analytical approaches available to define, manage and conserve Farm Animal Genomic Resources (FAnGR). However, these rapid changes pose challenges for FAnGR conservation in terms of technological continuity, analytical capacity and integrative methodologies needed to fully exploit new, multidimensional data. The final conference of the ESF Genomic Resources program aimed to address these interdisciplinary problems in an attempt to contribute to the agenda for research and policy development directions during the coming decade. By 2020, according to the Convention on Biodiversity's Aichi Target 13, signatories should ensure that "... *the genetic diversity of ... farmed and domesticated animals and of wild relatives ... is maintained, and strategies have been developed and implemented for minimizing genetic erosion and safeguarding their genetic diversity.*" However, the real extent of genetic erosion is very difficult to measure using current data. Therefore, this challenging target demands better coverage, understanding and utilization of genomic and environmental data, the development of optimized ways to integrate these data with social and other sciences and policy analysis to enable more flexible, evidence-based models to underpin FAnGR conservation. At the conference, we attempted to identify the most important problems for effective livestock genomic resource conservation during the next decade. Twenty priority questions were identified that could be broadly categorised into challenges related to methodology, analytical approaches, data management and conservation. It should be acknowledged here that while the focus of our meeting was predominantly around genetics, genomics and animal science, many of the practical challenges facing conservation of genomic resources are societal in origin and are predicated on the value (e.g. socio-economic and cultural) of these resources to farmers, rural communities and society as a whole. The overall conclusion is that despite the fact that the livestock sector has been relatively well-organised in the application of genetic methodologies to date, there is still a large gap between the current state-of-the-art in the use of tools to characterise genomic resources and its application to many non commercial and local breeds, hampering the consistent utilisation of genetic and genomic data as indicators of genetic erosion and diversity. The livestock genomic sector therefore needs to make a concerted effort in the coming decade to enable to the democratisation of the powerful tools that are now at its disposal, and to ensure that they are applied in the context of breed conservation as well as development.

Introduction

Understanding current technical, infrastructural and policy challenges and assessing the likely benefits of overcoming them in the future is essential for any field of scientific endeavour and especially those with clear societal consequences and potential benefits. In this context, the concept of horizon scanning has been developed and applied annually in the field of biodiversity conservation since 2009 (Sutherland and Woodroof 2009), using a variety of systematic and semi-systematic methods to mine trending issues from web engines and social media and by analyzing focused questionnaires. Similar approaches have also been taken to identify emerging issues in agriculture (Pretty et al 2010) and related fields such as soil science, food systems and pollination (Adewopo et al 2014; Ingram et al 2013; Dicks et al 2013). Such exercises have identified a number of issues of relevance to the conservation of FAnGR, such as genetic control of invasive species (Sutherland et al 2014) and sustainable intensification of high yielding agriculture (Sutherland et al 2015). In 2010, Pretty et al's article pinpointing the '*Top 100 questions of importance to the future of global agriculture*' identified genetic issues in crop improvement (e.g. gains in improvement that could result from breeding for stress tolerance) but identified no such pressing agendas for livestock genomic resources. Since Cardellino and Boyazoglu (2009) no attempt has been

published to identify research priorities for FAnGR conservation, despite genetic erosion (*sensu* Aichi Target 13) continuing apace (e.g. Berthouly-Salazar et al 2012; FAO 2015a) and the step-change that has occurred in molecular breed characterization since the routine implementation of livestock Single Nucleotide Polymorphism (SNP) arrays. To fill this gap, a central activity of the Final Conference of the European Science Foundation's Genomic Resources program, held at Cardiff University June 17th – 19th 2014 was to pick out a series of pressing questions that could form part of a research and policy agenda for FAnGR conservation for the next decade. While not following the standard systematic approaches adopted by conventional Horizon Scanning exercises, all 43 attendees of this focused meeting took part in the exercise, including scientists and policy-makers from South and East Asia, North America, Europe and Africa involved in a range of disciplines from genomics to animal breeding, genetic resource management, economic and social sciences and global agricultural policy development.

Methods and Results

During the course of the conference, attendees were asked to contribute up to five questions of highest priority for research, infrastructure and policy development during the coming decade. Eighty-six suggestions were received. The issue identified with highest frequency (18 times) was the need for 'next generation phenotyping (i.e. high-throughput methods to collect and summarise detailed phenotypic data from domestic animals). A summary of the top 20 questions is found in Table 1, a subset of which are presented below (some are amalgamated). All responses were categorized into four major groups, "Methodological Challenges", "Analytical Challenges", "Data Management" and "Conservation Management and Prioritisation". Four working groups were convened to cover these categories and their findings are presented below.

1. Methodological Challenges

Next generation phenotyping: The need for high-resolution phenotypic data to be collected for in-depth characterisation of FAnGR was identified, especially in light of the rapid advances that have been made in molecular breed characterisation. Developing methods for phenotypic characterisation was also identified by Cardellino and Boyazoglu (2009) following from FAO recommendations (FAO, 2007a) and has clearly remained an under-explored research area. However, with the richness of molecular data increasing dramatically since 2009, the mismatch between molecular and phenotypic data is widening for all except highly commercial transboundary breeds and lines with genomic breeding values. Inherent in high-resolution breed characterisation is a need to define key phenotypic traits and characteristics (particularly those potentially involved in local adaptation) based on guidelines that can be used as common measures for such studies with stringent field protocols for their collection. FAO published guidelines on phenotypic characterization (FAO 2012a). In this way more comparable data can be generated, and breed characterisation can have a more functional basis, especially with the urgent need to understand breed characteristics in the face of climate change (Hoffmann 2010). Also an improved description of the specific production environment and epidemiological history in which populations of a breed are kept would allow better comparison of phenotypes and performances (e.g. FAO, 2009). Since breed characterisation can be a costly exercise, especially for remote regions of the world, as many phenotypic traits as possible should be collected following well documented and reproducible procedures, a process that calls for the need for standardized methods to measure/collect data and ultimately for training of people on how to do it. Where possible, data should be made publicly available through a repository such as FAO's global Domestic Animal Diversity Information System DAD-IS (<http://dad.fao.org>) for comparative purposes. The establishment of a working group to define guidelines, protocols and tools for collecting such data under the auspices of the FAO, International Society for Animal Genetics or the International Committee for Animal Recording (www.icar.org) would accelerate this process.

Omics data and association studies: The dramatic acceleration in genome sequencing means that all domesticated species and their few remaining wild relatives will become genome-enabled in the coming decade (e.g. Qiu et al 2012; Wu et al 2014). Reference genomes provide the basis for development of genome-wide assays for variation in less commonly farmed and/or more regionally distributed livestock species and populations using SNP arrays, as have been developed and made available for commercial livestock in the past five years (e.g. Matukumalli et al 2009). The choice of SNPs for inclusion in arrays for less commercial populations may be expected to focus on a wider array of traits than for commercial/transboundary breeds, such as those related to local adaptation, disease resistance, drought tolerance and niche product characters, but in practice this could be hampered by a lack of reliable phenotypic data. To enable SNP arrays to be developed in a rapid, cost effective and widely applicable manner, the identification of common reference genomes and test panels of individuals for array development and diversity studies is key. However, it is important to note that with the rapidly falling cost of whole genome resequencing (eg Lee et al 2013; Zhang et al 2015) using next generation technologies and the availability of even lower cost genotyping by sequencing (GBS: De Donato et al 2013) being available, the problem of ascertainment bias can be mitigated against since they allow the identification and direct estimation of SNP diversity for FAnGR populations, breeds or species at reasonable prices. Indeed these methods are sufficiently cost-effective now, that they can be in principle used as standard assaying approaches, with a cost in the low tens of dollars for GBS now feasible for analysis of tens of thousands of SNPs.

A major issue identified for genome-wide association studies (GWAS) is experimental design including, but not confined to, sample size considerations (Kadarmideen 2014) and the availability of different SNP genotyping arrays for some species and their compatibility or lack thereof (Nicolazzi et al 2015). Characterisation of environmental parameters in extensive production systems is another key challenge for GWAS but may be assisted by the application of E(environment)WAS methodologies as applied in humans (e.g. Patel et al 2010). Additionally, understanding the role of the epigenome and its role in environment-dependent phenotypic diversity and plasticity is becoming an increasing focus in livestock genetics (e.g. Jammes and Renard 2010; Magee et al 2011, 2014). Ultimately, the integration of genomic, epigenomic, transcriptomic and environmental data will be required if meaningful large-scale studies are to be successful in identifying selection and conservation targets in heterogeneous environments (Jones et al 2013; Wu et al 2014) and in scrutinizing the biological basis for adaptation, resilience, and even animal improvement.

Non-autosomal inheritance: Non-autosomal inheritance (Y-chromosomal, X-chromosomal and mitochondrial) is a comparatively neglected area of research in livestock conservation. While studies of non-autosomal genetic markers have been extensively used in studies of evolutionary history, both singly and combined (e.g. Götherström et al 2005; Svensson and Götherström 2008; Meadows and Kijas 2008; Pereira et al 2009; Ramirez et al 2009; Ginja et al 2010; Groeneveld et al 2010), their exploitation in genomic studies has been somewhat overlooked in comparison to autosomal markers in many livestock species. This oversight is surprising given the well-documented links between mitochondrial sequence variation and fitness in human populations (e.g. Wallace 2005) and the increasingly recognised role that Y-chromosomal variation plays in male fertility in livestock (e.g. Chang et al 2013; Yue et al 2014). Technical challenges have long been acknowledged with finding polymorphic markers on the Y-chromosome in mammals and W-chromosome in birds, however such markers, although elusive, have been shown to provide novel insights into livestock diversity when available (e.g. Edwards et al 2011; Wallner et al 2013), and should be used as a matter of course to provide a male/female perspective on livestock genomic diversity.

Ancient DNA studies: Although firmly established as a major route into a deeper understanding of livestock evolution and diversity (e.g. Larson et al 2010), ancient DNA (aDNA) studies have been hampered by a number of constraints. These include limited

access to samples from geographic areas where (local) domestication may have taken place (e.g. Africa, Near East, Asia, South America), limited data sharing among those groups working on samples from critical sites (but see Arbuckle et al 2014) and limited success rates, especially for genome-wide studies. Nonetheless, recently developed methodological and bioinformatics tools allowed for increased accuracy in the analysis of high-throughput ancient DNA data and even the characterization of complete genomes of Pleistocene horses (Orlando et al 2014). Also, alternative sources of material such as parchment are, however, providing promising outcomes (Teasdale et al 2015). Exciting opportunities have recently been opened up by the discovery of livestock DNA in lake sediment samples in Lake Anterne, Switzerland (Giguet-Covex et al 2014), which enabled a direct comparison to be made of the paleoenvironment with changes in this environment due to the arrival of farming and domestic livestock, and could be applied to describe historic fluctuations in agricultural intensity and practice and, excitingly, may even allow the possibility of predictive modelling for the presence/absence of suitable agri-habitat under future climate change scenarios.

2. Analytical Challenges

Conservation of genomic diversity: The concept of genome conservation has been discussed extensively in the literature but advances in genome data and technologies only now allow the development of breed management programs able to achieve this aim. For example, Herrero-Medrano et al (2014), using genome resequencing and SNP arrays discovered almost 100 non-synonymous polymorphic nucleotides nearly fixed in commercial pig breeds but with an alternative allele in non-commercial populations, affecting 65 genes in total. Such genomic polymorphisms could fall into a category of those that '*cannot afford to be lost*' from less commercial local breeds, given their distinctiveness and the value they potentially represent as a genetic resource for alternative selection should the production environment change (Kristensen et al 2015). However, to design a management program that evaluates genomic regions for conservation, not only do polymorphisms need to be identified, the functional architecture of those genomic regions and the genes they contain needs to be assessed and the interaction among those genes needs to be considered. Recently, a study of chicken breeds examined functional variation in copy number variants (CNV) at over 200 genes overlapping 1,000 quantitative trait loci, including some putatively involved in traits such as skin colour and skeletal characteristics (Han et al 2014).

Haplotype blocks versus individual SNPs: Obtaining an accurate description of the genetic polymorphisms explaining a trait of evolutionary, adaptive and/or economic importance is not a trivial task, as traits substantially vary in the number of polymorphisms involved in their phenotype and where these occur across the genome (Goddard and Hayes 2009; Olson-Manning et al 2012). For example, many of such traits are polygenic and distributed around the genome, making whole-genome resequencing, and medium and high-density SNP arrays a powerful approach to locating them and elucidating their variation (e.g. Huang et al 2010). However, for certain linked traits, haplotypes may provide a more efficient unit of assessing diversity in QTL regions than individual SNPs (e.g. Kijas et al 2013; Mokry et al 2014, Bosse et al 2014a, Bosse et al 2014b), reflecting local genomic architecture in a more accurate fashion. Consequently, at the initial stages of studies aiming to identifying the genetic basis of phenotypic variation, general genome-wide SNP analyses may be more suitable. It is worth noting, however, that phasing haplotypes in divergent populations lacking complementary pedigree data presents a non-trivial challenge. Haplotype analysis can provide an especially powerful tool to investigate the hybrid origin of domesticated populations. For instance, modern Western commercial pig genomes are a mosaic of Eastern and Western Eurasian biogeographic origin. Admixture mapping allows the 'sorting' of haplotype segments for their putative origin. In addition, this strategy has been shown to be powerful to infer selection on specific haplotypes post-hybridization (Bosse et al. 2014a, Bosse et al. 2014b).

Managing the transition from microsatellite to SNP data: The transition from microsatellite markers to SNPs has happened rapidly in FAnGR for commercial/transboundary breeds due to the availability of relatively inexpensive 50K SNP genotyping arrays for most common livestock species (Matukumalli et al 2009). However, SNP arrays are not yet affordable tools for much of the world's FAnGR and are not yet available for all species (see above). This therefore raises the immediate problem of how to integrate data from the two marker types and how to manage the transition from microsatellite-based FAnGR characterisation (much of which has been carried out using markers recommended by ISAG, FAO 2011) to SNP-based characterisation. One option is to re-genotype many of the breeds that already have microsatellite genotypes with SNPs (Ajmone-Marsan et al 2014), but this would be expensive and if implemented would raise the question as to whether the new data would again be replaced by a newer technology (e.g. whole-genome resequencing). Pragmatically, it seems that microsatellite data are perfectly adequate for estimating genetic diversity and describing demographic relationships (e.g. Ferrando et al 2014). However, for cost reasons the full set of microsatellite markers was frequently not applied, especially in developing countries. Also, microsatellite data will not be as efficient for enabling the identification and targeted conservation of genomic regions under selection since data are usually produced with a few tens of quasi-neutral markers (e.g. Herrero-Medrano et al 2013).

Nevertheless, it is becoming clear that data produced using SNP arrays are more repeatable and do not suffer from scoring differences that have made the combination of microsatellite datasets sometimes problematic and requiring statistical evaluation (Lenstra et al 2012). Paradoxically, whole genome resequencing may become the most reliable and cost effective way to analyse genomic diversity in the future, even for non-commercial breeds, if the cost comes down by another order of magnitude (as may happen with portable sequencers such as Oxford Nanopore's MinION system), providing the advantage of no longer needing to use a set of SNP markers ascertained from commercial populations.

Genome-wide diversity statistics: The emergence of whole genome sequencing and medium-high density SNP arrays means that summarising genetic diversity can now be a more nuanced and genomic region-specific exercise. It is well known that ascertainment bias of SNP arrays can strongly underestimate the diversity of the (usually autochthonous and less commercial) breeds not used to design the arrays (Neto and Barendse 2010). This phenomenon does not impact on whole-genome resequencing as all polymorphisms are captured provided sufficient sequence depth is achieved. A combination of parameters will be required to adequately summarise genome diversity (e.g. heterozygosity and effective population size and inbreeding), as no single all encompassing statistic to summarise all of a population's genomic diversity and history exists, despite of how tempting it may be to define such statistic (e.g. for policy makers). Effective population size (N_e) estimates can be obtained with as little as a single genome using methods such as the Pairwise Sequential Markovian Coalescent, although these analyses can prove inconclusive if genome coverage is insufficient or if admixture pertains (Li and Durbin 2010; Schiffels and Durbin 2014; Frantz et al 2015; Orozco-terWengel and Bruford 2014). For recently evolved populations, such as many domestic species, linkage disequilibrium-based (LD) estimates may be more accurate and methods are now emerging to carry out these analysis (e.g. Barbato et al 2015). Runs of homozygosity (ROH; e.g. Bosse et al 2012; Scraggs et al 2014) functions describing the distribution of homozygosity throughout the genome may also serve as a robust genome-scale N_e estimator in the future, although interpretation and scaling depends on the local recombination. ROH are already used as a genomic proxy for inbreeding (e.g. Purfield et al 2012; Curik et al 2014), including for specific genome-located traits (Pryce et al 2014). This approach promises to be an efficient way to avoid the production of offspring homozygous for deleterious alleles at specific genomic regions that are associated with inbreeding depression (Pryce et al 2014).

3. Data Management

Data accessibility: As also identified by Cardellino and Boyazoglu (2009) there remains a major need to provide much better *links* between the major FAnGR databases, which have largely been set up independently and are breed-focused (Groeneveld et al 2010). The livestock genomics community needs either to build on an existing platform (such as the ARKDB– <http://www.thearkdb.org/arkdb/> and the European Nucleotide Archive <http://www.ebi.ac.uk/ena/>), that have some level of connectivity, eg with Ensembl (<http://www.ensembl.org/index.html>) or to establish an independent community-based initiative(s) under the form of a user-friendly global web portal and would include web services able to federate resources and act as an *educational* central point. Such resources are already being developed, including the Adaptmap project for goats (<http://www.goatadaptmap.org/>). Information on livestock related data should be made available and useful recommendations are required to inform stakeholders on how to record data, and where to store what type of information. In particular, it is important to promote within the community of users that raw and meta-data are key components and that they should be made available in public datasets together with elaborated datasets. When there are existing public resources for a given datatype such as those listed above, they should be used for their ability set standards and centralise data access. For other data types, open digital repositories such as Dryad (<http://datadryad.org/>), Zenodo (<https://zenodo.org/>), or figshare (<http://figshare.com/>) comprise invaluable tools acting as incentives for people to maintain and upgrade their datasets as data can be submitted and authors are provided with a reference which can be cited. This data ecosystem becomes especially important with the myriad of SNP array datasets that are now available and the incompatibility among different versions of these arrays within the same species (Nicolazzi et al 2015). Moreover, to add value to genetic resources, federating gene bank resources is one step that needs to be completed by explicit connection – through geographical coordinates – with phenotypic data, but also with socio-economic, socio-demographic, climatic, environmental, and policy information. This requires links to existing online digital resources (Joost et al. 2010) that are currently rarely used by the FAnGR community often and need to be listed on such a global portal.

Data availability: While many genotyping projects on commercial livestock breeds are funded by industry, rendering all except summary data unavailable in many cases, in principle raw data from publicly funded projects should be made publicly available. Indeed, when data are open, it first makes the information more credible, makes data re-usable, and also enables reproducibility an important scientific principle (Ertz et al. 2014). Increasingly, international consortia, such as FAANG on animal functional genomics follow the Toronto protocol and immediately place data in the public domain (www.faaang.org; The Faang Consortium 2015; The Toronto International Data Release Workshop Authors 2009). A next generation phenotyping database should also be established, including GIS and anonymized farm level data, animal photographs and meta-data – this could partly follow the format of the EU FP5 project Econogene (<http://www.econogene.eu>) and would be most efficiently linked with FAO's DAD-IS and EFABIS (<http://efabis.tzv.fal.de>). The ownership and hosting of such a resource would be logistically and financially challenging, and could provide an opportunity for the agri-industry to contribute towards conservation of the genetic resources it has utilized in the past and may need again in the future. This could also be part of the community-based action mentioned above, with many advantages (logistic and funding), but requiring a strong leadership. An approach to data resourcing such has been exemplified with human data by the 1,000 Genomes project (<http://www.1000genomes.org>) and the 1,001 Arabidopsis genomes resource (<http://1001genomes.org>) with data being publicly available either immediately or after an agreed embargo period could be very applicable to livestock studies. For example, the resequencing data from the EU Framework 7 Nextgen project was made available shortly after the project's completion at the European Bioinformatics Institute's FTP site (<ftp://ftp.ebi.ac.uk/pub/databases/nextgen/>).

Participatory projects: Many individuals who are interested in FAnGR are involved in

agriculture as smallholders, farmers, breeders and producers and many of these are not formally involved in breeding programs and livestock conservation, yet maintain an interest through agricultural shows and farmer's markets (eg Zimmerer 2010; Johns et al 2013). At the same time, the role of participatory approaches and mobile technology potentially enables robust data collection on a previously unimaginable scale (Teacher et al 2013; Lisson et al 2010; Sambo et al 2015). Use of crowdsourcing should therefore be encouraged in FAnGR as should use of smart-phone apps and technologies for photography, data storage and sampling (eg 'do-forms' <http://www.doforms.com>). A logical combination of these initiatives lies in the possibility of a livestock community independent initiative, including website, web services to federate these data sources, to carry out quality control and providing a central access point for data but also information to educate people on how to record FAnGR data. Such approaches could also help in securing funds for projects in FAnGR populations and breeds, which often face the problem of securing funds to carry out this necessary research.

4. Conservation, management and prioritization

Is prioritization a priority?: A paradigm within FAnGR for the past 15 years concerns the use of genetic data, alongside other information in prioritization of livestock populations and breeds for conservation (Weitzman 1992; Simianer et al 2002; Boettcher et al 2010; Ginja et al 2013). However, there is limited evidence that this approach is being applied systematically across countries reporting to the FAO, although the second report on the State of the World's Animal Genetic Resources has documented activities to some extent (FAO, 2015b,c). If, however, prioritization methods are not being applied by managers and policy-makers, the question needs to be asked as to why? A number of explanations may pertain: first, the method(s) may have not gained enough traction with policy makers to ensure its/their implementation, which may indeed be because genomic methods, which have yet to be systematically implemented, will largely supersede the microsatellite-based approaches implemented thus far and enable conservation prioritization to include genes important in functionally valuable traits (eg Toro et al 2014). Furthermore, prioritization on the basis of genetic distances (Weitzman 1992) is confounded by genetic isolation of breeds (European Cattle Genetic Diversity Consortium, 2006). Second, prioritization may not actually be needed, at least in certain regions, where breed societies are active and all or most of the breeds can be maintained. However, recent animal health emergencies (eg outbreaks of transmissible spongiform encephalopathies, TSEs) have cast doubt on this simplistic scenario and required the application of careful genetic management during and after the outbreak. While prioritization may be less of a priority the in the world's richest regions, it is not expected to be the case in developing countries, where extinction may take a number of forms, including genetic erosion (eg Berthouly-Salazar et al 2012, FAO, 2015a,b). Finally, the methods developed may not have been applied because policy makers and managers are unaware of their availability, which could be due to a lack of dissemination or penetrance of educational material to the decision makers.

Utilization in practice: While research and application of genomic tools in livestock is occurring in many commercial/transboundary breeds (eg Pryce et al 2014; Scraggs et al 2014), its application in less commercial populations is sporadic and the scientific basis of decisions on management of indigenous livestock, for example in which germplasm to store, assessing the effects of upgrading or evaluating ongoing genetic management is therefore highly variable (eg Brown et al 2014; FAO 2015b). This points to the reality that genetically-based prioritization is unlikely to be operational in the absence of other considerations, including commercial reality and the ecosystem/production environment (eg Sanderson et al 2014). The use of genomic data to manage FAnGR *within* breeds is however, continuing apace (see above) and can be demonstrated to be assisting conservation, production and management in many cases (e.g. Scraggs et al 2014; Herrero-Medrano et al 2014). However, for many breeds the cost of genetic/genomic analysis versus the potential

economic returns on genotyped stock (with a few exceptions such as TSE resistance) makes its application uneconomic, and therefore it is often not applied. It is unlikely that genotyping costs will reach the level of economic viability for many FAnGR, however this assumption should be tested by some targeted research across the sector.

Defining goals: The Convention on Biological Diversity's Aichi Target 13, which recommends that: "*strategies have been developed and implemented for minimizing genetic erosion and safeguarding genetic diversity*" is reflected in the Target for Strategic Priority Area 4 of the Global Plan of Action for Animal Genetic Resources (FAO, 2007b). These resource indicators contribute to the measurement of progress towards Aichi Target 13 (FAO, 2012b) and are calculated at national, regional and global levels, based on data entered by National Coordinators for the Management of Animal Genetic Resources¹ (172 countries had nominated a National Coordinator as of July 2014) into the Domestic Animal Diversity Information System (DAD-IS). The following indicators have been agreed by the Commission on Genetic Resources for Food and Agriculture:

- the number of locally adapted breeds;
- the proportion of the total population accounted for by locally adapted and exotic breeds; and
- the number of breeds classified as at risk, not at risk and unknown.

The Global Databank for Animal Genetic Resources, the backbone of DAD-IS, enables National Coordinators to enter breed-specific data, including data on the size and structure of breed populations, required to calculate their risk status. FAO produces biannual Status and Trends Reports (FAO, 2015a). For the first report on The State of the World's Animal Genetic Resources, a risk status classification based on population size data was used. The (lack of) availability of global data currently makes a more elaborate system involving, for example, molecular diversity indices, population structure/fragmentation, pedigree data, number and size of herds, and geographic distribution inoperable. While genomic methods might help to overcome these data deficiencies, if they are to be applied to livestock conservation, it is important to define the goals of such approaches and how the data could be used to improve or augment the current set of indicators using data that could be collected on trends in effective population size, admixture, inbreeding and genome-wide diversity. The wider application of such data hinges on their applicability to autochthonous, less-commercial breeds. Unfortunately, the data currently provided to FAO does not even allow the reliable calculation of basic trends currently measured via the above indicators (Tittensor et al 2014, FAO 2015a) yet the livestock genetics and conservation community possess many of the tools needed to directly evaluate whether signatories to the CBD are '*...minimizing genetic erosion*' and '*safeguarding genetic diversity*' (CBD Target 13). Two key developments are required to enable the current approach to more directly use genetic or genomic data in the future: first, the livestock conservation genetics community must therefore insist that data are collected and analysed in such a way that results are directly comparable and second to help develop better indicators applied to monitoring genetic trends in domestic populations.

Conclusion

Any exercise designed to assess the state-of-the-art in a scientific field only manages to capture a brief moment in time, which is why the Horizon scanning exercises carried out in biodiversity conservation are repeated every year (see Sutherland et al 2015). Here, we attempted to take a longer-term (decadal) view of genomic resources conservation, and during this period, some major milestones will be passed. Chief among these is the imminent release of the Second Report on the State of the World's Animal Genetic Resources (final publication due Autumn 2015, FAO 2015b,d) and the Convention on Biological Diversity's 2020 deadline halting the loss of biodiversity Aichi targets. In the context of the dramatic

advances in omics technology that are expected during the next decade, the field is expected to move fast. But structural changes in the livestock sector that will bring further erosion during this period are likely to be equally rapid. However, this makes it critically important that a strategic approach is taken to incorporating these technological advances into real world FAnGR conservation. Such an approach has been taken in the past (e.g. with the implementation of approved microsatellite marker sets) and, we would argue, is needed now to ensure that practical conservation of farm animal agricultural biodiversity is not left behind. The FAnGR community therefore needs to make best use of new genomic tools, and at the same time continue and augment its classical phenotyping efforts. Both, genomic and phenotypic tools need to be applied more consistently, at a much wider scale and for more breeds, to describe, utilise and conserve the world's genomic/breed diversity for future generations.

¹ The list of National Coordinators for the Management of Animal Genetic Resources is found at dad.fao.org/cgi-bin/EfabisWeb.cgi?sid=-1,contacts.

In review

References

- Adewopo JB, VanZomeren C, Bhomia RK, Almaraz M, Bacon A, Eggleston E, Judy JD, Lewis RW, Lusk M, Miller B, Moorberg C, Snyder EH, Tiedeman M (2014) Top-Ranked Priority Research Questions for Soil Science in the 21st Century. *Soil Sci Soc Am. J.* 78: 337-347.
- Ajmone-Marsan P, Colli L, Han J, Achilli A, Lancioni H, Joost S, Crepaldi P, Pilla F, Stella A, Taberlet P, Boettcher P, Negrini R, Lenstra JA (2014) The characterization of goat genetic diversity: Towards a genomic approach. *Small Ruminant Res* 121: 58-72.
- Arbuckle BS, Kansa SW, Kansa E, Orton D, Cakirlar C, Gourichon L, Atici L, Galik A, Marciniak A, Mulville J, Buitenhuis H, Carruthers D, De Cupere B, Demirergi A, Frame S, Helmer D, Martin L, Peters J, Poellath N, Pawlowska K, Russell N, Twiss K, Wuertenberger D (2013) Data Sharing Reveals Complexity in the Westward Spread of Domestic Animals across Neolithic Turkey. *PLoS ONE* 9: e99815.
- Barbato M, Orozco-terWengel PA, Tapio M, Bruford MW (2015) SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Frontiers Genet.* 6: 109.
- Berthouly-Salazar C, Thevenon S, Van TN, Nguyen BT, Pham LD, Chi CV, Maillard J-C (2012) Uncontrolled admixture and loss of genetic diversity in a local Vietnamese pig breed. *Ecol Evol* 2: 962-975.
- Bosse M, Megens HJ, Madsen O, Paudel Y, Frantz LA, Schook LB, Crooijmans RP, Groenen MA (2012) Regions of homozygosity in the porcine genome: consequence of demography and the recombination landscape. *PLoS Genet.* 8: e1003100.
- Bosse M, Megens HJ, Madsen O, Frantz LA, Paudel Y, Crooijmans RP, Groenen MA (2014a) Untangling the hybrid nature of modern pig genomes: a mosaic derived from biogeographically distinct and highly divergent *Sus scrofa* populations. *Mol Ecol* 23:4089-102.
- Bosse M, Megens HJ, Frantz LA, Madsen O, Larson G, Paudel Y, Duijvesteijn N, Harlizius B, Hagemeyer Y, Crooijmans RP, Groenen MA (2014b) Genomic analysis reveals selection for Asian genes in European pigs following human-mediated introgression. *Nat Comms* 5:4392.
- Cardellino RA, Boyazoglu J (2009) Research opportunities in the field of animal genetic resources. *Livestock Sci* 120: 166-173.
- Chang TC, Yang Y, Retzel EF, Liu WS (2013) Male-specific region of the bovine Y chromosome is gene rich with a high transcriptomic activity in testis development. *Proc Natl Acad Sci USA* 110: 12373-12378.
- Curik I, Ferencakovic M, Soelkner, J (2014) Inbreeding and runs of homozygosity: A possible solution to an old problem. *Livestock Sci* 166: 26-34.
- De Donato, M, Peters SO, Mitchell SE, Hussain T, Imumorin IG (2013) Genotyping-by-Sequencing (GBS): A Novel, Efficient and Cost-Effective Genotyping Method for Cattle Using Next-Generation Sequencing. *PLoS ONE* 8: e62317.
- Dicks LV, Abrahams A, Atkinson J, Biesmeijer J, Bourn N, Brown C, Brown MJF, Carvell C, Connolly C, Cresswell JE, Croft P, Darvill B, De Zylva P, Effingham P, Fountain M, Goggin A, Harding D, Harding T, Hartfield C, Heard MS, Heathcote R, Heaver D, Holland J, Howe M,

Hughes B, Huxley T, Kunin WE, Little J, Mason C, Memmott J, Osborne J, Pankhurst T, Paxton RJ, Pocock MJO, Potts SG, Power EF, Raine NE, Ranelagh E, Roberts S, Saunders R, Smith K, Smith RM, Sutton P, Tilley LAN, Tinsley A, Tonhasca A, Vanbergen AJ, Webster S, Wilson A, Sutherland WJ (2013) Identifying key knowledge needs for evidence-based Conservation Of Wild Insect Pollinators: A Collaborative Cross-Sectoral Exercise. *Insect Cons Diversity* 6: 435-446.

Edwards CJ, Ginja C, Kantanen J, Perez-Pardal L, Tresset A, Stock F, Gama LT, Penedo MCT, Bradley DG, Lenstra JA, Nijman IJ and European Cattle Genetic Diversity Consortium (2011) Dual Origins of Dairy Cattle Farming - Evidence from a Comprehensive Survey of European Y-Chromosomal Variation. *PLoS ONE* 6: e15922.

Ertz, O., Rey, S.J. & Joost, S. (2014). The open source dynamics in geospatial research and education. *J. Spatial Inf. Sci.* 0: 67–71.

European Genetic Cattle Diversity Consortium (2006) Marker-assisted conservation of European cattle breeds: an evaluation. *Anim. Genet.* 37: 475-481.

The FAANG Consortium, Andersson L, Archibald AI, Bottema CD, Brauning R, Burgess SS, Burt DW, Casas E, Cheng HH, Clarke L, Couldrey C, Dalrymple BP, Elsik CG, Foissac S, Guiffra E, Groenen MA, Hayes BJ, Huang LS, Khatib H, Kijas JW, Kim H, Lunney JK, McCarthy FM, McEwan JC, Moore S, Naduri B, Notredame C, Palti Y, Plastow GS, Reecy GM, Rohra GA, Sarropoulou E, Schmidt CJ, Silverstein J, Tellam RL, Tixier-Boichard M, Tosser-Klopp G, Tuggle CK, Vilkki J, White SN, Zhao S, Zhou H (2015) Coordinated international action to accelerate genome-to-phenome with FAANG, the Functional Annotation of Animal Genomes project. *Genome Biol* 16: 57.

FAO (2007a) "The Global Plan of Action for Animal Genetic Resources and the Interlaken Declaration." Rome. www.fao.org/docrep/010/a1404e/a1404e00.htm.

FAO (2007b) "The State of the World's Animal Genetic Resources for Food and Agriculture". Rome: Food and Agriculture Organization of the United Nations, 2007. 511. edited by B. Rischkowsky & D. Pilling. www.fao.org/3/a-a1250e.pdf.

FAO (2009) Report of the FAO/WAAP Workshop on Production Environment Descriptors for Animal Genetic Resources. dad.fao.org/cgi-bin/getblob.cgi?sid=-1,593

FAO (2011) *Molecular genetic characterization of animal genetic resources*. FAO Animal Production and Health Guidelines. No. 9. Rome.

FAO (2012a) Phenotypic characterization of animal genetic resources. FAO Animal Production and Health Guidelines No. 11. Rome.

FAO (2012b) "Targets and indicators for animal genetic resource." Seventh Session of the Intergovernmental Technical Working Group on Animal Genetic Resources for Food And Agriculture, Rome, 24–26 October 2012 (CGRFA/WG-AnGR-7/12/7). Rome. www.fao.org/docrep/meeting/026/me514e.pdf.

FAO (2015a) "Status and trends of animal genetic resources – 2014." Fifteenth Regular Session of the Commission on Genetic Resources for Food and Agriculture, Rome, 19–23 January 2015 (CGRFA-15/15/Inf.18). Rome. www.fao.org/3/a-mm278e.pdf.

FAO (2015b) Draft Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture (Part 3), Commission on Genetic Resources for Food and Agriculture, CGRFA-15/15/Inf.17.2. www.fao.org/3/a-mm310e.pdf

FAO (2015c) "Synthesis progress report on the implementation of the Global Plan of Action for Animal Genetic Resources – 2014." Fifteenth Regular Session of the Commission on Genetic Resources for Food and Agriculture, Rome, 19–23 January 2015 (CGRFA-15/15/Inf.19). Rome. www.fao.org/3/a-mm282e.pdf.

FAO. 2015d. Draft Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture (Part 1 and 2), Commission on Genetic Resources for Food and Agriculture, CGRFA-15/15/Inf.17.1. www.fao.org/3/a-mm313e.pdf

Ferrando A, Goyache F, Pares P-M, Carrion C, Miro J, Jordana J (2014) Genetic relationships between six eastern Pyrenean sheep breeds assessed using microsatellites. *Spanish J. Ag. Res.* 12: 1029-1037.

Frantz LAF, Madsen O, Megens H-J, Schraiber JG, Paudel Y, Bosse M, Crooijmans RPMA, Larson G, Groenen MAM (2015) Evolution of Tibetan wild boars. *Nature Genet* 47: 188-189.

Giguet-Covex C, Pansu J, Arnaud F, Rey P-J, Griggo C, Gielly L, Domaizon I, Coissac E, David F, Choler P, Poulénard J, Taberlet P (2014) Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nature Comms.* 5: 3211.

Ginja C, Penedo MCT, Melucci L, Quiroz J, Martínez López OR, Revidatti MA, Martínez-Martínez A, Delgado JV, Gama LT (2010) Origins and genetic diversity of New World Creole cattle: inferences from mitochondrial and Y chromosome polymorphisms. *Anim. Genet.* 41: 128–141.

Goddard MT, Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programs. *Nature Revs Genet* 10: 381- 391.

Götherström A, Anderung C, Hellborg L, Elburg R, Smith C, Bradley DG, Ellegren H (2005) Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. *Proc. R. Soc. Lond B* 272: 2345-2350.

Groeneveld LF, Lenstra JA, Eding H, Toro MA, Scherf B, Pilling D, Negrini R, Jianlin H, Finlay EK, Groeneveld E, Weigend S. & GlobalDiv Consortium (2010) Genetic diversity in livestock breeds. *Anim. Genet.* 41 (suppl. 1), 6-31.

Han R, Yang P, Tian Y, Wang D, Zhang Z, Wang L, Li Z, Jiang R, Kang, X (2014) Identification and functional characterization of copy number variation in diverse chicken breeds. *BMC Genomics* 15: 934.

Herrero-Medrano JM, Megens HJ, Crooijmans RP, Abellaneda JM, Ramis G (2013) Farm-by-farm analysis of microsatellite, mtDNA and SNP genotype data reveals inbreeding and crossbreeding as threats to the survival of a native Spanish pig breed. *Anim Genet* 44:259-266.

Herrero-Medrano J-M, Megens H-J, Groenen MAM, Boss M, Pérez-Enciso M, Crooijmans RPMA (2014) Whole-genome sequence analysis reveals differences in population management and selection of European low-input pig breeds. *BMC Genomics* 15: 601.

Hoffmann I (2010) Climate change and the characterization, breeding and conservation of animal genetic resources. *Anim Genet* 41 (Suppl 1): 32-46.

Huang W, Kirkpatrick BW, Rosa GJM, Khatib H (2014) A genome-wide association study using selective DNA pooling identifies candidate markers for fertility in Holstein cattle. *Anim. Genet.* 41: 570-578.

Ingram JS, Wright HL, Foster L, Aldred T, Barling D, Benton TG, Berryman P, Bestwick, CS, Bows-Larkin A, Brocklehurst TF, Buttriss J, Casey J, Collins H, Crossley DS, Dolan CS, Dowler E, Edwards R, Finney KJ, Fitzpatrick JL, Fowler M, Garrett DA, Godfrey JE, Godley, A, Griffiths W, Houlston EJ, Kaiser, MJ, Kennard R, Knox JW, Kuyk A, Linter BR, Macdiarmid JI, Martindale W, Mathers JC, McGonigle DF, Mead A, Millar SJ, Miller A, Murray C, Norton, IT, Parry S, Pollicino M, Quested TE, Tassou S, Terry LA, Tiffin R, van de Graaf P, Vorley W, Westby A, Sutherland WJ (2013) Priority research questions for the UK food system. *Food Security* 5: 617-636.

Kadarmideen HN (2014) Genomics to systems biology in animal and veterinary sciences: Progress, lessons and opportunities. *Livestock Sci* 166 : 232-246.

Kijas JW, Ortiz JS, McCulloch R, James A, Brice B, Swain B, Tosser-Klopp G (2013) Genetic diversity and investigation of polledness in divergent goat populations using 52088 SNPs. *Anim Genet* 44: 325-335.

Kristensen TN, Hoffmann AA, Pertoldi C, Stronen AV (2015) What can livestock breeders learn from conservation genetics and vice versa? *Frontiers Genet.* 6: 38.

Jammes H, Renard J. -P (2010) Epigenetics and phenotype construction, a challenge for livestock production? *Product Anim.* 23: 23-42.

Jones MR, Forester BR, Teufel AI, Adams RV, Anstett DN, Goodrich BA, Landguth BL, Joost S, Manel S (2013) Integrating landscape genomics and spatially explicit approaches to detect loci under selection in clinal populations. *Evolution* 67: 3455-3468.

Joost, S., Colli, L., Baret, P.V., Garcia, J.F., Boettcher, P.J., Tixier-Boichard, M., Ajmone-Marsan, P. & The GLOBALDIV Consortium. (2010). Integrating geo-referenced multiscale and multidisciplinary data for the management of biodiversity in livestock genetic resources. *Anim Genet.* 41: 47-63.

Larson G, Lui R, Zhao X, Yuan J, Fuller D, Barton L, Dobney K, Fan Q, Gu Z, Liu XH, Luo Y, Lv P, Andersson L, Li N (2010) Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. *Proc Natl Acad Sci USA* 107: 7686-7691.

Lee KT, Chung WH, Lee SY, Choi JW, Kim J, Lim D, Lee S, Jang G, Kim B, Choy YH, Liao X, Stothard P, Moore SS, Lee SH, Ahn S, Kim N, Kim TH (2013) Whole-genome resequencing of Hanwoo (Korean cattle) and insight into regions of homozygosity. *BMC Genomics* 14: 519.

Lenstra JA, Groeneveld LF, Eding H, Kantanen J, Williams JL, Taberlet P, Nicolazzi EL, Soelkner J, Simianer H, Ciani E, Garcia JF, Bruford MW, Ajmone-Marsan P, Weigend S (2012) Molecular tools and analytical approaches for the characterization of farm animal genetic diversity. *Anim. Genet.* 43: 483-502.

Lisson, S., MacLeod, N., McDonald, C., Corfield, J., Pengelly, B., Wirajaswadi, L., Rahman, R., Bahar, S., Padjung, R., Razak, N., Puspadi, K., Dahlanuddin, Sutaryono, Y., Saenong, S., Panjaitan, T., Hadiawati, L., Ash, A. & Brennan, L. (2010). A participatory, farming systems approach to improving Bali cattle production in the smallholder crop-livestock systems of Eastern Indonesia. *Agric. Systems* 103: 486-497.

Li H, Durbin Richard (2011) Inference of human population history from individual whole-genome sequences. *Nature* 475: 493-U84.

Magee DA, Sikora KM, Berkowicz EW, Berry DP, Howard DJ, Mullen MP, Evans RD, Spillane C, MacHugh DE (2011) DNA sequence polymorphisms in a panel of eight candidate bovine imprinted genes and their association with performance traits in Irish Holstein-Friesian cattle. *BMC Genet* 11: 93.

Magee DA, Spillane C, Berkowicz EW, Sikora KM, MacHugh, DE (2014) Imprinted loci in domestic livestock species as epigenomic targets for artificial selection of complex traits. *Anim Genet* 45 (S1): 25-39.

Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, O'Connell J, Moore SS, Smith TPL, Sonstegard TS, Van Tassell CP (2009) Development and Characterization of a High Density SNP Genotyping Assay for Cattle. *PLoS ONE* 4: e5350.

Meadows JRS, Kijas JW (2008) Re-sequencing regions of the ovine Y chromosome in domestic and wild sheep reveals novel paternal haplotypes. *Anim. Genet.* 40: 119–123.

Mokry FB, Buzanskas ME, Mudadu MA, Grossi D A, Higa RH, Ventura RV, Lima AO, Sargolzaei M, Conceicao Meirelles, SL, Schenkel, FS, Gualberto Barbosa da Silva MV, Meo Niciura, SC, de Alencar MM, Munari DP, de Almeida Regitano LC (2014) Linkage disequilibrium and haplotype block structure in a composite beef cattle breed. *BMC Genomics* 15: S6.

Pereira F, Queirós S, Gusmão L, Nijman IJ, Cuppen E, Lenstra JA, Econogene Consortium, Davis SJM, Nejmeddine F, Amorim A (2009) Tracing the history of goat pastoralism: new clues from mitochondrial and Y chromosome DNA in North Africa. *Mol. Biol. Evol.* 26: 2765-2773.

Porto Neto LR, Barendse W (2010) Effect of SNP origin on analyses of genetic diversity in cattle. *Anim. Prod. Sci.* 50: 792-800.

Nicolazzi EL, Caprera A, Nazzicari N, Cozzi P, Strozzi F, Lawley C, Pirani A, Soans C, Brew F, Jorjani H, Evans G, Simpson B, Tosser-Klopp G, Brauning R, Williams JL, Stella A. (2015). SNPchiMp v.3: integrating and standardizing single nucleotide polymorphism data for livestock species. *BMC genomics*, 16:283.

Olson-Manning CF, Wagner MR, Mitchell-Odds T (2012) Adaptive evolution: evaluating empirical support for theoretical predictions. *Nature Revs Genet* 13: 867 – 877.

Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, Stiller M, Schubert M, Cappellini E, Petersen B, Moltke I, Johnson PLF, Fumagalli M, Vilstrup JT, Raghavan M, Korneliussen T, Malaspinas AS, Vogt J, Szklarczyk D, Kelstrup CD, Vinther J, Dolocan A, Stenderup J, Velazquez AMV, Cahill J, Rasmussen M, Wang X, Min J, Zazula GD, Seguin-Orlando A, Mortensen C, Magnussen K, Thompson JF, Weinstock J, Gregersen K, Røed KH, Eisenmann V, Rubin CJ, Miller DC, Antczak DF, Bertelsen MF, Brunak S, Al-Rasheid KAS, Ryder O, Andersson L, Mundy J, Krogh A, Gilbert MTP, Kjær K, Sicheritz-Ponten T, Jensen LJ, Olsen JV, Hofreiter M, Nielsen R, Shapiro B, Wan J, Willerslev E (2013) Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* 499: 74-78.

Orozco-terWengel PA, Bruford MW (2014) Mixed signals from hybrid genomes. *Mol Ecol* 23: 3941-3943.

Patel CJ, Bhattacharya J, Butte AJ (2010) An Environment-Wide Association Study (EWAS) on Type 2 Diabetes Mellitus. *PLoS ONE* 5: e10746.

Pretty J, Sutherland WJ, Ashby J, Auburn J, Baulcombe D, Bell M, Bentley J, Bickersteth S, Brown K, Burke J, Campbell H, Chen K, Crowley E, Crute I, Dobbelaere D, Edwards-Jones G, Funes-Monzote F, Godfray HJC, Griffon M, Gypmantisiri P, Haddad L, Halavatau S, Herren H, Holderness M, Izac A-M, Jones M, Koohafkan P, Lal R, Lang T, McNeely J, Mueller A, Nisbett N, Noble A, Pingali P, Pinto Y, Rabbinge R, Ravindranath NH, Rola A, Roling N, Sage C, Settle W, Sha JM, Shiming L, Simons T, Smith P, Strzepeck K, Swaine H, Terry E, Tomich TP, Toulmin C, Trigo E, Twomlow S, Kees Vis J, Wilson J, Pilgrim S (2010) Top 100 questions of importance to the future of global agriculture. *Int. J. Agric. Sustainability* 8: 219-238.

Pryce JE, Haile-Mariam M, Goddard ME, Hayes B (2014) Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genet Select. Evol* 46: 71.

Purfield DC, Berry DP, McParland S, Bradley DG (2012) Runs of homozygosity and population history in cattle. *BMC Genet* 13: 70.

Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z, Cao C, Hu Q, Kim J, Larkin DM, Auvil L, Capitanu B, Ma J, Lewin HA, Qian X, Lang Y, Zhou R, Wang L, Wang K, Xia J, Liao S, Pan S, Lu X, Hou H, Wang Y, Zang X, Yin Ye, Ma H, Zhang J, Wang Z, Zhang Y, Zhang D, Yonezawa T, Hasegawa M, Zhong Y, Liu W, Zhang Y, Huang Z, Zhang S, Long R, Yang H, Wang J, Lenstra JA, Cooper DN, Wu Y, Wang J, Shi P, W, Liu J (2012) The yak genome and adaptation to life at high altitude. *Nat Genet* 44: 946.

Ramirez O, Ojeda A, Tomas A, Gallardo D, Huang LS, Folch JM, Clop A, Sanchez A, Badaoui B, Hanotte O, Galman-Omitogun O, Makuza SM, Soto H, Cadillo J, Kelly L, Cho IC, Yeghoyan S, Perez-Enciso M, Amills, M (2009) Integrating Y-Chromosome, Mitochondrial, and Autosomal Data to Analyze the Origin of Pig Breeds. *Mol. Biol. Evol.* 26: 2061-2072.

Sambo, E., Bettridge, J., Dessie, T., Amare, A., Habte, T., Wigley, P. & Christley, R.M. (2015). Participatory evaluation of chicken health and production constraints in Ethiopia. *Prevent. Vet. Med.* 118: 117–127.

Schiffels S, Durbin R (2014) Inferring human population size and separation history from multiple genome sequences. *Nature Genet* 46: 919-925.

Scraggs E, Zanella R, Wojtowicz A, Taylor JF, Gaskins CT, Reeves JJ, de Avila JM, Neibergs HL (2014) Estimation of inbreeding and effective population size of full-blood wagyu cattle registered with the American Wagyu Cattle Association. *J. Anim Breed Genet* 131: 3-10.

Sutherland WJ, Woodroof HJ (2009) The need for environmental horizon scanning. *Trends Ecol. Evol.* 24 : 523–527.

Sutherland WJ, Aveling R, BrooksTM, Clout M, Dicks LV, Fellman L, Fleishman E, Gibbons DW, Keim B, Lickorish F, Monk KA, Mortimer D, Peck LS, Pretty J, Rockstrom J, Rodriguez JP, Smith RK, Spalding MD, Tonneijck FH, Watkinson AR (2014) A horizon scan of global conservation issues for 2014. *Trends Ecol Evol* 29: 1.

Sutherland WJ, Clout M, Depledge M, Dicks LV, Dinsdale J, Entwistle AC, Fleishman E, Gibbons DW, Keim B, Lickorish FA, Monk KA, Ockendon N, Peck LS, Pretty J, Rockstrom J, Spalding MD, Tonneijck FH, Wintle BC (2015) A horizon scan of global conservation issues for 2015. *Trends Ecol Evol* 30: 1.

Svensson E, Gotherstrom A (2008) Temporal fluctuations of Y-chromosomal variation in *Bos taurus*. *Biol. Lett.* 4: 752–754.

Teasdale MD, van Doorn NL, Fiddymment S, Webb CC, O'Connor T, Hofreiter M, Collins MJ, Bradley DG (2015) Paging through history: parchment as a reservoir of ancient DNA for next generation sequencing. *Phil Trans R Soc Lond B* 370: 20130379.

Tittensor DP, Walpole M, Hill SLL, Boyce DG, Britten GL, Burgess ND, Butchart SHM, Leadley PW, Regan EC, Alkemade R, Baumung R, Bellard C, Bouwman Le, Bowles-Newark NJ, Chenery AM, Cheung WWL, Christensen V, Cooper HD, Crowther AR, Dixon MJR, Galli A, Gaveau V, Gregory R, Gutierrez NL, Hirsch TL, Hoefft R, Januchowski-Hartley SR, Karmann M, Krug CB, Leverington FJ, Loh J, Lojenga RK, Malsch K, Marques A, Morgan DHW, Mumby PJ, Newbold T, Noonan-Mooney K, Pagad SN, Parks BC, Pereira HM, Robertson T, Rondinini C, Santini L, Scharlemann JPW, Schindler S, Sumaila UR, Teh LSL, van Kolck J, Visconti P, Ye Y (2014) A mid-term analysis of progress toward international biodiversity targets. *Science* 346: 241-244.

The Toronto International Data Release Workshop Authors (2009) Prepublication data sharing. *Nature* 461: 168-170.

Wallace DC (2005) A mitochondrial paradigm of metabolic and degenerative disease, aging and cancer : a dawn for evolutionary medicine. *Ann Rev Genet* 39 : 359-407.

Wallner B, Vogl C, Shukla P, Burgstaller JP, Druml T, Brem, G (2013) Identification of Genetic Variation on the Horse Y Chromosome and the Tracing of Male Founder Lineages in Modern Breeds. *PLoS ONE* 8: e60015.

Weitzman ML (1992) On diversity. *Quart J Econ.* CVII, 363– 405.

Wu H, Guang X, Al-Fageeh MB, Cao J, Pan S, Zhou H, Zhang L, Abutarboush MH, Xing Y, Xie Z, Alsharqeti AS, Zhang Y, Yao Q, Al-Shomrani BM, Zhang D, Li J, Manee MM, Yang Z, Yang L, Liu Y, Zhang J, Altammami MA, Wang S, Yu L, Zhang W, Liu S, Ba L, Liu C, Yang X, Meng F, Wang S, Li L, Li E, Li X, Wu K, Zhang S, Wang J, Yin Y, Yang H, Al-Swailem AM, Wang J (2014) Camelid genomes reveal evolution and adaptation to desert environments. *Nat Comms* 5: 5188.

Yue XP, Dechow C, Chang TC, DeJarnette JM, Marshall CE, Lei CZ, Liu Wan-S (2014) Copy number variations of the extensively amplified Y-linked genes, HSFY and ZNF280BY, in cattle and their association with male reproductive traits in Holstein bulls. *BMC Genomics* 15: 113.

Zhang Q, Calus MPL, Guldbbrandtsen B, Lund, MS, Sahana G (2015) Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. *BMC Genetics* 16: 88.

Table 1: Summary of the Top 20 questions in farm animal genomics research identified by the participants of the Cardiff symposium. Frequencies are not included for each question and the questions are not listed in rank order.

Question #	
1. Next generation phenotyping	The mismatch between molecular and phenotypic data has increased dramatically. Which key phenotypic traits should be used as common measures for diversity studies to define breed characteristics in the face of climate change?
2. Genome-wide SNP assays	The identification of common reference genomes and test panels of individuals for SNP array development in less commercial and/or local populations is key. Which strategy shall be used to enable SNP arrays to be developed in a rapid, cost effective and widely applicable manner?
3. Reference genomes	Which common reference genomes and test panels of individuals should be used for array development and diversity studies?
4. E(environment)WAS	How to characterise environmental parameters in extensive production systems?
5. Epigenetics	How can epigenomic information be integrated with phenotypic and genomic data to scrutinize the biological basis for adaptation and plasticity/resilience in livestock populations?
6. Male-mediated genetic diversity	Which methodological approach can be applied to promote reliable assembly of the Y-chromosome, still lacking for many livestock species, as well as to develop polymorphic Y-chromosome markers?
7. Ancient DNA & paleoenvironmental analyses	Which strategies should be followed to collect zooarchaeological specimens from critical geographic sites and promote the analysis of ancient genomes?
8. Conservation of genomic diversity	How to design a management program that evaluates genomic regions for conservation?
9. Polygenic adaptive & economic traits	Haplotypes versus SNPs: in which situations do one or the other provide a more efficient unit of diversity in QTL regions?
10. Microsatellites (STRs) versus SNPs	How to integrate data from the STRs and SNPs, and how to manage the transition from STR- to SNP-based characterisation of FAnGR?
11. GW diversity statistics	Which combination of parameters will be required to adequately summarise genome diversity?
12. Data management	How can links between major FAnGR databases be promoted to be able to federate resources and act as an educational central point?
13. Data availability	Which format should be used to make NGS, phenotyping and GIS data publicly available, and how can industry contribute towards population and maintenance of such database?
14. Participatory projects	How can participatory projects, including citizen science, for example, the use of smart-phone technologies be encouraged to enable data collection on FAnGR at a large scale?
15. Prioritization for conservation	Why are prioritization methods not being applied by policy makers and managers and is there a lack of dissemination or penetrance?
16. Genomic	How to implement genomic approaches systematically in conservation

prioritization	prioritization to include genes important in functionally valuable traits?
17. Utilization in practice	How to reconcile the cost of genomic analysis versus the economic returns on genotyped stock to allow for a wider use of genomic data to assist conservation, production and management of FAnGR? What is the demand and willingness to pay within the sector?
18. Systematic collection	How to ensure that genetic and genomic data are collected sufficiently systematically to be applied to new indicators?
19. Defining goals	Which indicators can be applied to most efficiently monitor genetic trends in domestic populations?
20. Strategic approach	How will the latest advances in 'omics technology contribute to achieve the ultimate goal of halting the loss of biodiversity of FAnGR?

In review