University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Department of Animal Science: Faculty Publications

Department of Animal Science

9-18-2023

Animal board invited review: Practical applications of genomic information in livestock

D. P. Berry

Matthew L. Spangler

Follow this and additional works at: https://digitalcommons.unl.edu/animalscifacpub

 Part of the Genetics and Genomics Commons, and the Meat Science Commons

This Article is brought to you for free and open access by the Department of Animal Science at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Department of Animal Science: Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Animal 17 (2023) 100996

Contents lists available at ScienceDirect

Animal
The international journal of animal biosciences

Animal board invited review: Practical applications of genomic information in livestock

D.P. Berry [a,*], M.L. Spangler [b]

[a] Animal & Grassland Research and Innovation Centre, Teagasc, Moorepark, Fermoy, Cork, Ireland
[b] Department of Animal Science, University of Nebraska-Lincoln, Lincoln, NE, United States

ARTICLE INFO

Article history:
Received 29 August 2023
Revised 8 September 2023
Accepted 11 September 2023
Available online 18 September 2023

Keywords:
Decision support
DNA
Genotype
Inbreeding
Parentage

ABSTRACT

Access to high-dimensional genomic information in many livestock species is accelerating. This has been greatly aided not only by continual reductions in genotyping costs but also an expansion in the services available that leverage genomic information to create a greater return-on-investment. Genomic information on individual animals has many uses including (1) parentage verification and discovery, (2) traceability, (3) karyotyping, (4) sex determination, (5) reporting and monitoring of mutations conferring major effects or congenital defects, (6) better estimating inbreeding of individuals and coancestry among individuals, (7) mating advice, (8) determining breed composition, (9) enabling precision management, and (10) genomic evaluations; genomic evaluations exploit genome-wide genotype information to improve the accuracy of predicting an animal's (and by extension its progeny's) genetic merit. Genomic data also provide a huge resource for research, albeit the outcome from this research, if successful, should eventually be realised through one of the ten applications already mentioned. The process for generating a genotype all the way from sample procurement to identifying erroneous genotypes is described, as are the steps that should be considered when developing a bespoke genotyping panel for practical application.

© 2023 The Author(s). Published by Elsevier B.V. on behalf of The Animal Consortium. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Implications

Genotypes are now available on an ever-growing population of breeding animals with the ever-reducing costs contributing also to a greater availability of genomic information on production animals. Technology adoption is strongly influenced by the associated return-on-investment; such return-on-investment from genotyping is a function of the information that can be generated from the genotype and how this can be used in supporting value-creating decisions making.

Introduction

Genomics is the study of the structure, function and intragenomic interactions and editing of the genome (i.e., DNA sequence in every cell). The first draft of the human genome was published in 2004 (International Human Genome Sequencing Consortium, 2004) while that of the bovine, porcine, and ovine genomes were published in 2009 (Zimin et al., 2009), 2012 (Groenen et al., 2012) and 2014 (Jiang et al., 2014), respectively. Having a draft genome for a species has many uses including helping determine the location of genes and mutations of likely large effect. Knowing which genomic markers are segregating (i.e., varying) in a population and their location relative to genes is useful for research studies attempting to identify genomic variants causing phenotypic differences. Accurate knowledge of genomic positions also aids in determining co-located single nucleotide polymorphisms (**SNPs**) which has implications for genomic evaluations (using haplotypes; Calus et al., 2008) and the process of imputation which involves predicting missing genotypes from adjacent genotype arrangements (i.e., haplotypes).

Many different types of genomic markers exist but SNPs are currently the genomic marker of choice. An SNP may be defined as a variation (sometimes called a polymorphism) between individuals of the same species at a single location of the DNA. Other variations to the structure of the DNA also exist including copy number variants (Rafter et al., 2018), insertions, and deletions (Väli et al., 2008). Nonetheless, microsatellites continue to be used in some species and countries to undertake both parentage testing and forensics as well as being used in some population genetic studies. The general shift towards SNPs being the marker of choice is attributable to their amenability for automatic calling of the

* Corresponding author.
E-mail address: Donagh.Berry@teagasc.ie (D.P. Berry).

https://doi.org/10.1016/j.animal.2023.100996
1751-7311/© 2023 The Author(s). Published by Elsevier B.V. on behalf of The Animal Consortium.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

genotypes which leads to higher throughput and thus lower cost per SNP. Schlötterer (2004) described the other advantages of SNPs namely their low mutation rate, high frequency within the genome and their excellent reproducibility across genotyping platforms and laboratories (Berry et al., 2016) making them more amenable to sharing and inclusion in meta-analyses.

Arguably one of the most impactful developments in practical genomics applied to livestock was the commercial availability of what are commonly termed SNP chips, beadchips or SNP microarrays. SNP chips are premade microarrays which enable the establishment of the genomic variant (i.e., allele) at (hundreds of) thousands of locations (i.e., loci) across the genome from just a single sample and hybridisation step. While a plethora of genotype panels now exist across livestock species, the density of panels most common in the past two decades is 50 000 SNPs (generally termed *50 K chips*; https://assets.thermofisher.com/TFS-Assets/GSD/brochures/axiom-genotyping-arrays-agrigenomics-brochure.pdf; https://www.illumina.com/content/dam/illumina-marketing/documents/applications/agrigenomics/information-sheet-array-consortia-agrigenomics-web.pdf). The availability of these SNP chips has enabled high-throughput, low-cost genotyping of livestock with many downstream practical applications.

In an attempt to lower the cost per animal genotype, efforts were made to evaluate the potential to reduce the number of SNPs on the genotyping panels (Judge et al., 2016; Boichard et al., 2012). Nonetheless, the cost of the SNP chip itself represented a relatively small percentage of the entire cost of generating a genotype for an animal since other costs like sample acquisition, DNA extraction, actually generating the genotypes from the DNA and running and publishing the genomic evaluations are also part of the entire cost. Hence, the cost savings from reducing panel density and thus the cost of the actual genotype panel itself were diluted over all other costs. If interested in simply generating accurate within-breed genomic evaluations, then circa. 50 000 DNA markers are adequate (VanRaden et al., 2011a; 2011b). Genotyping SNP chips could, however, soon be replaced by genotype-by-sequencing based on low-pass sequencing approaches (Li et al., 2021) although initially clients may still elect to extract circa. 50 000 markers for routine use but with a greater density of markers being available on request (at a cost).

The objective of this review is to outline the steps to generate a usable genotype on an individual followed by a description of each of the possible practical applications a genotype can have in livestock breeding and management. The review concludes with steps to consider if developing a bespoke genotype panel. Many of the points made are agnostic to the genotyping technology used (i.e., SNP arrays, whole-genome sequencing, low-pass sequencing).

From sample to genotype

The word genotype, which originates from the word *Gynotypus* meaning "type of genes", is often used to refer to the genetic variant an individual possesses at a given position (i.e., locus) on its genome or indeed across its entire genome (i.e., DNA sequence). Several steps make up the pipeline to generate a genotype with each step not only incurring a cost but also introducing a possible source of error. Here, the genotyping process refers to any technology (i.e., SNP arrays, genotype-by-sequencing) that generates a genotype. The steps are as follows:

Sample acquisition

The quality but especially the quantity of DNA differs not only by biological sample type but also how it was obtained, including storage (Abraham et al., 2012). The sample type used in domesticated species was traditionally hair follicles (cattle), semen (artificial insemination males) or Flinders Technology Associates filter cards (sheep); this has been replaced in many jurisdictions by ear biopsies obtained when tagging an animal for identification purposes. The cost of sample acquisition includes the logistics of delivering the sampling tool, the cost of the mechanism itself in the case of a tag or the cost of the semen straw, and finally the cost of returning the sample to a central depot for genotyping. Cost per sample is expected to reduce as the number of samples increases because the cost of delivering and returning the samples is diluted over more units. Sources of error include the sampling of an animal different to that intended and incorrectly labelling the sample.

Deoxyribonucleic acid extraction

The preferred standard operating procedure (**SOP**) for DNA extraction differs not only by biological sample (Edson et al., 2021) but also among laboratories; the extraction SOP also has a major bearing on the resulting DNA quality (Psifidi et al., 2015). Minimum quality and quantity thresholds of DNA required are specified for each genotyping or sequencing method. There is a cost associated with DNA extraction especially if cleaning of the sample or amplification of the DNA is required. The cost of DNA extraction generally benefits from economies of scale.

In most instances when samples are sent for genotyping or sequencing, all of the tissue sample is used when extracting DNA and the extracted DNA is usually stored, at least for some time, either in the laboratory of the service provider or returned to the customer for storage. Whether every sample should be stored indefinitely is questionable albeit having a biobank of samples of interest should be considered such as those from influential animals, animals with defects, a pool of samples segregating for specific alleles or indeed rare breeds.

Genotyping and sequencing

The decision on which genotyping platform to use (which includes sequencing as a strategy for genotyping) is a function of several factors including:

- motivation for genotyping (i.e., parentage testing, candidate gene (research), genomic evaluations, fine mapping),
- the platform(s) available,
- the available budget as a function of the cost per sample, and
- the value and role (i.e., seed stock or commercial) of the individual to be genotyped.

Parentage analysis was traditionally undertaken using microsatellite markers but this has either already transitioned or is in the process of transitioning to SNPs. Candidate gene studies or genotyping just a few markers for use in selection decisions (i.e., marker-assisted selection; Dekkers, 2004) is rapidly being replaced by genotyping tens (or hundreds) of thousands of SNPs. One reason for such a transition is the minimal cost differential between genotyping a few markers versus several thousand; the same is true for parentage testing with most now opting to simply genotype animals for tens of thousands of SNPs which almost always include the International Society for Animal Genetics panel SNPs for parentage testing (McClure et al., 2018). Generating data on tens of thousands of SNPs provides multiples more information enabling a greater return-on-investment. Many commercially available genotype panels now exist for a range of different species (https://assets.thermofisher.com/TFS-Assets/GSD/brochures/axiom-genotyping-arrays-agrigenomics-brochure.pdf; https://www.illumina.com/content/dam/illumina-marketing/documents/applications/

agrigenomics/information-sheet-array-consortia-agrigenomics-web.pdf), with some bespoke genotype panels also in existence albeit not always freely available (e.g., Eurogenomics, International Dairy and Beef, Neogen). In general, the greater the number of markers being genotyped, the greater the cost of the genotyping process although the costs of sample procurement and DNA extraction are relatively constant, the exception being for sequencing.

Interest is growing on the use of genotyping-by-sequencing as a means of generating genotype information on individual animals, especially as a means of 'animal-side' rapid genotyping (Lamb, 2023). This is being made possible with the rapid development of technological solutions developed by companies like Oxford Nanopore (Jain et al., 2016). The usefulness of these developing technologies in the field both in relation to practical utility but also cost relative to SNP arrays is an area of active research (Lamb, 2023). The promise of such approaches is to enable capturing rare variation, avoiding the need to re-build SNP assays, and to ultimately increase the accuracy of genomic predictions by including putative causative variants directly into prediction models. Irrespective, cross-compatibility between genotyping and sequencing platforms is paramount enabling imputation to a common set of genomic markers. This is especially true for some sequencing approaches whereby the SNPs that exist on most arrays are also present in the generated sequence data.

There are three main costs to the genotyping. In the case of chip genotyping, it is the cost of the SNP chip hardware itself, the cost of the associated reagents, and the cost of the service; the latter includes the return-on-capital (i.e., infrastructure, people, time) and profit. For sequencing, the costs include the reagents and the costs of the service. Many sequencing service providers also offer downstream bioinformatic analyses as a service including data transfer.

While hybridising DNA across an SNP chip is one strategy to generate genotypes, it is also possible to generate genotypes of individuals in-silico where genotypes are also available on their progeny (cattle – Berry et al., 2014; sheep – Berry et al., 2018a; pigs – See et al., 2022). In the absence of genotypes from the mate, genotypes from >10 progeny can be used to reconstruct the genotype of the parent with high accuracy; this number of progeny reduces to $\geq$7 if the mate's genotype is also available. Such a strategy is particularly useful when a biological sample of one parent is no longer available or in cases where the budget might prohibit an entire tier of the breeding pyramid (e.g., multiplier) from being genotyped.

Bioinformatics

The output from most commercially available SNP chips is relatively standard. It is also possible to generate such a format from sequencing approaches based on the most likely genotype. The basic information is the sample identification, the (customer) SNP name and the called genotype. Additional information in the genotype output files include quality metrics of the called genotypes which are at an SNP level (e.g., GenTrain score for Illumina genotypes), but also at the individual SNP-by-individual sample level (i.e., GenCall score for Illumina and Confidence score for Thermoscientific). The GenTrain score of genotypes called by the Illumina platform considers the quality and shape of the clusters of called genotypes (i.e., homozygous versus heterozygous) and relative distance from each other (More et al., 2019). SNPs with poor genotype cluster resolution indicated by a GenTrain score < 0.55 (Judge et al., 2016; Zhao et al., 2015) to <0.60 (Erbe et al., 2012) are often discarded. The Gencall score generated by the Illumina software is a measure of the confidence assigned to the called genotype for each SNP on each individual and is used to filter poor-quality calls. A Gencall score lower than 0.15 suggests failed genotypes and is considered the default threshold in Illumina's Genome Studio (https://www.illumina.com/Documents/products/technotes/technote_infinium_genotyping_data_analysis.pdf). Berry et al. (2021) documented how the concordance rate between genotypes called with high confidence (i.e., high Gencall score) versus genotypes called for the same SNP in a duplicate animal eroded as the Gencall score of duplicate genotype reduced. Nonetheless, the impact was not very noticeable until a Gencall score of <0.55 was achieved.

Other data fields in the genotyping output files are also available including the X and Y signal intensity for each SNP-by-individual sample, the B allele frequency and the logarithm normalised R ratios. Each intensity value per SNP represents the intensity for each channel of the fluorescent dyes associated with the two alleles of the SNP under investigation. The B allele and logarithm normalised R ratios outputs can be used in the detection of copy number variants, indels, or karyotype abnormalities. Additional output files summarise call rates per individual and mean frequencies of different genotypes per animal but these can be derived from the other output files themselves.

The downstream bioinformatic analysis needed is a function of the requirements of the client. Many clients simply want the called genotypes which are the input for already established pipelines for actions such as parentage testing (Dodds et al., 2005) or genomic evaluations (VanRaden, 2008). Nonetheless, for producers, these services still incur a cost for data processing.

Genotype quality control

The genotype quality control procedures adopted depend on the end use but also on the statistical procedures that are employed to arrive at the end goal. Service providers sometimes impose their own genotype quality control measures prior to sending the genotypes to the client – these could include the omissions of some SNP probes with poor performance or those of proprietary value. One of the first quality control measures imposed by the customer is the genotype call rate per individual. The call rate is the proportion of genotypes on the SNP panel where a genotype was called (with confidence); this is usually a measure of DNA quality or concentration unless the majority of samples have a poor call rate which may necessitate deeper investigation. The minimum call rate threshold per animal imposed in livestock generally varies from 85 to 95% (Wiggans et al., 2010; Purfield et al., 2016; Croué et al., 2019) but the end use of the data, and possibly even the relative importance of the individual, should be taken into consideration when setting the threshold. For example, a lower call rate (e.g., 85%; Purfield et al., 2016) could be used for routine genomic evaluations of commercial animals while a higher call rate may be imposed when individuals are participants in research studies. Moreover, some apply different call rates per SNP for autosomes versus sex chromosomes (e.g., Wiggans et al., 2010) as well as by minor allele frequency (e.g., Wiggans et al., 2010). Irrespective, sporadically missing genotypes remaining after the call rate edit are generally filled in using a process of imputation (Li et al., 2009). Whereas an individual may have a high call rate, excessive heterozygosity in samples should also be examined; excessive heterozygosity may be indicative of sample contamination.

Where ancestry information exists, autosomal SNPs not adhering to expected Mendelian inheritance are often discarded. This should ideally be undertaken iteratively since if, for example, 10% parentage errors exist in a population then many SNPs may exhibit high Mendelian errors, but also if many SNP genotypes are poorly called and subsequently flagged as Mendelian errors then recorded ancestral relationships may be severed. Calus et al. (2011) described a procedure to test for Mendelian inconsistencies in SNP genotypes between putative parent-offspring pairs. Conflicts

between genotype-predicted and recorded parentage may warrant further interrogation, especially where a dispute with the owner exists. One such further evaluation could consider hemizygosity or indeed a deletion; procedures and evidence for such have been described by Berry et al. (2019) for sheep.

Studies often impose a restriction on the frequency of the least common allele – this is termed the minor allele frequency (**MAF**). The rational for including a minimum threshold is three-fold: (1) like any feature in a statistical model, large sample sizes are required to detect (small) associations in features demonstrating very little variability, (2) algorithms can have difficulty properly calling genotypes in the calibration process where low numbers of heterozygotes and homozygotes for one allele exist (Anderson et al., 2010), and (3) low MAF SNPs tend not to contain much information for polygenic traits and thus excluding them could reduce the computational requirements. Nonetheless, rare alleles may contribute to genetic variability (Jang et al., 2022) so should be discarded with caution. The cited minimum minor allele frequency in studies is usually on a percentage basis but in reality, this should be based on the actual count of the allele since for the same percentage allele frequency, more individuals will be carrying the allele as the total population size increases. Examples of minimum MAF imposed in livestock vary from 0.1% (Zhou et al., 2014) to 5% (Chen et al., 2013). Population sub-structure is important to consider when calculating MAF as is the end use of the genotypes (e.g., breed assignment especially for rare breeds).

Most genomic studies also tend to discard SNPs where the genotype frequencies deviate from Hardy-Weinberg equilibrium. While deviations from Hardy-Weinberg equilibrium for reported genotype data can occur due to genotype errors, actual deviations from Hardy-Weinberg equilibrium can be due to the actions of evolutionary forces, one of which is selection. Given that many genomic studies attempt to identify loci under selection, care therefore needs to be exercised when discarding SNPs based on their deviation from Hardy-Weinberg. An extreme example is lethal recessive mutations where no homozygous mutant genotype should be detected despite the presence of heterozygotes; this SNP, which is arguably hugely informative, will deviate from Hardy-Weinberg equilibrium and therefore, in some studies, may not be included in the analysis. Relaxed thresholds are warranted such as $P < 10^{-6}$ (Panetto et al., 2017) and 10^{-7} (Junqueira et al., 2017). Imposition of Hardy-Weinberg edits in collated dataset across different populations should also be undertaken with caution – some populations may have a high frequency of homozygotes for one allele while another population may have a high frequency of homozygotes for the other allele; a combined analysis may reveal a deviation from Hardy-Weinberg equilibrium but this is simply due to population structure and this SNP would be very informative for assigning animals to their respective population.

Additional quality control tests that should be considered, but can depend on the analysis being performed, is to test for the presence of duplicate genotypes (which could be monozygotic twins). Checking gender of the individual (discussed later) as well as the provenance (e.g., population, breed) are also important. To reduce computational time and resources, removing redundant SNPs may also be undertaken. These are SNPs which are in complete linkage disequilibrium so that the marginal information in all SNPs over and above just one of those SNPs is zero; as the population of genotyped animals grow, fewer SNPs in complete linkage disequilibrium may exist. In such situations, the SNP with the higher call rate may be retained.

In all, many factors in the genotyping/sequencing pipeline can affect the actual called genotypes. The sampling procedure and DNA extraction SOP, for example, can affect the called genotype (e.g., call rate, heterozygosity rate). DNA quality, therefore, should not be compromised in the pursuit of reduced genotyping costs. While SNP arrays may be less sensitive to DNA quality, genotype-by-sequencing technologies may not (Lamb, 2023). Hence, all actors in the pipelines (i.e., manufacturers of sampling apparatuses, producers who take the sample, laboratories who extract the DNA, genotyping and sequencing service providers, bioinformaticians) should be engaged in the process with good communication channels and the appropriate (and timely) quality control mechanisms and protocols in place at different stages of the sample-to-genotype pipeline.

Practical uses of genomic information

The practical uses of genomic data in helping make more informed breeding and management decisions are in Fig. 1. Animal-level genomic information is also hugely informative for scientists in designing experiments or properly accounting for genetic structure in cross-sectional data analyses.

Parentage verification and discovery

Ensuring accurate parentage is not only important for generating accurate genetic evaluations to maximise genetic gain (Visscher et al., 2002; Sanders et al., 2006), but is also important for informing mating decisions and estimating the expected coancestry (i.e., half the genetic relationships) among individuals; the expected inbreeding of the mating of two individuals is equal to the coancestry between the two mates.

The impact on genetic gain from parentage errors is a function of both the heritability of the trait as well as the number of progeny records contributing to the genetic evaluation (Fig. 2; Visscher et al., 2002). Irrespective of heritability, the greater the rate of the parentage errors the greater the impact on genetic gain. Furthermore, the impact of parentage errors on an animal's estimate of genetic merit diminishes as the animal accumulates more (recorded) progeny with the impact reducing faster for higher heritability traits (Fig. 2). Inbreeding, on average, reduces performance, especially if between close relatives; this is termed inbreeding depression and the effects have been well publicised (McParland et al., 2007 and 2008; Selvaggi et al., 2010; Silió et al., 2013).

Accurate recording of parentage negates the requirement for genotyping (solely for parentage validation or discovery) but this is not always possible such as in extensive production systems, where mob mating occurs, where heterospermic inseminations are used, or in situations of superfecundation in polyovulatory species. Even in large well-managed herds (e.g., dairy herds in New Zealand), it is not always possible to match calves to their dams, and thus also their sires.

Using the knowledge that each parent transmits half its DNA to its offspring, having a genotype on both the offspring and its parents enables accurate parentage verification. If sufficient genetic markers as used, then parentage discovery can also be successful assuming the true parents' genotypes are also available (Dodds et al., 2005). Estimates of parent-to-offspring parentage errors vary from 7.6 to 10.0% in sheep (Berry et al., 2016), from 10.00 to 13.28% in cattle (Visscher et al., 2002; Purfield et al., 2016; Řehout et al., 2006), and from 8.4 to 14.6% in goats (Bolormaa et al., 2008).

If appropriately chosen, the number of SNPs needed for parentage verification is between 300 and 400 in most livestock (Berry et al., 2019; Van Doormaal, 2016; Strucken et al., 2017). Certainly, using the now almost ubiquitous 50 K SNP chips is sufficient to verify and discover parentage. Such dense genotypes can also be useful to confirm or assign grandparents and even great-grandparents (VanRaden et al., 2013). Importantly, however, to maintain integrity, sample mix-up should be avoided when taking

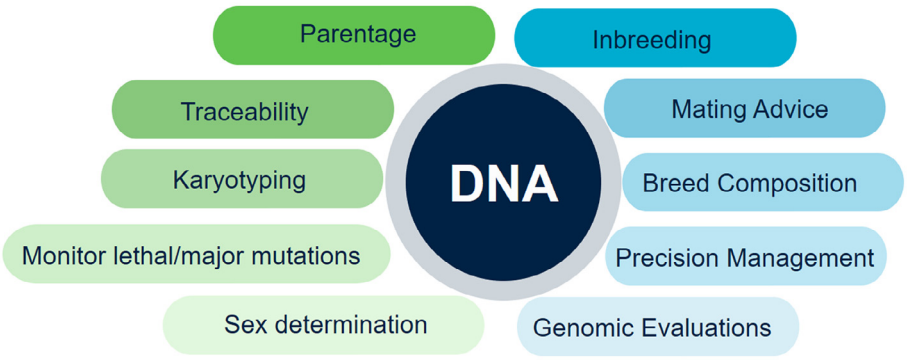

Fig. 1. Applications of genomic information in livestock breeding and management.

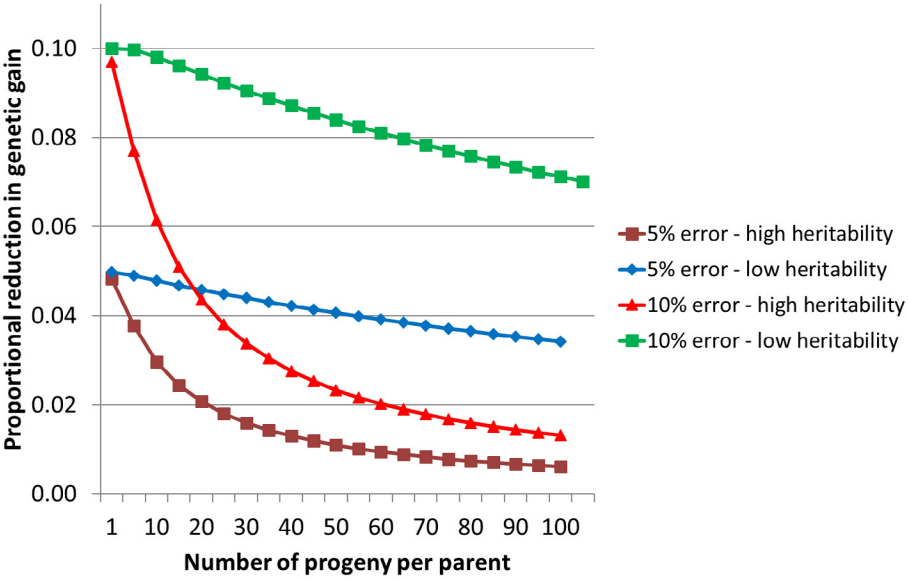

Fig. 2. Impact of parentage errors in any species on genetic gain for high heritability (0.30) and low heritability (0.02) traits with 5% or 10% parentage errors.

the biological sample. Having relatively accurate dates of birth is also important especially for parentage discovery; because parentage discovery is based on the pretence that parent-offspring pairs share half their genome, then blindly undertaking parentage discovery without considering the date of birth could result in a progeny being proposed as a parent.

Traceability

Traceability has uses not only for reassuring the consumer of the ability to trace a product (e.g., steak) to an individual, but is also useful in reconciling animals under situations where animal hustling (i.e., stealing) may be a problem, or in research settings by ensuring correct animal identification when biological samples are taken from commercial abattoirs. Like for other applications of genomic technologies, the number of SNPs required for traceability is dependent on their informativeness (Weller et al., 2006). The probability of two unrelated individuals carrying the same genotype at 10 SNPs with a minor allele frequency of 0.5 is 5.5×10^{-5} (Fig. 3). The number of SNPs on the commonly used 50 K chip is therefore well sufficient to enable accurate traceability.

Karyotyping

A karyotype is a term used to describe the appearance and number of chromosomes in a cell. Alterations to either the number of chromosomes or their appearance is called a karyotype abnormality, with the resulting repercussion varying from no observable external difference, to infertility or embryo death (Holečková et al., 2021). Like any abnormality, (early) detection is important so that the individual can enter an appropriate production cycle (i.e., to enter the breeding herd or reared for meat production). It is possible to identify aneuploidy, using SNP chips, by investigating the intensity values of the SNPs on each chromosome; this is akin to the detection of large copy number variants. An automated approach to detecting monosomy from SNP chips has been described in cattle (Berry et al, 2017), sheep (Berry et al, 2018b) and humans (Prakash et al., 2014). Is it not possible to identify all karyotype abnormalities (e.g., translocations and inversions) using SNP chips for some species. Nonetheless, identifying some karyotype abnormalities early in life is important in aiding management decisions since some karyotype abnormalities can result in infertility (Romano et al., 2015; Berry et al., 2017). Being able to identify (some) karyotype abnormalities requires the B allele

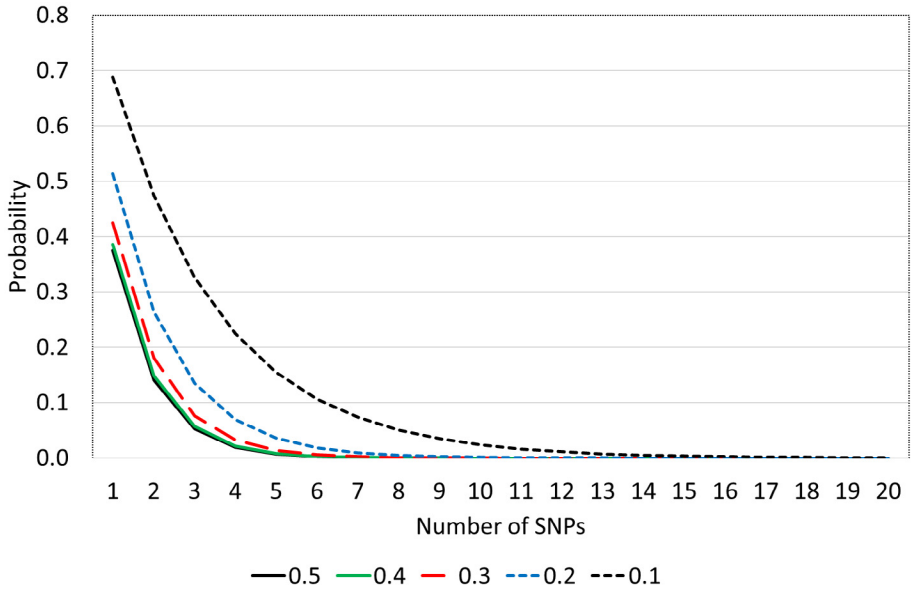

Fig. 3. Probability that two unrelated individuals of the same species will have the same genotype depending on the mean minor allele frequency (0.1–0.5) of the chosen single nucleotide polymorphisms (SNPs) and the number of SNPs derived using the equation provided by Weller et al. (2006).

frequency (i.e., normalised measure of the allelic intensity ratio of two alleles), logarithm normalised R ratios, and the X and Y signal intensity values of all called genotypes. The intensity values of each SNP along each chromosome can then either be visually examined or compared to the mean intensity values of the other chromosomes for the same individual (Berry et al., 2017; 2018b). Sufficient SNPs exist in the 50 K SNP chips to enable these discoveries. Practical applications of karyotyping individuals have been demonstrated for cattle (Bouwman and Mullaart, 2023).

Although not a karyotype abnormality, but considered a chromosomal abnormality, uniparental disomy (Engel, 1980) is when an individual inherits two copies of a chromosome, or even part of a chromosome, from one parent with no copy inherited from the other parent. Uniparental disomy can be the result of heterodisomy, which occurs when a pair of non-identical chromosomes are inherited from one parent or it can be due to isodisomy, where a single chromosome from one parent is duplicated. Most occurrences of uniparental disomy result in no phenotypic anomalies. Uniparental disomy can be detected where one, or ideally both of the parents are genotyped along with the individual itself and the 50 K SNP chip is sufficient to undertake the analysis; the genotype of the individual is compared to that of the parent(s). It should be noted that uniparental disomy and some other karyotype abnormalities could contribute to apparent parentage conflicts.

A freemartin is a chimeric (XX/XY) female that is infertile resulting from mixed-sex twin pregnancy in some species (e.g., cattle). The blood system of the two placentas is not fully disengaged so the blood of both twins can flow around the other. The presence of male Y chromosome SNP can be detected in the white blood cells of the suspected freemartin female soon postbirth. No evidence exists on the ability to detect freemartinism using current SNP array genotypes generated from DNA extracted from hair follicles or ear biopsies.

Sex determination

While visually determining the sex of an animal is not complicated, procedures like rubber-ring castration of young animals can make it a little more challenging in real time. Moreover, where preprogrammed automatic drafting options linked to radio frequency identification systems exist, drafting animals on gender, breed, and genetic merit can be readily undertaken if the genotype information and output are stored in the backend information systems. Furthermore, genotype-determined gender can also be a useful quality control measure in the genotyping pipeline process itself (McClure et al., 2018).

It is advisable to use the genotypes from both the Y chromosome and non-pseudoautosomal region of the X chromosome when determining animal sex. Because males in most livestock species have only one X chromosome, no heterozygous genotypes should exist in the non-pseudoautosomal region. Moreover, genotypes should be called for SNPs on the Y chromosome in males but not females. Relying just on the extent of homozygosity on the X chromosome may misidentify inbred females who inherited the same intact X chromosome from a common ancestor. If semen is the biological sample being genotyped, and if it is female-sex sorted, then very few Y chromosome carrying sperm may be present and leading to faint or no genotype called on the Y chromosome. Karyotype abnormalities can also cause discrepancies between genotyped-predicted gender and actual gender (e.g., Swyer syndrome; Berry et al., 2023).

Breed composition

Accurate knowledge of the breed composition of an individual is important for (1) developing a mating programme to fully exploit inter-breed non-additive genetic effects, (2) understanding the history leading to the breed composition of the individual (e.g. differentiating first cross vs stablised composite) which could be useful for modelling the non-additive effects (e.g., heterosis) in genetic evaluations models, (3) help in maintaining the integrity of breed society records and breed conservation programmes where a restriction is imposed on breed purity, (4) reassuring consumers of the authenticity of animal products that command a higher price based on breed origin, (5) as an adjustment factor in multi-breed genetic evaluations, (6) as an early warning system of errors in mislabelling during sample procurement or genotyping, and (7) experimental design where one of the factors of interest is breed/crossbred (Zimmermann et al., 2021; Schiermiester et al., 2015) differences. The expected breed composition of an

individual is the average of its respective parents' breed composition. The actual breed composition of an individual, however, may deviate from the expectation where one or both parents is crossbred owing to the segregation of alleles during gamete formation. Several approaches exist to estimate breed composition from genomic data (O'Brien et al., 2020; Kuehn et al., 2011; Alexander et al., 2009).

Tens of SNPs are required to assign purebred animals to individual breeds (Hulsegge et al., 2013) but several hundred SNPs are required to estimate breed composition (O'Brien et al., 2020; Judge et al., 2017; Strucken et al., 2017), especially in a highly admixed population. Informative SNPs are those generally with a high fixation index (Weir and Hill, 2002) scattered across the genome (Judge et al., 2017); the fixation index reflects the genetic differentiation between populations.

Monitoring major genes or mutations conferring congenital or large effects

A major gene may be defined as a gene with pronounced phenotypic expression. Casas and Kehrli (2016) provided a summary of the detected major genes of known effect in cattle with VanRaden et al. (2011a and 2011b) reporting additional recessive mutations in Holsteins. Some mutations are due to single or dinucleotide polymorphisms (e.g., K232A in DGAT1) while others can be due to deletions (e.g., nt821 deletion in the myostatin gene). The mode of action of the different mutations can differ although those conferring lethality are almost always recessive; in such situations, an individual must be carrying two copies of the mutant variant to be affected. Some alleles exhibit dominance action in that one allele masks or overrides the effect of the other allele at the same locus; an example is the dominant allele conferring polledness in cattle (Medugorac et al., 2012).

Monitoring the change in frequency of different alleles of such large effect over time can be of interest for future proofing; for example, a growing frequency of lethal recessives in a population may have long-term unfavourable consequences. Some breed societies refuse to register individuals carrying alleles for particular variants (e.g., some beef breeds do not accept animals carrying the nt821 myostatin variant). Knowledge of the carrier status of individual animals is also informative when deciding on matings to produce the next generation. It is not clear if including the actual causal mutation in genomic evaluations improves the accuracy of genomic predictions once genome-wide SNP information is also included (Oget et al., 2019) although the benefit, or lack thereof, will depend on a number of factors like (1) the ability to actually identify such variants and the allele substitution effect of the locus, (2) the density of genotyped SNPs in the vicinity of the major locus and their linkage disequilibrium structure, (3) whether evaluations are being undertaken within or across breed and the linkage disequilibrium structure (and haplotype phase) in that region, and (4) the size of the reference population of phenotyped and genotyped animals for the genome-wide markers.

The sequence flanking the mutation of interest can readily be included as primers on SNP chips for genotyping along with all other thousands of SNPs (e.g., https://www.icbf.com/wp-content/uploads/2014/06/Farmer-Genetic-Disease-and-Trait-Information-for-IDB-Genotyped-Animals-in-Ireland_9_20_16.pdf); this therefore requires just one biological sample and one process of genotyping. Some mutations are, however, royalty bearing.

Relatedness, coancestry, inbreeding and heterosis

Knowledge of the relationships among individuals, especially how the set of relationships change over time, is important in the pursuit of minimising the accumulation of inbreeding and its generally unfavourable repercussions for animal performance through inbreeding depression (McParland et al., 2007, 2008; Selvaggi et al., 2010; Silió et al., 2013). Whereas the expected relationship among full sibs is 50%, a SD of 4% has been reported (Wang et al., 2014; Kenny et al., 2023). This implies that for approximately 0.1% of full sibs (i.e., 1 in every 1 000), they are expected to share <38% of their genome while for another approximately 0.1%, the pair are expected to share >62% of their genome. The extent of this variability in relationships between full sibs is also dependent on the inbreeding coefficient of the parents and their relationship (Kenny et al., 2023). The commonly used 50 K SNP chip provides sufficient SNPs to enable accurate relationships to be quantified – in fact, the 50 K genotypes are used to generate genomic relationship matrices which form the basis of most genomic evaluations for livestock species globally. The genomic relationship matrix (VanRaden, 2008) represents the relationships among genotyped animals calculated from their respective genotype and several approaches exist to calculate the genomic relationship matrix (e.g., VanRaden, 2008).

In the absence of genomic information, the inbreeding statistic applied to an animal is the expected inbreeding based on the tracing of (sometimes incorrect as previously discussed) ancestry to common ancestors (if they existed). Genotype information can be used to more accurately estimate homozygosity in the genome (Purfield et al., 2012; Lopes et al., 2013) some of which can be due to inbreeding. Runs of homozygosity use genomic data to form sets of contiguous loci that are homozygous and offer the opportunity to predict inbreeding in, yet unknown, key regions of the genome that lead to trait-specific inbreeding depression as opposed to global estimates of inbreeding (Pilon et al., 2021). Nonetheless, inbreeding metrics per animal continue to be generally expressed based on homozygosity across the entire genome. However, inbreeding should only cause an effect if (a) it is for loci that affect performance, and (b) the homozygous genotype is inferior to the other possible genotypes in which case, if known, it may be purged from the population through selective breeding. Therefore, loci- or runs of homozygosity-specific measures of inbreeding (or indeed coancestry) may be more informative. Such information, though, could become unwieldy so therefore it may be more appropriate to integrate (expected) inbreeding depression into expected performance predictions when informing potential matings. Heterosis is simply the opposite of inbreeding. Therefore, the same principles of loci-specific heterosis effects apply.

Mating advice

The basis underpinning mating advice tools, or decision support systems, is to maximise genetic gain in the herd while also at the same time attempting to correct particular features of females with males excelling in those features and avoid the mating of (closely) related individuals (Kinghorn, 1998; Bérodier et al., 2021). The ability of genomic information to detect carriers of genetic markers conferring different phenotypic attributes have already been discussed and such markers can be considered in the mating algorithm. Similarly, the ability to better estimate the relationship among candidate parents from which to estimate the expected inbreeding of the potential offspring has been discussed. Linear programming is sometimes used in mating programmes (Bérodier et al., 2021; Bengtsson et al., 2022) where the goal is to maximise an objective function (e.g., total genetic merit) under a set of constraints such as upper threshold on coancestry among potential mates.

For commercial producers, an additional objective may be to minimise the expected heterogeneity of phenotypic performance in the offspring while the opposite might be true for seed stock breeders. Knowledge of the gametic variance of individuals can

also help identify animals, but also matings, that minimise (or maximise) the expected heterogeneity of the progeny. Santos et al. (2019) described a deterministic approach to estimate the gametic variance for individuals based on the available genotypes. While several studies have described the potential of such an approach to identify individuals that will create more heterogeneous progeny, and thus possibly create genetically extreme animals for inclusion in a breeding programme (Bijma et al., 2020), the technique can also be used to identify animals that are likely to create more homogeneous progeny more suited to consistent management by producers. Knowing the gametic variance of sires could be useful when evaluating the risk of a difficult birth ensuing from a given mating, especially for the mating of beef bulls to dairy females (Berry, 2021). For such beef-on-dairy matings, the dairy farmer may choose a sire with a greater expected mean genetic predisposition to a difficult birth of his progeny but with a lower chance of producing very large calves requiring veterinary intervention (i.e., a more homogenous group of progeny). The tool could also be particularly useful for polyovulatory species like pigs where minimising litter variability in birth weight and growth rates could be extremely helpful in management.

To estimate the gametic variance of an individual, genotypes for all SNPs included in the respective genomic evaluation should be available. This is necessary because SNP effects estimated in the genomic evaluation process form part of the calculation of gametic variance (Santos et al., 2019).

Genomic evaluations

The potential for genetic markers to enable more accurate genetic selection for complex traits, particularly those that were sex-limited or difficult/expensive to measure, led to both marketing and eventual deployment of marker panels to inform marker-assisted selection. In some species and in some countries, the results of some of the first commercially marketed tests were delivered to producers in the form of scores ranging from 1 to 10 or even more naively as a number of stars whereby more stars were assumed to be genetically superior animals. These initial tests were based on a limited number (a few to tens) of genomic markers. Although such tests were marketed, it became apparent that the results were not robust across populations (Van Eenennaam et al., 2007). Moreover, the results of such tests directly competed with traditional pedigree-based estimates of genetic merit (i.e., estimated breeding values; **EBVs**) as a selection criterion thereby only adding confusion for breeders. In other cases, limited numbers of markers were included in genetic evaluations as fixed effects (Fernando and Grossman, 1989). These initial marker panels evolved to include hundreds and eventually thousands of markers necessitating a different strategy to including DNA-based information into genetic evaluations.

To capture value from this new form of data, entities (i.e., breeding companies, breed associations/societies) began to develop their own within-breed or within-population reference populations comprised of both genotyped and phenotyped animals. These reference sets (also called training populations) were used to estimate the effects of individual SNPs that could then be applied to newly genotyped selection candidates to generate genomic EBV (**GEBV**). The methods for incorporating genomic predictors into EBV were based on either (a) including the GEBV as a correlated trait (Kachman, 2008; MacNeil et al., 2010) in a multi-trait evaluation or (b) employing a blending, or indexing, approach whereby the molecular and traditional estimates of genetic merit were combined proportional to their accuracy (VanRaden, 2008). Irrespective, these two methods were sub-optimal given SNP effects first needed to be estimated from the reference population and then applied to selection candidates and required combining two pieces of information (genomic and pedigree-based EBV) that arose from two different procedures.

The investment in genotyping to construct a representative reference population proved to be a critical first step to achieving the realisation of genomic selection (Meuwissen et al., 2001). While phenotyping is also costly, many of the initial reference populations exploited already existing phenotypic data. In fact, many of the reference populations were based on progeny-tested sires where the accuracy of the estimates of genetic merit was high thus requiring fewer genotyped animals to achieve a high accuracy of selection as would be needed if animals with only their own phenotype were considered. Retraining, or recalibration, which is the process of re-estimating SNP effects, became a necessity as the population of genotyped animals increased in size. The issue of robustness of predictions across breeds arose, and the use of genomic predictors trained in one breed/population could not be used with any degree of accuracy in a closely related breed/population (Kachman et al., 2013) unless the reference population accounted for all potential breeds in the target population. This issue of sensitivity of SNP effects to the composition of the reference population was largely a function of differences in linkage disequilibrium (i.e., how co-located SNPs are inherited) across breeds/populations and even within-breed/population across generations.

A paradigm shift in genetic evaluations occurred when statistical models were developed to enable the inclusion of genomic information in a 'single step' in the calculation of EBVs, avoiding the need for reference populations that were external to the system and the potential for bias that such systems created. One such approach considers genomic data as a means of quantifying kinship among animals (Legarra et al., 2009). A genomic relationship matrix (**G**) describing the actual relationships among genotyped individuals could easily replace the pedigree-based (expected) relationship matrix in genetic evaluation software, resulting in genomic best linear unbiased predictions (VanRaden, 2008). However, the majority of livestock populations contain a mixture of both genotyped and non-genotyped animals. Fundamentally, single-step genomic BLUP combines the traditional pedigree-based expected relationships among non-genotyped animals with relationships calculated from genomic information for genotyped animals. The combining of the pedigree-based relationship matrix (termed the **A** matrix) and the genomic relationship matrix (i.e., the **G** matrix) formed an **H** matrix (Aguilar et al., 2010; Christensen and Lund, 2010). As before, the matrix **H** can be easily substituted into the traditional genetic evaluation software suites.

The primary benefit of relationship estimates derived from genomic information is the ability to determine more refined estimates of relationship between individuals. Classic examples include relationships between grandparents and grand-offspring or between full siblings. The former relationship, based on pedigree and assuming no inbreeding, has an expected value of 0.25 but the range could be between 0 and 0.5 due to Mendelian sampling. The latter example has an expected value of 0.5 but the range could be between 0 and 1; the SD is 0.04 units (Wang et al., 2014; Kenny et al., 2023). Because the **G** matrix can partially capture Mendelian sampling and the recorded ancestry is often incorrect or indeed missing, the generated genomic relationships are a more accurate reflection of relationships thereby contributing to improved accuracy of the EBV (Hayes et al., 2009). A competing single-step method, single-step Bayesian regression as described by Fernando et al. (2014), is equivalent to single-step genomic BLUP under certain circumstances and can be thought of as a hybrid of pedigree-based BLUP and a marker effects model that estimates the effects of each SNP that is included (typically a few thousand) in one unifying statistical model. Regardless of the choice of single-step method, both hold advantages over the

two-step approach primarily because SNP effects are estimated internal to the system and do not rely on an external set of animals. Consequently, both 'single-step' methods have demonstrated benefits in increasing the accuracy of genomic breeding values and have become the methods of choice for the majority of commercial breeding programmes.

Precision management

Precision or personalised management is where the management of an individual is optimised to its genotype. Access to information on parentage, carriers of genes of major effects (including karyotype abnormalities) as well as more accurate prediction of genomic, and therefore phenotypic merit all contribute to more precise management decisions. For example, being able to more accurately estimate the genetic merit of an individual using genomic technologies facilitates more bespoke management strategies such as feeding animals differently contingent on their (estimated) genetic merit. Such precision management is certainly not a novel concept in livestock. For example, energy-rich diets are fed to animals with a genetic potential for rapid muscle growth. Such genetic differentiation of animals, however, has heretofore generally been based on breed. Genomic evaluations provide more granular and precise predictions of within-breed (as well as across breed) performance. Being able to estimate non-additive genetic effects is particularly important when predicting phenotypic performance; non-additive genetic effects are generally ignored in breeding because not all are directly transmitted from one generation to the next. Therefore, it is likely in the future that animals will be penned or managed (i.e., quantity and quality of feed) based on genomic breeding value rather than breed or coat colour. Producers might also use such predictions to inform buying/selling decisions, such that a feedlot might pay more for animals with a genetic predisposition for greater feed efficiency or carcass merit. Similarly, producers may actively seek to purchase animals with a similar genetic predisposition to some diseases; one example could be the avoidance of animals genetically less resistant to tuberculosis by producers farming in areas rife with tuberculosis.

As well as aiding the management of animals, more accurate estimates of genetic merit through genomic evaluations can be used to modify rules to invoke action within decision support systems. As an example, a cow that is predicted to be genetically predisposed to an infection of the uterus or indeed more prone to ovarian cysts may be recommended to receive an ultrasound examination of her reproductive tract prior to the start of the breeding season. Similarly, the threshold somatic cell count level in the milk of a lactating animal that triggers a recommended action may be lowered for cows predicted to be more genetically prone to sub-clinical mastitis. This is akin to life-style recommendations for humans differing in the family history of diseases such as cancers or cardiovascular disease. Genomic predictions could also be used to inform specific pharmaceutical regimes to maximise animal response to vaccines or to reduce the overall use of pharmaceuticals. Again evidence clearly exists in humans on the efficacy of different drugs (e.g. codeine) for people of different genotypes.

Nonetheless, other contributors to precision management from genomic information over and above that already described also exist. One such example is the identification of high milk somatic cell count individuals within a herd by comparing the genotype of the bulk milk to those of the contributing individuals; it is not the actual called genotype of the bulk milk that is used but instead the frequency of the B allele in the genotype results. Such a tactic is already commercially available (GénoCells®). The majority of somatic cells in milk are leukocytes which are elevated in lactating animals in response to some perturbation (e.g., infection). Therefore, high somatic cell count in an individual's milk can be a reflection of that individual's physiological and health status. Although technologies exist to routinely measure or approximate somatic cells or their proxies in milk (Suhren and Walte, 1998; Fragkou et al., 2014), the gold standard measurement of somatic cell count in milking animals is undertaken, at best, every 4 weeks. Milk collected from all females on a farm during the milking process is stored locally in bulk tanks; the milk in these tanks is usually collected every 1–3 days. Leveraging the DNA information contained within the somatic cells, the count of the alleles for a given SNP in the somatic cells of a representative bulk milk tank sample is reflective of the contribution of the DNA of each individual to that bulk milk sample. By comparing the genotype of the bulk milk sample to the stored genotypes of the individuals in the herd that contributed to that milk sample, a prediction of the somatic cell count of each individual can be generated (Blard et al., 2012).

Interest in animal-side genotype-by-sequencing approaches for rapid diagnostics is coming to the fore (Lamb et al., 2020). Such an approach has already demonstrated potential for the diagnosis of African swine fever (O'Donnell et al., 2020). Available technology has now been demonstrated to be able to identify viruses responsible for bovine respiratory disease in cattle with sufficient sequencing reads in the first hour to be able to correctly identify BoHV1, BRSV or BPI-3 present in their lung cultures. Being better able to target antimicrobials where necessary depending on the pathogen will not only reduce medicinal cost but also reduce pressures on microbial resistance due to improved stewardship of antimicrobial use. This could be complemented not only through pathogen identification but also through the screening for antimicrobial-resistant genes in the pathogen.

Development of genotype panels

Several studies have documented processes for developing lower-density genotype panels (Boichard et al., 2012; Judge et al., 2016; Corbin et al., 2014). Creating a bespoke genotyping panel has the advantage that the content is dictated by the developer. This can mean that the variants included are most informative for the population of interest, although this assumes that all variants were appropriately chosen. Proprietary variants can be included in the bespoke panel thus enabling the retention of trade secret variants while also providing an opportunity to include putatively interesting (research) variants for downstream interrogation. Flexibility also exists in the choice of not only the vendor of the genotyping platform (e.g., ThermoScientific, Illumina) but also the service provider; such flexibility creates opportunities for realising the benefits of competitive financial bids from providers although may be limited by the availability of the necessary technology in a region. Irrespective of species, high degrees of concordance exist in the called genotypes between SNP arrays (Berry et al., 2016; Wijesena et al., 2019). Moreover, there is evidence that high concordance exists between array genotypes and those called from imputed low-pass ($\sim$0.5$\times$) whole-genome sequencing (Sanglard et al., 2022a) although it is reasonable to expect that results could differ as the relationship between the reference set of animals and the target animals changes. It should be noted, however, that the panel here refers to the panel of SNPs as opposed to the technology (e.g., SNP chips) with the described approaches also being applicable to other technologies like target-enrichment sequencing.

Genotyping cost is dictated by, amongst others, the volume of chips or reagents ordered – all else being equal, the greater the volume, the lower the unit price of the genotyping platform (and service). Hence, the cost of panel construction and production could

be high unless large volumes are ordered; it could actually make more economic sense to use a commercially available (higher density) panel even if not fully optimised for the population where it will be used. Many of the commercially available genotype panels are very much up-to-date with the state of the art in informative genomic markers – keeping up-to-date requires considerable resources and, in some instances, trial and error where the primers flanking the genotyping site are not straight forward. Developers of bespoke panels, therefore, need to be up-to-date on all informative genomic markers some of which may not even yet be in the public domain. It is also important that the developed bespoke panels are compatible with other available panels (i.e., suitable overlap in content) to enable imputation so as to make use of animals previously genotyped and, as it is now relatively common in cattle and sheep at least, to share (subsets of) genotypes.

The number and choice of genomic markers to include on a panel is a function of the proposed use. If limited to using SNP array genotyping panels, there is generally only a small difference in price per sample for panels with less than 50 000 SNPs relative to the commonly used circa. 50 000 SNP panels; this is especially true when the full cost of genotyping (i.e., sample procurement, DNA extraction, genotyping service and downstream bioinformatic analyses) is considered. Cost reductions could be possible for fewer SNPs (<circa 3 000 to 5 000 SNPs) if using a different genotyping strategy or platform such as targeted genotype-by-sequencing. With circa 50 000 SNPs, then specifically choosing SNPs for traceability and parentage discovery is unlikely to be necessary since sufficient information will be contained within the several thousand SNPs. Nonetheless, the International Society for Animal Genetics parentage SNPs and other SNPs used in other populations for parentage testing should be included in any bespoke panel to facilitate (partial) genotype exchange (even if the exchange is not in the initial plans). Using commercially available panels with circa 50 000 SNPs, specific SNPs for determining breed composition are also unlikely to be needed unless there is a specific interest in rare breeds not represented in the original population to select the informative SNPs. In such situations, individual SNP statistics like the fixation index could be used to identify informative SNPs where ideally the chosen SNPs are fixed for one allele in the breed of interest and fixed (or low allele frequency) for the opposing allele in all other breeds (i.e., fixation index of 1).

Genomic variants may also be chosen for a specific population to aid imputation to higher density, including sequence, in that breed. This is best achieved by selecting informative SNPs from high-density genotypes/sequences in genetically unrelated and prominent animals for the breed in question; having such higher-density genotypes also enables the testing of imputation accuracy from the chosen lower-density panel to a higher density as well as identifying SNPs with a high call rate. Imputation across the genome tends to be best when SNPs are evenly spaced across the genome with a relatively high minor allele frequency in the population within which they will be used (Judge et al., 2016; Lashmar et al., 2021; Gualdrón Duarte et al., 2013); additional SNPs at the chromosome ends are also required to aid imputation (Boichard et al., 2012). Genomic regions where imputation is poor have been demonstrated (Lashmar et al., 2021; Gualdrón Duarte et al., 2013; Ventura et al., 2016); these regions could be co-located with recombination hot spots or could simply be due to poor annotation in these regions of the genome. Imputation from SNPs to microsatellites (McClure et al., 2013) may also be important for parentage testing where historical animals have been genotyped on microsatellites. Minimising the linkage disequilibrium among the chosen SNPs on the panel will minimise redundancy although, for imputation purposes, the SNPs not chosen should ideally be in strong linkage disequilibrium with co-located SNPs that did make the eventual panel.

In the end, the decision on whether or not to develop a bespoke genotyping panel will be a function of cost versus information content (i.e., return). While the ability to add proprietary information can be important, the information content of currently available genotype panels for a population can be approximated by simply genotyping several prominent individuals (possibly 200) on already available panels and quantifying the number of usable (i.e., called) SNPs that are segregating with a minor allele frequency of >0.05 (i.e., are informative). It should also be noted that while many genotype panels assay circa 50 000 SNPs, the actual number of SNPs postediting is circa 40 000 (Berry and Kearney, 2011; Wiggans et al., 2019; Duchemin et al., 2012; Howard et al., 2018). In fact, circa 40 000 SNPs seem sufficient for within-breed genomic predictions (VanRaden et al., 2011a; 2011b).

There is, nonetheless, a growing interest in genotype-by-sequencing or low-pass sequencing approaches (Elshire et al., 2011; Lamb, 2023). Low-pass, or skim sequencing is whole-genome sequencing at low depth (i.e., 0.5×) and, when coupled with imputation, can enable much denser genomic data at a fraction of the cost of deeper sequence. Generating lower-density data coupled with imputation is not new, and is akin to genotyping with low-density assays and imputing to higher (e.g., 50 000 SNP) content. Although SNP arrays that are of sufficient density are in use at a price point that has encouraged wide-spread genotyping throughout many livestock species, the appeal of low-pass sequencing solutions are two-fold: more variants including potentially causative mutations, and the ability to change the set of variants actually fitted in genetic evaluations without the need to redesign a custom assay making the process of including new variants more dynamic. However, there are also two primary shortcomings of such an approach: the need for a representative reference population and the ability to determine genotype calls for variants of importance with greater certainty than imputation might provide. A representative reference set of animals sequenced at greater depth (e.g., 10×) would ideally enable high-accuracy imputation for all animals in the target population, and thus the reference set should initially contain contemporary high-use sires and representative animals from any less connected subpopulations to ensure the reference can capture the diversity of haplotypes in the current population. The reference over time would need to be dynamic. If imputation-mediated sequencing is not desired, then either sequencing depths will need to increase, increasing the cost of data generation, or a targeted capture or genotype-by-sequencing (De Donato et al. 2013) approach would need to be implemented to generate genotypes. Important variants might include diagnostic markers, such as those for coat colour, horned/polled, or those for genetic conditions. Such variants require targeted deep sequencing approaches to ensure genotype calls are made without error. Initial proof of concept work has illustrated that such an approach can yield accurate genotype calls (Snelling et al., 2020; Lamb, 2023) and some genotyping service providers seem inclined to move in the direction of providing genotypes through low-pass sequencing-based products. The end goal of using imputed variants from sequence is to enable more accurate prediction of genetic merit (Veerkamp et al., 2016; Warburton et al., 2020). Moreover, should the desire be to generate genomic evaluations across breeds or diverse strains, then higher-density genotypes may be needed (de Roos et al., 2008; Erbe et al., 2012). SNPs on SNP chips do suffer from ascertainment bias in that those chosen for inclusion on the panel are from a sample population and are usually chosen to be segregating in the breeds represented. Sequencing approaches, either genotype-by-sequencing approaches or whole-genome sequencing, however, can overcome ascertainment bias. It is also possible that such an approach could yield novel indicators, or bio-markers, of traits related to efficiency

by leveraging mitochondrial sequence that is generated through (low-pass) sequencing approaches (Sanglard et al., 2022b).

The future

Genomic technology, and its application, continue to evolve as does omics in general including epigenomics (Triantaphyllopoulos et al., 2016). The cost, particularly the cost per genetic variant genotyped, will continue to decline. The density of genomic information through the greater use of whole-genome sequencing will necessitate fundamental changes to the statistical models that are in use currently to exploit the explosion in data quantity. The use of genomics will likely expand more into commercial sectors of livestock production to enable more informed management decisions including more individualised use of pharmaceuticals, economic valuation of animals for sale, and to connect commercial-level phenotypes to seed stock/nucleus breeding programmes in the absence of known kinship. Such an expansion of genomic data into commercial sectors will require coordination among sectors to avoid unnecessary duplication of costs and mechanisms to share in the cost/benefit. The potential to not only identify causal genomic variants but also to rapidly propagate them in multiple populations through the combination of technologies such as whole-genome sequencing and gene-editing (Jenko et al., 2015) could enable more efficient selection for disease susceptibility and fitness and reduce any undesirable effects associated with more traditional introgression approaches. Traditionally genetic prediction has focused on evaluating animals, but as deeper knowledge of the genome becomes available and the potential to strategically introgress alleles becomes more prevalent, the notion of an 'animal' will need to change and evaluations will be of sets of alleles that could be in common with multiple animals (surrogates). Developments in genotyping technologies including affordability may also contribute to a proliferation of the sample types being routinely genotyped including the metagenome of different locations like the rumen, vagina and udder teats (Ross and Hayes, 2022) as well as different samples for epigenomic analyses.

The routine generation of millions of genotypes instead of tens or hundreds of thousands presents practical hardware and software computational challenges (Jiang et al., 2021). Efficient data storage and management along with efficient data retrieval mechanisms, while also addressing robust data security practices, will be a significant challenge. Integrating data from different genomic sources (including vendors and genotyping/sequencing platforms) with other data including those generated from other omics disciplines will grow in importance as will be the development of scalable, sophisticated computational strategies to integrate these diverse, hierarchical and often unstructured datasets and extract meaningful insights for understanding the complex interactions between genomics and various phenotypes. No doubt, much of the advancements in the near future in genomic analyses will be achieved through integrating domain expertise across a range of disciplines (e.g., biology, data science, breeding).

Conclusions

Genotyping, both the number of genomic markers simultaneously genotyped and the cost per genotype, has changed dramatically in the past 15 years. This was achieved by the commercial availability of microarray-type genotyping platforms coupled with the ever-accelerating growth in demand for such panels thereby reducing the costs. While genotype data from microarray technology is relatively simple to collate and analyse, alternative technologies like genotype-by-sequencing or low-pass sequencing may come to the fore in future, adding new challenges. Parentage determination and screening for major genes or congenital defects were the justification for genotyping animals in the past and such genotyping was confined to valuable genetically elite animal at the top of breeding pyramid. This has now been replaced by the desire for genomic evaluations of candidate sires and dams. The future is likely to involve the genotyping of (almost) all animals at birth and using the wealth of information gleaned from the genotype to deliver value-creating management and breeding decisions.

Ethics approval

Not applicable.

Data and model availability statement

Not applicable. Data were not deposited in an official repository. No new datasets were created.

Declaration of Generative AI and AI-assisted technologies in the writing process

The authors did not use any artificial intelligence-assisted technologies in the writing process.

Author ORCIDs

Berry, D.P.: https://orcid.org/0000-0003-4349-1447.
Spangler M.L.: https://orcid.org/0000-0001-5184-501X.

Author contributions

Berry, D.P.: Conceptualisation, Writing- Reviewing and Editing.
Spangler, M.L.: Conceptualisation, Writing- Reviewing and Editing.

Declaration of interest

None.

Acknowledgements

None.

Financial support statement

None.

References

Abraham, E.E., Maranian, M.J., Spiteri, I., Russell, R., Ingle, S., Luccarini, C., Earl, H.M., Pharoah, P.P.D., Dunning, A.M., Caldas, C., 2012. Saliva samples are a viable alternative to blood samples as a source of DNA for high throughput genotyping. BMC Medical Genomics 5, 19.
Aguilar, I., Misztal, I., Johnson, D.L., Legarra, A., Tsuruta, S., Lawlor, T.J., 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. Journal of Dairy Science 93, 743–752.
Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Research 19, 1655–1664.
Anderson, C.A., Pettersson, F.H., Clarke, G.M., Cardon, L.R., Morris, A.P., Zondervan, K. T., 2010. Data quality control in genetic case-control association studies. Nature Protocology 5, 1564–1573. https://doi.org/10.1038/nprot.2010.116.
Bengtsson, C., Stålhammar, H., Thomasen, J.R., Eriksson, S., Fikse, W.F., Strandberg, E., 2022. Mating allocations in Nordic Red Dairy Cattle using genomic information. Journal of Dairy Science 105, 1281–1297.
Bérodier, M., Berg, P., Meuwissen, T.H.E., Boichard, D., Brochard, M., Ducrocq, V., 2021. Improved dairy cattle mating plans at herd level using genomic information. Animal 15, 100016.

Berry, D.P., 2021. Beef-on-dairy—The generation of crossbred beef × dairy cattle. Journal of Dairy Science 104, 3789–3819.
Berry, D.P., Dunne, F.L., Evans, R.D., McDermott, K., O'Brien, A.C., 2021. Concordance rate in cattle and sheep between genotypes differing in Illumina GenCall quality score. Animal Genetics 52, 208–213.
Berry, D.P., Herman, E.K., Carthy, T.R., Jennings, R., Bandi-Kenari, N., O'Connor, R.E., Mee, J.F., O'Donovan, J., Mathews, D., Stothard, P., 2023. Characterisation of eight cattle with Swyer syndrome by whole-genome sequencing. Animal Genetics 54, 93–103.
Berry, D.P., Kearney, J.F., 2011. Imputation of genotypes from low-to high-density genotyping platforms and implications for genomic selection. Animal 5, 1162–1169.
Berry, D.P., McParland, S., Kearney, J.F., Sargolzaei, M., Mullen, M.P., 2014. Imputation of ungenotyped parental genotypes in dairy and beef cattle from progeny genotypes. Animal 8, 895–903.
Berry, D.P., O'Brien, A., Wall, E., McDermott, K., Randles, S., Flynn, P., Park, S., Grose, J., Weld, R., McHugh, N., 2016. Inter-and intra-reproducibility of genotypes from sheep technical replicates on Illumina and Affymetrix platforms. Genetics, Selection, Evolution 48, 86.
Berry, D.P., Wolfe, A., O'Donovan, J., Byrne, N., Sayers, R.G., Dodds, K.G., McEwan, J.C., O'Connor, R.E., McClure, M., Purfield, D.C., 2017. Characterization of an X-chromosomal non-mosaic monosomy (59, X0) dairy heifer detected using routinely available single nucleotide polymorphism genotype data. Journal of Animal Science 95, 1042–1049.
Berry, D.P., McHugh, N., Randles, S., Wall, E., McDermott, K., Sargolzaei, M., O'Brien, A.C., 2018a. Imputation of non-genotyped sheep from the genotypes of their mates and resulting progeny. Animal 12, 191–198.
Berry, D.P., O'Brien, A., O'Donovan, J., McHugh, N., Wall, E., Randles, S., McDermott, K., O'Connor, R.E., Patil, M.A., Ho, J., Kennedy, A., Byrne, N., Purfield, D.C., 2018b. Aneuploidy in dizygotic twin sheep detected using genome-wide single nucleotide polymorphism data from two commonly used commercial vendors. Animal 12, 2462–2469.
Berry, D.P., McHugh, N., Wall, E., McDermott, K., O'Brien, A.C., 2019. Low-density genotype panel for both parentage verification and discovery in a multi-breed sheep population. Irish Journal of Agriculture and Food Research 58, 1–12.
Bijma, P., Wientjes, Y.C.J., Calus, M.P.L., 2020. Breeding top genotypes and accelerating response to recurrent selection by selecting parents with greater gametic variance. Genetics 214, 91–107.
Blard, G., Zhang, Z., Coppieters, W., Georges, M., 2012. Identifying cows with subclinical mastitis by bulk single nucleotide polymorphism genotyping of tank milk. Journal of Dairy Science 95, 4109–4113.
Boichard, D., Chung, H., Dassonneville, R., David, X., Eggen, A., Fritz, S., Gietzen, K.J., Hayes, B., Lawley, C.T., Sonstegard, T.S., Van Tassell, C.P., VanRaden, P.M., Viaud-Martinez, K.A., Wiggans, G.R., 2012. Design of a bovine low-density SNP array optimized for imputation. PLoS One1 7, e34130.
Bolormaa, S., Ruvinsky, A., Walkden-Brown, S., van der Werf, J., 2008. DNA-based parentage verification in two Australian goat herds. Small Ruminant Research 80, 95–100.
Bouwman, A.C., Mullaart, E., 2023. Screening of in vitro-produced cattle embryos to assess incidence and characteristics of unbalanced chromosomal aberrations. Journal of Dairy Science Communications 4, 101–105.
Calus, M.P.L., Meuwissen, T.H.E., de Roos, A.P.W., Veerkamp, R.F., 2008. Accuracy of genomic selection using different methods to define haplotypes. Genetics 178, 553–561.
Calus, M.P.L., Mulder, H.A., Bastiaansen, J.W.M., 2011. Identification of Mendelian inconsistencies between SNP and pedigree information of sibs. Genetics Selection Evolution 43, 34.
Casas, E., Kehrli, M.E., 2016. A review of selected genes with known effects on performance and health of cattle. Frontiers in Veterinary Science 15, 1–11.
Chen, L., Schenkel, F., Vinsky, M., Crews Jr., D.H., Li, C., 2013. Accuracy of predicting genomic breeding values for residual feed intake in Angus and Charolais beef cattle. Journal of Animal Science 91, 4669–4678.
Christensen, O.F., Lund, M.S., 2010. Genomic prediction when some animals are not genotyped. Genetics Selection Evolution. 42, 2.
Corbin, L.J., Kranis, A., Swinburne, J.E., Vaudin, M., Bishop, S.C., Woolliams, J.A., 2014. The utility of low-density genotyping for imputation in the Thoroughbred horse. Genetics, Selection, Evolution 46, 9. https://doi.org/10.1186/1297-9686-46-9.
Croué, I., Michenet, A., Leclerc, H., Ducrocq, V., 2019. Genomic analysis of claw lesions in Holstein cows: Opportunities for genomic selection, quantitative trait locus detection, and gene identification. Journal of Dairy Science. 102, 6306–6318.
De Donato, M., Peters, S., Mitchell, S., Hussain, T., Imumorin, I., 2013. Genotype-by-sequencing(GBS): A novel, efficient and cost-effective genotyping method for cattle using next-generation sequencing. PLoS One 8, e62137.
de Roos, A.P.W., Hayes, B.J., Spelman, R.J., Goddard, M.E., 2008. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus Cattle. Genetics 179, 1503–1512.
Dekkers, J.C.M., 2004. Commercial application of marker- and gene assisted selection in livestock – Strategies and lessons. Journal of Animal Science 82, E313–E328.
Dodds, K.G., Tate, M.L., Sise, J.A., 2005. Genetic evaluation using parentage information from genetic markers. Journal of Animal Science 83, 2271–2279.
Duchemin, S.I., Colombani, C., Legarra, A., Baloche, G., Larroque, H., Astruc, J.-M., Barillet, F., Robert-Granié, C., Manfredi, E., 2012. Genomic selection in the French Lacaune dairy sheep breed. Journal of Dairy Science 95, 2723–2733.
Edson, J., Brown, J., Miller, W.L., Walter, W.D., 2021. Comparison of sample types from white-tailed deer (*Odocoileus virginianus*) for DNA extraction and analyses. Scientific Reports 11, 10003.
Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6, e19379.
Engel, E., 1980. A new genetic concept: Uniparental disomy and its potential effect, isodisomy. American Journal of Medical Genetics 6, 137–143.
Erbe, M., Hayes, B.J., Matukumalli, L.K., Goswami, S., Bowman, P.J., Reich, C.M., Mason, B.A., Goddard, M.E., 2012. Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. Journal of Dairy Science 95, 4114–4129.
Fernando, R.L., Grossman, M., 1989. Marker assisted selection using best linear unbiased prediction. Genetics, Selection, Evolution 21, 467–477.
Fernando, R.L., Dekkers, J.C.M., Garrick, D.J., 2014. A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. Genetics Selection, Evolution 46, 59.
Fragkou, A., Boscos, C.M., Fthenakis, G.C., 2014. Diagnosis of clinical or subclinical mastitis in ewes. Small Ruminant Research 118, 86–92.
Groenen, M.A., Archibald, A.L., Uenishi, H., Tuggle, C.K., Takeuchi, Y., Rothschild, M. F., Rogel-Gaillard, C., Park, C., Milan, D., Megens, H.-J., Li, S., Larkin, D.M., Kim, H., Frantz, L.A.F., Caccamo, M., Ahn, H., Aken, B.L., Anselmo, A., Anthon, C., Auvil, L., Badaoui, B., Beattie, C.W., Bendixen, C., Berman, D., Blecha, F., Blomberg, J., Bolund, L., Bosse, M., Botti, S., Bujie, Z., Bystrom, M., Capitanu, B., Carvalho-Silva, D., Chardon, P., Chen, C., Cheng, R., Choi, S.-H., Chow, W., Clark, R.C., Clee, C., Crooijmans, R.P.M.A., Dawson, H.D., Dehais, P., De Sapio, F., Dibbits, B., Drou, N., Du, Z.-Q., Eversole, K., Fadista, J., Fairley, S., Faraut, T., Faulkner, G.J., Fowler, K.E., Fredholm, M., Fritz, E., Gilbert, J.G.R., Giuffra, E., Gorodkin, J., Griffin, D.K., Harrow, J.L., Hayward, A., Howe, K., Hu, Z.-L., Humphray, S.J., Hunt, H., Jeon, J.T., Jern, P., Jones, M., Jurka, J., Kanamori, H., Kapetanovic, R., Kim, J., Kim, J.-H., Kim, K.-W., Kim, T.-H., Larson, G., Lee, K., Lee, K.-T., Leggett, R., Lewin, H.A., Li, Y., Liu, W., Loveland, J.E., Lu, Y., Lunney, J.K., Ma, J., Madsen, O., Mann, K., Matthews, L., McLaren, S., Morozumi, T., Murtaugh, M.P., Narayan, J., Nguyen, D.T., Ni, P., Oh, S.-J., Onteru, S., Panitz, F., Park, E.-W., Park, H.-S., Pascal, G., Paudel, Y., Perez-Enciso, M., Ramirez-Gonzalez, R., Reecy, J.M., Rodriguez-Zas, S., Rohrer, G.A., Rund, L., Sang, Y., Schachtschneider, K., Schraiber, J.G., Schwartz, J., Scobie, L., Scott, C., Searle, S., Servin, B., Southey, B.R., Sperber, G., Stadler, P., Sweedler, J.V., Tafer, H., Thomsen, B., Wali, R., Wang, J., Wang, J., White, S., Xu, X., Yerle, M., Zhang, G., Zhang, J., Zhang, J., Zhao, S., Rogers, J., Churcher, C., Schook, L.B., 2012. Analyses of pig genomes provide insight into porcine demography and evolution. Nature 491, 393–398.
Gualdrón Duarte, J.L., Bates, R.O., Ernst, C.W., Raney, N.E., Cantet, R.J.C., Steibe, J.P., 2013. Genotype imputation accuracy in a F2 pig population using high density and low density SNP panels. BMC Genetics 14, 38.
Hayes, B., Visscher, P., Goddard, M., 2009. Increased accuracy of artificial selection by using the realized relationship matrix. Genetics Research 91, 47–60.
Holečková, B., Schwarzbacherová, V., Galdíková, M., Koleničová, S., Halušková, J., Staničová, J., Verebová, V., Jutková, A., 2021. Chromosomal aberrations in cattle. Genes 12, 1330. https://doi.org/10.3390/genes12091330.
Howard, J.T., Rathje, T.A., Bruns, C.E., Wilson-Wells, D.F., Kachman, S.D., Spangler, M. L., 2018. The impact of truncating data on predictive ability for single-step genomic best linear unbiased prediction. Journal of Animal Breeding and Genetics 135, 251–262.
Hulsegge, B., Calus, M.P.L., Windig, J.J., Hoving-Bolink, A.H., Maurice-van Eijndhoven, M.H.T., Hiemstra, S.J., 2013. Selection of SNP from 50K and 777K arrays to predict breed of origin in cattle. Journal of Animal Science 91, 5128–5134.
International Human Genome Sequencing Consortium, 2004. Finishing the euchromatic sequence of the human genome. Nature 431, 931–945.
Jain, M., Olsen, H.E., Paten, B., Akeson, M., 2016. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biology 17, 239.
Jang, S.K., Evans, L., Fialkowski, A., Arnett, L.K., Ashley-Koch, A.E., Barnes, K.C., Becker, D.M., Bis, J.C., Blangero, J., Bleecker, E.R., Boorgula, M.P., Bowden, D.W., Brody, J.A., Cade, B.E., Jenkins, B.W.C., Carson, A.P., Chavan, S., Cupples, L.A., Custer, B., Damrauer, S.M., David, S.P., de Andrade, M., Dinardo, C.L., Fingerlin, T. E., Fornage, M., Freedman, B.I., Garrett, M.E., Gharib, S.A., Glahn, D.C., Haessler, J., Heckbert, S.R., Hokanson, J.E., Hou, L., Hwang, S.-J., Hyman, M.C., Judy, R., Justice, A.E., Kaplan, R.C., Kardia, S.L.R., Kelly, S., Kim, W., Kooperberg, C., Levy, D., Lloyd-Jones, D.M., Loos, R.J.F., Manichaikul, A.W., Gladwin, M.T., Martin, L.W., Nouraie, M., Melander, O., Meyers, D.A., Montgomery, C.G., North, K.E., Oelsner, E.C., Palmer, N.D., Payton, M., Peljto, A.L., Peyser, P.A., Preuss, M., Psaty, B.M., Qiao, D., Rader, D.J., Rafaels, N., Redline, S., Reed, R.M., Reiner, A.P., Rich, S.S., Rotter, J.I., Schwartz, D.A., Shadyab, A.H., Silverman, E.K., Smith, N.L., Smith, J.G., Smith, A. V., Smith, J.A., Tang, W., Taylor, K.D., Telen, M.J., Vasan, R.S., Gordeuk, V.R., Wang, Z., Wiggins, K.L., Yanek, L.R., Yang, I.V., Young, K.A., Young, K.L., Zhang, Y., Liu, D. J., Keller, M.C., Vrieze, S., 2022. Rare genetic variants explain missing heritability in smoking. Nature Human Behaviour 6, 1577–1586.
Jenko, J., Gorjanc, G., Cleveland, M.A., Varshney, R.K., Whitelaw, C.B.A., Woolliams, J. A., Hickey, J.M., 2015. Potential of promotion of alleles by genome editing to improve quantitative traits in livestock breeding programs. Genetics Selection Evolution 47, 55.
Jiang, M., Bu, C., Zeng, J., Du, Z., Xiao, J., 2021. Applications and challenges of high performance computing in genomics. CCF Transactions on High Performance Computing 3, 344–352.

Jiang, Y., Xie, M., Chen, W., Talbot, R., Maddox, J.F., Faraut, T., Wu, C., Muzny, D.M., Li, Y., Zhang, W., Stanton, J.-A., Brauning, R., Barris, W.C., Hourlier, T., Aken, B.L., Searle, S.M.J., Adelson, D.L., Bian, C., Cam, G.R., Chen, Y., Cheng, S., DeSilva, U., Dixen, K., Dong, Y., Fan, G., Franklin, I.A., Fu, S., Fuentes-Utrilla, P., Guan, R., Highland, M.A., Holder, M.E., Huang, G.L., Ingham, A.B., Jhangiani, S.N., Kalra, D., Kovar, C.L., Lee, S.L., Liu, W., Liu, X., Lu, C., Lv, T., Mathew, T., McWilliam, S., Menzies, M., Pan, S., Robelin, D., Servin, B., Townley, D., Wang, W., Wei, B., White, S.N., Yang, X., Ye, C., Yue, Y., Zeng, P., Zhou, Q., Hansen, J.B., Kristiansen, K., Gibbs, R.A., Flicek, P., Warkup, C.C., Jones, H.E., Oddy, V.E., Nicholas, F.W., McEwan, J.C., Kijas, J.W., Wang, J., Worley, K.C., Archibald, A.L., Cockett, N., Xu, X., Wang, W., Dalrymple, B.P., 2014. The sheep genome illuminates biology of the rumen and lipid metabolism. Science 344, 1168–1173.

Judge, M.M., Kearney, J.F., McClure, M.C., Sleator, R.D., Berry, D.P., 2016. Evaluation of developed low-density genotype panels for imputation to higher density in independent dairy and beef cattle populations. Journal of Animal Science 94, 949–962.

Judge, M.M., Kelleher, M.M., Kearney, J.F., Sleator, R.D., Berry, D.P., 2017. Ultra-low-density genotype panels for breed assignment of Angus and Hereford cattle. Animal 11, 938–947.

Junqueira, V.S., Cardoso, F.F., Oliveira, M.M., Sollero, B.P., Silva, F.F., Lopes, P.S., 2017. Use of molecular markers to improve relationship information in the genetic evaluation of beef cattle tick resistance under pedigree-based models. Journal of Animal Breeding and Genetics 133, 1–13.

Kachman, S., 2008. Incorporation of marker scores into national cattle evaluations. In: Proceedings of the 9th Genetic Prediction Workshop, 08 December 2008, Kansas City, MO, pp. 88–91.

Kachman, S.D., Spangler, M.L., Bennett, G.L., Hanford, K.J., Kuehn, L.A., Pollak, E.J., Snelling, W.M., Thallman, R.M., Saatchi, M., Garrick, D.J., 2013. Comparison of within and across breed trained molecular breeding values in seven breeds of beef cattle. Genetics Selection Evolution 45, 30.

Kenny, D., Berry, D.P., Pabiou, T., Rafter, P., 2023. Variation in the proportion of the segregating genome shared between full-sibling cattle and sheep. Genetics Selection Evolution 55, 1–7.

Kinghorn, B.P., 1998. Mate selection by groups. Journal of Dairy Science 81, 55–63.

Kuehn, L.A., Keele, J.W., Bennett, J.W., Mc Daniel, T.G., Smith, T.P.L., Snelling, W.M., Sonstegard, T.S., Thallman, R.M., 2011. Predicting breed composition using frequencies of 50,000 markers from the US Meat Animal Research Centre 2,000 Bull Project. Journal of Animal Science 89, 1742–1750.

Lamb, H.J., 2023. On-farm genotyping to accelerate genetic gain in Australia's northern beef industry PhD thesis. The University of Queensland, Queensland, Australia.

Lamb, H.J., Hayes, B.J., Nguyen, L.T., Ross, E.M., 2020. The future of livestock management: A review of real-time portable sequencing applied to livestock. Genes 11, 1478. https://doi.org/10.3390/genes11121478.

Lashmar, S.F., Berry, D.P., Pierneef, R., Muchadeyi, F.C., Visser, C., 2021. Assessing single-nucleotide polymorphism selection methods for the development of a low-density panel optimized for imputation in South African Drakensberger beef cattle. Journal of Animal Science 99, skab118.

Legarra, A., Aguilar, I., Misztal, I., 2009. A relationship matrix including full pedigree and genomic information. Journal of Dairy Science 92, 4656–4663.

Li, J.H., Mazur, C.A., Berisa, T., Pickrell, J.K., 2021. Low-pass sequencing increases the power of GWAS and decreases measurement error of polygenic risk scores compared to genotyping arrays. Genome Research 31, 529–537.

Li, Y., Willer, C., Sanna, S., Abecasis, G., 2009. Genotype imputation. Annual Review of Genomics and Human Genetics 10, 387–406.

Lopes, M.S., Silva, F.F., Harlizius, B., Duijvesteijn, N., Lopes, P.S., Guimarães, S.E.F., Knol, E.F., 2013. Improved estimation of inbreeding and kinship in pigs using optimized SNP panels. BMC Genetics 14, 92.

MacNeil, M.D., Nkrumah, J.D., Woodward, B.W., Northcutt, S.L., 2010. Genetic evaluation of Angus cattle for carcass marbling using ultrasound and genomic indicators. Journal of Animal Science 88, 517–530.

McClure, M.C., Sonstegard, T.S., Wiggans, G.R., Van Eenennaam, A.L., Weber, K.L., Penedo, C.T., Berry, D.P., Flynn, J., Garcia, J.F., Carmo, A.S., Regitano, L.C.A., Albuquerque, M., Silva, M.V.G.B., Machado, M.A., Coffey, M., Moore, K., Boscher, M.-Y., Genestout, L., Mazza, R., Taylor, J.F., Schnabel, R.D., Simpson, B., Marques, E., McEwan, J.C., Cromie, A.R., Coutinho, L.L., Kuehn, L.A., Keele, J.W., Piper, E.K., Cook, J., Williams, R., Bovine HapMap Consortium, Van Tassell, C.P., 2013. Imputation of microsatellite alleles from dense SNP genotypes for parentage verification across multiple Bos taurus and Bos indicus breeds. Frontiers in Genetics 4, 176.

McClure, M.C., McCarthy, J., Flynn, P., McClure, J.C., Dair, E., O'Connell, D.K., Kearney, J.F., 2018. SNP data quality control in a national beef and dairy cattle system and highly accurate SNP based parentage verification and identification. Frontiers in Genetics 9, 84. https://doi.org/10.3389/fgene.2018.00084.

McParland, S., Kearney, J.F., Rath, M., Berry, D.P., 2007. Inbreeding effects on milk production, calving performance, fertility, and conformation in Irish Holstein-Friesians. Journal of Dairy Science 90, 4411–4419.

McParland, S., Kearney, J.F., MacHugh, D.E., Berry, D.P., 2008. Inbreeding effects on postweaning production traits, conformation, and calving performance in Irish beef cattle. Journal of Dairy Science 86, 3338–3347.

Medugorac, I., Seichter, D., Graf, A., Russ, I., Blum, H., Göpel, K.H., Rothammer, S., Förster, M., Krebs, S., 2012. Bovine polledness – an autosomal dominant trait with allelic heterogeneity. PLoS One 7, e39477.

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819–1829.

More, M., Gutierrez, G., Rothschild, M., Bertolini, F., Ponce, A., de Leon, F., 2019. Evaluation of SNP genotyping in Alpacas using the bovine HD genotyping beadchip. Frontiers in Genetics 10, 361.

O'Brien, A.C., Purfield, D.C., Judge, M.M., Long, C., Fair, S., Berry, D.P., 2020. Population structure and breed composition prediction in a multi-breed sheep population using genome-wide single nucleotide polymorphism genotypes. Animal 14, 464–474.

O'Donnell, V.K., Grau, F.R., Mayr, G.A., Samayoa, T.L.S., Dodd, K.A., Barrette, R.W., 2020. Rapid sequence-based characterization of African swine fever virus by use of the Oxford Nanopore MinION sequence sensing device and a companion analysis software tool. Journal of Clinical Microbiology 58, e01104–e01119.

Oget, C., Teissier, M., Astruc, J.M., Tosser-Klopp, G., Rupp, R., 2019. Alternative methods improve the accuracy of genomic prediction using information from a causal point mutation in a dairy sheep model. BMC Genomics 20, 719.

Panetto, J.C.C., Machado, M.A., Silva, M.V.G.B., Barbosa, R.S., Santos, G.G., Leite, R.M. H., Peixoto, M.G.C.D., 2017. Parentage assignment using SNP markers, inbreeding and population size for the Brazilian Red Sindhi cattle. Livestock Science 204, 33–38.

Pilon, B., Hinterneder, K., Hay, E.A., Fragomeni, B., 2021. Inbreeding calculated with runs of homozygosity suggests chromosome-specific inbreeding depression regions in Line 1 Hereford. Animals 11, 3105.

Prakash, S., Guo, D., Maslen, C.L., Silberbach, M., Milewicz, D., Bondy, C.A., 2014. Single-nucleotide polymorphism array genotyping is equivalent to metaphase cytogenetics for diagnosis of Turner syndrome. Genetics in Medicine 16, 53–59.

Psifidi, A., Dovas, C.I., Bramis, G., Lazou, T., Russel, C.L., Arsenos, G., Banos, G., 2015. Comparison of eleven methods for genomic DNA extraction suitable for large-scale whole-genome genotyping and long-term DNA banking using blood samples. PLoS One 10, e0115960.

Purfield, D.C., Berry, D.P., McParland, S., Bradley, D.G., 2012. Runs of homozygosity and population history in cattle. BMC Genetics 13, 70.

Purfield, D.C., McClure, M., Berry, D.P., 2016. Justification for setting the individual animal genotype call rate threshold at eighty-five percent. Journal of Animal Science 94, 4558–4569.

Rafter, P., Purfield, D.C., Berry, D.P., Parnell, A.C., Gormley, I.C., Kearney, J.F., Coffey, M.P., Carthy, T.R., 2018. Characterization of copy number variants in a large multibreed population of beef and dairy cattle using high-density single nucleotide polymorphism genotype data. Journal of Animal Science 96, 4112–4124.

Řehout, V., Hradecká, E., Čítek, E., 2006. Evaluation of parentage testing in the Czech population of Holstein cattle. Czech Journal of Animal Science 51, 503–509.

Romano, J.E., Raussdepp, T., Mulon, P.Y., Villadóniga, G.B., 2015. Non-mosaic monosomy 59, X in cattle: A case report. Animal Reproductive Science 156, 83–90.

Ross, E.M., Hayes, B.J., 2022. Metagenomic Predictions: A Review 10 years on. Frontiers in Genetics 13, 865765. https://doi.org/10.3389/fgene.2022.865765.

Sanders, K., Bennewitz, J., Kalm, E., 2006. Wrong and missing sire information affects genetic gain in the Angeln dairy cattle population. Journal of Dairy Science 89, 315–321.

Sanglard, L.P., Kuehn, L.A., Snelling, W.M., Spangler, M.L., 2022a. Influence of environmental factors and genetic variation on mitochondrial DNA copy number. Journal of Animal Science 100, skac059.

Sanglard, L.P., Kuehn, L.A., Snelling, W.M., Spangler, M.L., 2022b. Genotype concordance between SNP chip and imputed low-pass whole-genome sequence in beef cattle. Journal Animal Science 100, 7–8.

Santos, D.J.A., Cole, J.B., Lawlor Jr, T.J., VanRaden, P.M., Tonhati, H., Ma, L., 2019. Variance of gametic diversity and its application in selection programs. Journal of Dairy Science 102, 5279–5294.

Schiermiester, L.N., Thallman, R.M., Kuehn, L.A., Kachman, S.D., Spangler, M.L., 2015. Estimation of breed-specific heterosis effects for birth, weaning, and yearling weight in cattle. Journal of Animal Science 93, 46–52.

Schlötterer, C., 2004. The evolution of molecular markers – just a matter of fashion? Nature Reviews Genetics 5, 63–69.

See, G.M., Fix, J.S., Schwab, C.R., Spangler, M.L., 2022. Imputation of non-genotyped F_1 dams to improve genetic gain in swine crossbreeding programs. Journal of Animal Science 100, skac148.

Selvaggi, M., Darioa, C., Peretti, V., Ciotola, F., Carnicella, D., Dario, M., 2010. Inbreeding depression in Leccese sheep. Small Ruminant Research 89, 42–46.

Silió, L., Rodríguez, M.C., Fernández, A., Barragán, C., Benítez, R., Óvilo, C., Fernández, A.I., 2013. Measuring inbreeding and inbreeding depression on pig growth from pedigree or SNP-derived metrics. Journal of Animal Breeding and Genetics 130, 349–360.

Snelling, W.M., Hoff, J.L., Li, J.H., Kuehn, L.A., Keel, B.N., Lindholm-Perry, A.K., Pickrell, J.K., 2020. Assessment of imputation from low-pass sequencing to predict merit of beef steers. Genes 11, 1312.

Strucken, E.M., Al-Mamun, H.A., Esquivelzeta-Rabell, C., Gondro, C., Mwai, O.A., Gibson, J.P., 2017. Genetic tests for estimating dairy breed proportion and parentage assignment in East African crossbred cattle. Genetic Selection Evolution 49, 67.

Suhren, G., Walte, H.-G., 1998. First experiences with automatic flow cytometric determination of total bacterial count in raw milk. Kiel Milchwirtschaftliche Forschungsberichte 50, 249–275.

Triantaphyllopoulos, K.A., Ikonomopoulos, I., Bannister, A.J., 2016. Epigenetics and inheritance of phenotype variation in livestock. Epigenetics & Chromatin 9, 31. https://doi.org/10.1186/s13072-016-0081-5.

Väli, U., Brandström, M., Johansson, M., Ellegren, H., 2008. Insertion-deletion polymorphisms (indels) as genetic markers in natural populations. BMC Genetics 22, 8.
Van Doormaal, B.J., 2016. Parentage analysis services for dairy cattle in Canada. In: Proceedings of the 2016 Interbull Meeting, 24–28 October 2016, Puerto Varas, Chile, pp. 94–107.
Van Eenennaam, A.L., Li, J., Thallman, R.M., Quaas, R.L., Dikeman, M.E., Gill, C.A., Franke, D.E., Thomas, M.G., 2007. Validation of commercial DNA tests for quantitative beef quality traits. Journal of Animal Science 85, 891–900.
VanRaden, P.M., 2008. Efficient methods to compute genomic predictions. Journal of Dairy Science 91, 4414–4423.
VanRaden, P.M., Olson, K.M., Null, D.J., Hutchison, J.L., 2011a. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. Journal of Dairy Science 94, 6153–6161.
VanRaden, P.M., O'Connell, J.R., Wiggans, G.R., Weigel, K., 2011b. Genomic evaluations with many more genotypes. Genetics Selection Evolution 43, 10. https://doi.org/10.1186/1297-9686-43-10.
VanRaden, P.M., Cooper, T.A., Wiggans, G.R., O'Connell, J.R., Bacheller, L.R., 2013. Confirmation and discovery of maternal grandsires and great-grandsires in dairy cattle. Journal of Dairy Science 96, 1874–1879.
Veerkamp, R.F., Bouwman, A.C., Schrooten, C., Calus, M.P.L., 2016. Genomic prediction using preselected DNA variants from a GWAS with whole-genome sequence data in Holstein-Friesian cattle. Genetics Selection Evolution 48, 95.
Ventura, R.V., Miller, S.P., Dodds, K.G., Auvray, B., Lee, M., Bixley, M., Clarke, S.M., McEwan, J.C., 2016. Assessing accuracy of imputation using different SNP panel densities in a multi-breed sheep population. Genetics Selection Evolution 48, 71.
Visscher, P.M., Woolliams, J.A., Smith, D., Williams, J.L., 2002. Estimation of pedigree errors in the UK dairy population using microsatellite markers and the impact on selection. Journal of Dairy Science 85, 2368–2375.
Wang, H., Misztal, I., Legarra, A., 2014. Differences between genomic-based and pedigree-based relationships in a chicken population, as a function of quality control and pedigree links among individuals. Journal of Animal Breeding and Genetics 131, 445–451.
Warburton, C.L., Engle, B.N., Ross, E.M., Costilla, R., Moore, S.S., Corbet, N.J., Allen, J. M., Laing, A.R., Fordyce, G., Lyons, R.E., McGowan, M.R., Burns, B.M., Haye, B.J., 2020. Use of whole-genome sequence data and novel genomic selection strategies to improve selection for age at puberty in tropically-adapted beef heifers. Genetics Selection Evolution 52, 28.
Weir, B.S., Hill, W.G., 2002. Estimating F-statistics. Annual Review of Genetics 36, 721–750.
Weller, J.I., Seroussi, E., Ron, M., 2006. Estimation of the number of genetic markers required for individual animal identification accounting for genotyping errors. Animal Genetics 37, 387–389. https://doi.org/10.1111/j.1365-2052.2006.01455.x.
Wiggans, G.R., VanRaden, P.M., Bacheller, L.R., Tooker, M.E., Hutchison, J.L., Cooper, T.A., Sonstegard, T.S., 2010. Selection and management of DNA markers for use in genomic evaluation. Journal of Dairy Science 93, 2287–2292.
Wiggans, G.R., Sonstegard, T.S., VanRaden, P.M., Matukumalli, L.K., Schnabel, R.D., Taylor, J.F., Schenkel, F.S., Van Tassell, C.P., 2019. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. Journal of Dairy Science 92, 3431–3436.
Wijesena, H.R., Rohrer, G.A., Nonneman, D.J., Keel, B.N., Petersen, J.L., Kachman, S.D., Ciobanu, D.C., 2019. Evaluation of genotype quality parameters for SowPro90, a new genotyping array for swine. Journal of Animal Science 97, 3262–3273.
Zhao, F., McParland, S., Kearney, F., Du, L., Berry, D.P., 2015. Detection of selection signatures in dairy and beef cattle using high-density genomic information. Genetics Selection Evolution 47, 1–12.
Zhou, L., Heringstad, B., Su, G., Guldbrandtsen, B., Meuwissen, T.H.E., Svendsen, M., Grove, H., Nielsen, U.S., Lund, M.S., 2014. Genomic predictions based on a joint reference population for the Nordic Red cattle breeds. Journal of Dairy Science 97, 4485–4496.
Zimin, A.V., Delcher, A.L., Florea, L., Kelley, D.R., Schatz, M.C., Puiu, D., Hanrahan, F., Pertea, G., Van Tassell, C.P., Sonstegard, T.S., Marçais, G., Roberts, M., Subramanian, P., Yorke, J.A., Salzberg, S.L., 2009. A whole-genome assembly of the domestic cow, Bos Taurus. Genome Biology 10, R42.
Zimmermann, M.J., Kuehn, L.A., Spangler, M.L., Thallman, R.M., Snelling, W.M., Lewis, R.M., 2021. Breed and heterotic effects for mature weight in beef cattle. Journal of Animal Science 99, skab209.