

Review

The Future of Livestock Management: A Review of Real-Time Portable Sequencing Applied to Livestock

Harrison J. Lamb ^{*}, Ben J. Hayes, Loan T. Nguyen  and Elizabeth M. Ross

Centre for Animal Science, Queensland Alliance for Agriculture and Food Innovation,
The University of Queensland, St. Lucia, QLD 4067, Australia; b.hayes@uq.edu.au (B.J.H.);
t.nguyen3@uq.edu.au (L.T.N.); e.ross@uq.edu.au (E.M.R.)

* Correspondence: Harrison.lamb@uqconnect.edu.au

Received: 28 August 2020; Accepted: 1 December 2020; Published: 9 December 2020



Abstract: Oxford Nanopore Technologies' MinION has proven to be a valuable tool within human and microbial genetics. Its capacity to produce long reads in real time has opened up unique applications for portable sequencing. Examples include tracking the recent African swine fever outbreak in China and providing a diagnostic tool for disease in the cassava plant in Eastern Africa. Here we review the current applications of Oxford Nanopore sequencing in livestock, then focus on proposed applications in livestock agriculture for rapid diagnostics, base modification detection, reference genome assembly and genomic prediction. In particular, we propose a future application: 'crush-side genotyping' for real-time on-farm genotyping for extensive industries such as northern Australian beef production. An initial in silico experiment to assess the feasibility of crush-side genotyping demonstrated promising results. SNPs were called from simulated Nanopore data, that included the relatively high base call error rate that is characteristic of the data, and calling parameters were varied to understand the feasibility of SNP calling at low coverages in a heterozygous population. With optimised genotype calling parameters, over 85% of the 10,000 simulated SNPs were able to be correctly called with coverages as low as 6×. These results provide preliminary evidence that Oxford Nanopore sequencing has potential to be used for real-time SNP genotyping in extensive livestock operations.

Keywords: crush-side genotyping; Oxford Nanopore sequencing; MinION; long-read sequencing; livestock; genetics; SNP calling; structural variant; genome assembly

1. Introduction

DNA sequencing allows us to examine the underlying genomic information that affects important traits in agriculture. Sequencing technology can broadly be divided into three distinct generations. First-generation technology, such as Sanger sequencing, was characterised by the use of chain termination, second-generation sequencers, such as Illumina, by high-throughput short reads and third-generation sequencers by high-throughput long reads [1,2]. Two technologies have led third-generation sequencing: Oxford Nanopore Technologies (ONT) and Pacific Biosciences (PacBio). A feature of third-generation sequencing technology is that the native DNA is sequenced directly without amplification. This has the advantage of removing the nucleotide biases and alterations in relative abundance of DNA templates that are observed in some short-read sequence data [1,3,4]. Single-molecule sequencing also produces longer reads which can map and characterise complex genomic regions, such as interspersed repeats, with greater specificity. Long reads have greatly improved the de novo assembly of a number of complex genomes. Their ability to span repetitive regions and complex structural variants (SVs) has led to the publication of more contiguous and

accurate reference genomes [5–8]. However, single-molecule sequencing has an associated increase in error rate as a result of not being amplified against a DNA template.

ONT sequencing uses a biological nanopore embedded in an electrically resistant membrane [9]. A motor protein, which “unzips” the dsDNA to allow it to pass through the nanopore, is added to the DNA template during library preparation [10]. A current flows through the nanopore, which is disrupted in a characteristic manner by nucleotides passing through the pore (Figure 1, [10]). ONT base calling algorithms use the characteristic disruption in electrical signal by 5 base pair k-mers passing through the pore to interpret the DNA sequence [10]. ONT released the MinION in 2014 [10] and it remains the smallest sequencer on the market at 105 mm × 23 mm × 33 mm [11] while also being powered by the USB port of a computer [10]. Recently, ONT also released the MinION MK1c which combines the MinION sequencer with a high-performance computer for base calling, a high-capacity solid state drive and a touch screen display [12]. The MinION and MinION MK1c are to date the only portable sequencing platforms on the market, a significant advantage over other third-generation sequencing platforms. Some systematic errors remain a problem with ONT sequencing, in particular, the under-representation of homopolymer repeats, caused by an inability for the base caller to identify the breakpoints in signal between homopolymers [13]. However, ONT have continued to release new flow cell chemistry and base calling algorithms to attempt to address this problem. Initially, single-pass accuracies of only 68.4% (Phred score of 5) were reported with the MinION in 2015 [13]. Using the R9.4.1 flow cell chemistry and Guppy base caller, raw read accuracies of 85–90% have been reported [14]. This is considerably less than the 99.999% accuracy in sanger sequencing [15] and 99.92% in Illumina [16]. The release of the R10 flow cell chemistry has reportedly increased the raw read accuracy. However, this has come at a cost to the overall flow cell yield. The error rate in ONT sequencing makes single-nucleotide variant (SNV) and insertion/deletion (INDEL) detection difficult and for this reason it has been largely avoided [17]. Instead, ONT sequencing’s long reads have given it a natural application in the detection of large variants, such as SVs—defined here as deletions, insertions, translocations, duplications and inversions greater than 50 bp [18]—in place of short-read sequencing technologies [17].

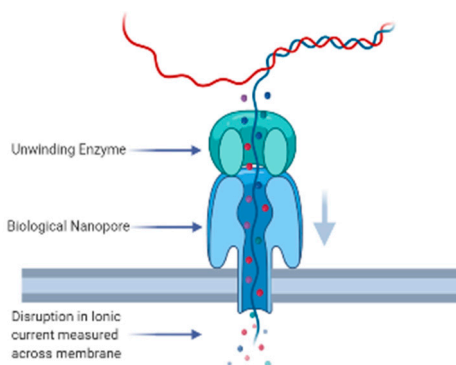


Figure 1. Stylized representation of Nanopore’s sequencing mechanism. Biological nanopores are embedded in a membrane. Double-stranded DNA is enzymatically unwound and the resulting ssDNA passes through the nanopore, disrupting the ionic flow through the nanopore in a characteristic manner dependent on the nucleotide bases within the pore.

ONT have also released a number of other sequencing platforms, which use the same underlying technology. For example, the GridION and PromethION, larger, benchtop cousins to the MinION, provide higher throughput, producing 75–150 Gbp and 2.4–4.3 Tbp, respectively, per run [12,19]. Additionally, the Flongle is a flow cell compatible with both the MinION and GridION which provides a cheaper alternative to traditional MinION flow cells where high yield is not required [12]. ONT has also developed LamPORE, their first diagnostic assay which will be aimed at detecting SARS-CoV-2 [12]. LamPORE is compatible with both the MinION and GridION sequencers and

combines barcoded multitarget amplification using a 15 min barcoded library preparation [20]. ONT estimate that LamPORE is capable of processing 96 samples in just over one hour [12]. To complement their line of sequencing platforms, ONT have a number of software packages for raw read base calling (Guppy), data quality control (EPI2ME and Guppy) and analysis (EPI2ME). An android application, F5N, has also been released to allow sequencing and rudimentary data analysis using a smart phone paired with a MinION [21].

The MinION has already been widely used as a tool for rapid diagnostics within human and microbial genetics [22–25]. Quick, et al. [23] identified, at a species level, a *Salmonella enterica* outbreak in a hospital within 20 min of sequencing on the MinION. However, in another study, the MinION demonstrated that accurate identification of bacteria was possible within 2 h of receiving a DNA sample [22]. Perhaps most impressively the MinION decreased the diagnosis time in a 62-year-old female patient with bacterial sepsis from 6.25 days to just 19 h [26]. The MinION's portability and ability to base call in real time has also made it popular for tracking disease outbreaks [24,27,28], including recent outbreaks of ebola [24], zika [28] and COVID-19 [29,30] as well as more isolated outbreaks of human metapneumovirus [31] and the fungal pathogen *Candida auris* [32].

Here we review a number of current applications of the MinION such as rapid diagnostics and de novo genome assembly and discuss their potential applications to the livestock industry, taking lessons from medical research. We discuss and review the potential for the MinION as a rapid diagnostic tool for diseases relating to pigs, chickens and cattle as well as benefits associated with using the MinION to genotype causative mutations including structural variants for genomic prediction. In particular, we present a future application in “crush-side genotyping”: the rapid, on-farm low-coverage genotyping of cattle for genomic prediction.

2. Current and Future Applications in Livestock

Within the field of agriculture, the uptake of Nanopore sequencing has been relatively slow by comparison to medical research, and slower still within livestock. The applications of the sequencer within livestock agriculture have largely been extensions of disease tracing and rapid diagnostics.

2.1. Rapid Diagnostics of Pathogens in Livestock

African swine fever virus is the causative agent behind the 2019 global outbreak of African Swine Fever (ASF), a highly contagious disease endemic to several sub-Saharan African countries [33]. ASF has up to a 100% mortality rate [34], and modelling of the economic impacts of the disease indicate that pork prices could increase by 17–85% [35] due to the current outbreak which has put an enormous strain on global pork production [36]. Early and rapid diagnosis of ASF is a crucial step in controlling the impact of the disease given the absence of a vaccine [37]. Results from current diagnosis methods may take several days or even weeks due to significant logistical and economic constraints [37]. O'Donnell, et al. [38] demonstrated the potential ability for the MinION to provide rapid diagnosis of this virus with the development of the African swine fever fast analysis sequencing tool (ASF-FAST). ASF-FAST monitors the designated sequencing output directory for .fastq files [38]. New .fastq files are transferred into a “Que” folder where reads are separated by barcode using Porechop [39] and processed using the Burrows–Wheeler alignment tool [40] to perform a reference-guided assembly [38]. ASF-FAST then performs a local BLASTN of the consensus sequence against known ASFV reference strains [38]. An output report with summary statistics in a PDF is then sent by SMS upon request [38]. Sufficient viral sequence was captured for complete resolution of the 192 kb genome [41] within 10 min of sequencing using ASF-FAST [38]. Accurate characterisation and diagnosis of porcine viral enteric disease in piglets has also been demonstrated using the MinION [42]. The ASF-FAST pipeline demonstrates the benefits in the MinION's real-time sequencing ability, and could be expanded to meet other diagnostic needs within livestock agriculture.

Avian cholera caused by the bacterium *Pasteurella multocida* has significant economic impacts in the poultry industry, principally in free range operations [43]. It has had a particularly severe impact on the

Australian poultry industry with a number of outbreaks and subsequent re-infections [43]. Whole killed cell vaccines are currently used to prevent outbreaks. However, these only provide protection against strains with identical or near identical lipopolysaccharide (LPS) structure [43,44]. The MinION could be used to not only provide rapid diagnosis of diseases such as fowl cholera in poultry flocks but also provide real-time characterisation of the LPS structure for accurate targeted vaccination.

Bovine respiratory disease (BRD) is the most common cause of mortality of cattle in North American and Australian feedlots [45]. In North America alone BRD is estimated to cost the agriculture industry over \$1 billion USD [46]. BRD is typically induced by a combination of stress factors as well as exposure to a range of infectious agents [47,48]. Rapid diagnosis of BRD and characterisation of the causative agent is a key mitigation strategy for disease control in feedlots [49]. The MinION is capable of identifying the viruses responsible for BRD using Oxford Nanopore's rapid sequencing kit [50,51]. McCabe, et al. [50] reported that library preparation was complete in approximately 10 min and enough reads were produced in the first hour of sequencing to correctly identify three viruses present in the sample. Moreover, they also concluded the MinION has potential for the rapid diagnosis of viruses in regional veterinary settings [50]. The cassava virus action project has already demonstrated that the MinION can be used for rapid diagnostics in agriculture in the field where laboratory access is limited [52]. Live export is another area in which the MinION's ability to rapidly diagnose a range of diseases could be exploited to improve animal welfare. For example, rapid sequencing diagnostics could be applied to animals before they enter the live export supply chain or a feedlot, to prevent disease outbreaks or identify animals not fit to travel.

In agriculture, the emergence of antibiotic resistant bacterial strains is of significant concern. Strains of multidrug-resistant *Mannheimia haemolytica* have already been identified in cattle with BRD using the MinION [53]. The MinION could reportedly predict resistance to β -lactams, tetracyclines, lincosamides, phenicols, aminoglycosides, sulfonamides and macrolides [53]. An immediately evident application for such abilities is mastitis in dairy cattle, where antibiotic resistant genes within an infection could be diagnosed within hours. Effectively targeting treatments to the specific bacterial infection has the potential to both reduce the treatment cost as well as improve veterinary outcomes by saving needless use of broad-spectrum antibiotics and increasing the efficacy of treatment.

Other real-time analysis tools have also been developed for real-time mapping and coverage statistics. The Read Assignment, Mapping and Phylogenetic Analysis in Real Time (RAMPART) tool [54] runs concurrently with ONT's MinKNOW to provide a real-time graphical overview of genome coverage and reference matching of reads as they are base called and mapped by MinKNOW [54]. RAMPART has been adopted by the ARTIC network project which aims to develop an end-to-end system for tracking viral outbreaks using real-time epidemiological information [55]. The ARTIC network pipeline is a user-friendly bioinformatics pipeline that combines reference alignment, post-processing, variant calling and consensus building using a number of popular Nanopore-compatible tools [55]. Despite being initially developed for use in human epidemiology, reference genomes for strains of any number of economically important livestock pathogens could be used with RAMPART to monitor sequencing depth in real time for rapid diagnosis.

A current limitation of viral detection using ONT is the titre of the viral particle. This is because the likelihood of sequencing viral DNA compared to that of the host genome is directly related to the proportion of viral DNA in the sample. Because viral genomes are many orders of magnitude smaller than mammalian genomes, there needs to be many orders of magnitude more viral genomes in the sample than host genomes [50]. One possible solution to this limitation is the upcoming release of adaptive sequencing by ONT [56]. Adaptive or read-until sequencing describes the selective sequencing of particular genomic regions by reversing the flow of ions through specific nanopores [56,57]. Through this, non-target DNA molecules can be ejected from the nanopore increasing the number of available pores for target sequences. This allows enrichment of regions to a desired coverage without additional library preparation procedures [56,57]. By using graphics processing unit (GPU)-compatible base calling, sequences can be base called in real time and either rejected or allowed

to continue sequencing [56]. Another approach by Kovaka, et al. [57] involves the targeting of k-mers with an estimated similar electrical signal to the desired enrichment regions. This eliminates the need for GPU base calling which could potentially lend itself towards regional applications of the MinION with adaptive sequencing. As a result, adaptive sequencing could be used to target viral diseases with significant economic effects on the livestock industry, including swine flu, foot-and-mouth, brucellosis, tuberculosis and bovine viral diarrhoea virus [58–60]. To date, these protocols have only been applied to medical research. However, they hold great promise for the realisation of in-field diagnostics.

Pipelines such as RAMPART and ASF-FAST demonstrate the proven ability for the MinION to significantly decrease diagnostic times and increase treatment specificity. Applications from human medical research can be applied directly to agricultural diagnostics. Importantly, the MinION not only presents an opportunity to lower disease burden, but also provides a mechanism for the reduction in broad-spectrum antibiotic treatment in agriculture. Currently, the high sequencing error rate as well as the relatively higher cost per sequencing run are two hurdles on the path to the wider adoption of ONT sequencing in agriculture. However, it is hoped that the release of adaptive sequencing will address these issues.

2.2. Reference Genomes

Reference genome assemblies are essential in understanding an organism's genetic diversity. Genome assemblies allow us to characterise genetic variants and also reduce the computation time required to analyse sequence data [61]. Through characterising genetic variants, they have allowed us to better understand genetic diseases and in agriculture, in particular, to leverage heritability for accelerated genetic gain. In livestock, genome assemblies exist for all of the major domesticated species: cattle [62], chicken [63], sheep [64], goats [61] and pigs [65]. A number of sequencing platforms have been used to develop reference genomes, each with their respective advantages and disadvantages (Table 1). Historically, reference assemblies were created using short-read sequencing technology and were less accurate than what can be achieved now with long-read technologies. An estimated 14.4 Mbp of the chicken (*Gallus gallus*) reference assembly was discovered to be misassembled contigs; while in the bovine (*Bos taurus*) reference assembly, 39 Mbp was estimated to be misassembled contigs [66].

ONT sequencing's unparalleled read length has allowed for the assembly of a number of complex genomes such as the protozoan parasite *Trypanosoma cruzi* [7] and the mountain-spinach *Atriplex hortensis* [8]. The ability to sequence complex genomes is in part due to long reads being able to span large repetitive regions which second-generation sequencing technologies typically fail to accurately assemble. The inability for short reads to map complex regions results in uneven sequencing coverage in comparison to long reads (Figure 2). A comparison of two genome assemblies using Nanopore and PacBio sequencing in rice found the Nanopore assembly to be more contiguous, having only 18 contigs [67], whereas the PacBio assembly featured 394 contigs [67]. The Nanopore assembly also reportedly had fewer assembly errors caused by long repetitive regions [67]. The individual base accuracy of the ONT assembly, however, was lower than the PacBio assembly, despite polishing the ONT assembly with $70 \times$ Illumina short reads [67]. Single-nucleotide errors in the ONT assembly formed clusters of which 94% were in complex genomic regions and therefore had Illumina short-read coverage of $<5 \times$ [67]. This highlights the importance of selecting an appropriate sequencing platform to meet the desired level of assembly contiguity and individual base accuracy.

Relatively high sequencing coverage is necessary for accurate assemblies when using Nanopore data as a result of the high sequencing error rate. For small genomes, this is easily achieved. However, for larger genomes, the higher cost per Gb for Nanopore sequencing means cheaper, highly accurate platforms such as Illumina are now used alongside Nanopore data, such as in the case of the snow sheep (*Ovis nivicola*) genome assembly [68]. This decreases the cost of sequencing and increases the individual base accuracy of genome assemblies. Reference genomes for the endangered giant redwood (*Sequoia sempervirens*) and giant sequoia (*Sequoiadendron giganteum*), both trees native to California, were also created using Illumina and Nanopore data [69]. These genomes were 8.2 and 26.5 Gbp,

respectively, demonstrating an ability for the MinION to be used for very large genomes [69]. The giant sequoia's reference genome also features the longest scaffolds assembled in any organism to date with the largest being 985 Mbp and the scaffold N50 being 690.55 Mb [69].

Table 1. Comparison of bench top second-generation sequencers as well as some of PacBio's and ONT's platforms.

Sequencing Platform	Read Length	Data Output	Run Time	Single-Pass Error Rate (%)	Reference
Illumina HiSeq 2000	2 × 150 bp	150–200 Gb	1–6 days	0.16	[70,71]
Illumina NextSeq 550	2 × 150 bp	120 Gb	15–18 h	0.16	[70,71]
PacBio RS II	10–15 kb on average	0.5–1 Gb	0.5–4 h	10	[16,71]
PacBio Sequel	10–15 kb on average	15 Gb	Up to 20 h	10	[72]
PacBio Sequel II	10–15 kb on average	80–100 Gb	Up to 30 h	10	[73]
Oxford Nanopore MinION	10–20 kb on average	20–30 Gb (per flow cell) Up to 1 flow cell/run	Up to 96 h	5–20	[12,74]
Oxford Nanopore GridION	10–20 kb on average	20–30 Gb (per flow cell) Up to 5 flow cells/run	Up to 96 h	5–20	[12,74]
Oxford Nanopore PromethION	10–20 kb on average	100–180 Gb (per flow cell) Up to 48 flow cells/run	Up to 96 h	5–20	[12,74]

ONT's ability to characterise complex genomic regions has the potential to increase the accuracy of reference genome assemblies in livestock. The ability of all three major ONT sequencing platforms: the MinION, PromethION and GridION, to facilitate accurate genome assemblies may help identify targets for future breeding and selection. More accurate assemblies will also facilitate the incorporation of more complex genomic variants into genomic prediction and increase the power of genomic variant association studies [75].

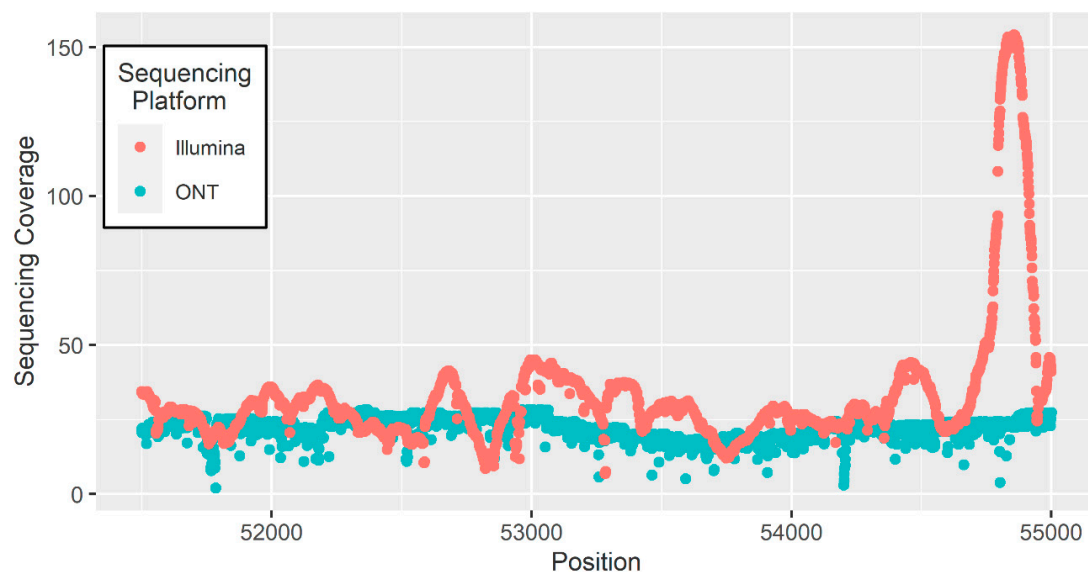


Figure 2. Sequencing coverage across a 3000 bp window for Illumina and ONT sequencing data aligned to the *B. taurus* reference genome [76].

2.3. Structural Variants

Beyond applications in improving the accuracy of reference genomes and rapid diagnostics, the unique mixture of read length and real-time data acquisition leans towards potential uses in calling genomic variants in real time. In particular, the MinION lends itself towards accurate SV characterisation and calling, which has already been demonstrated in humans [77,78]. SVs are estimated to contribute between 0.5 and 1% of genomic differences between individuals while comparatively single-nucleotide variations contribute only 0.1% [18]. This highlights the potential genetic variation that, for the most part, is poorly characterised to date, particularly in species other than model organisms and humans.

The proportion of genetic variation explained by SNP markers is often referred to as h_m^2 , while variance explained by genome-wide-significant SNPs is referred to as h_{GWS}^2 [79]. Heritability studies within families reveal a difference between total additive genetic variance (h^2) and h_{GWS}^2 which is referred to as “missing heritability” [79]. Ultimately these studies reveal $h_{GWS}^2 < h_m^2 < h^2$ [79], which is hypothesised to be the result of genomic variants that are not well tagged by SNPs [79]. Even traits such as fertility in dairy cattle, which could be assumed to be under selective pressure with reasonable confidence, have some missing heritability [80]. A possible source of this missing heritability is larger genomic variants such as SVs [81,82]. SVs range in size from tens to millions of base pairs [83] or even entire chromosomal restructuring [84]. This means that the possible genetic variation explained by SVs may be magnitudes more than that of single-nucleotide variants.

SVs, such as SNPs, can alter the gene itself [85], but may also affect the enhancer or promoter or change its proximity to the gene [86]. The best understood mechanism through which SVs may alter phenotypes is by copy number alteration of a dosage sensitive gene or cluster of genes [86–88], such is the case in Williams–Beuren syndrome [89], Smith–Magenis syndrome [90] and systemic lupus erythematosus [91]. SVs have also been found to alter 3D chromatin structure [92,93]. Large deletions and inversions demonstrate detectable effects on chromatin contacts; however, we do not yet understand the impacts of these alterations on an individual’s phenotype [92]. Significant work has been carried out in human genetics to ascertain relationships between a number of complex diseases and SVs [94–96]. For example, the age of onset of Alzheimer’s disease is poorly modelled by SNPs alone, but with consideration of structural variants [97,98], the age of onset of Alzheimer’s disease can now be accurately predicted in over 90% of the at-risk population [98]. Similarly, it is hypothesised that SVs are accountable for missing heritability in sporadic amyotrophic lateral sclerosis [81]. Generally speaking the mechanistic and molecular links between SVs and phenotypes are still poorly understood, particularly outside of humans and model organisms [18]. More accurate characterisation of SVs may help to clarify these links.

Short-read technologies struggle to identify SVs, in particular copy number variants in comparison to the MinION. In 14 *Drosophila melanogaster* genome assemblies, more than 20,000 SVs were identified by long-read sequencing of which almost 40% were reportedly undetectable using short-read sequencing [99]. The SV allele frequency in the *D. melanogaster* population relative to amino acid polymorphisms also suggested that SVs are more likely to impact phenotype than non-synonymous SNPs [99] making them candidates for being rare alleles with large phenotypic effect [99]. This highlights the potential power of using long-read sequencing platforms such as the MinION to better characterise SVs.

Voichek, et al. [100] compared the correlation between SNPs and phenotypic traits to the correlation between larger genomic variants such as insertions, deletions and rearrangements to phenotypic traits in a genome wide association study. They developed a range of k-mers that represented structural variations and tested the association between the k-mers and desirable phenotypes in *Arabidopsis thaliana*, tomato and maize [100]. K-mers typically had stronger statistical association with phenotype than SNPs and a number of new associations were found exclusively using k-mers [100]. K-mers have the potential to reveal new alleles associated with desirable phenotypic traits in maize [100]. However,

the short-read sequencing which was used produced a high fraction of non-uniquely mapped reads in the maize population [100]; a challenge that would likely be overcome using long Nanopore reads [101].

In livestock, there are only a handful of well-characterised structural variants with known economically significant effects. A 10 bp duplication in the *Pax7* gene is associated with increased body weight, body height and hip width in a number of Chinese cattle breeds [102]. Another phenotype of economic importance in agriculture caused by an SV is the poll phenotype in cattle. The poll phenotype, a desirable trait in both the dairy and beef industries, describes a lack of horns in cattle [103–105]. It is controlled at the poll locus on bovine chromosome one [105,106] by at least three known autosomal dominant poll (P) alleles [107,108]. All three alleles are SVs ranging in size from a 212 bp duplication replacing 10 bp [108] to an 80 kb duplication [109]. The poll phenotype is an example of well-characterised SVs with both economic and animal welfare importance in agriculture [109]. Being a highly desirable Mendelian dominant trait, homozygous poll (PP) animals are of greater value than heterozygous (Pp) cattle. This makes it an ideal candidate for real-time calling as there is no phenotypic way to distinguish poll homozygotes from heterozygotes. Recently, another example of an economically important SV was characterised in sheep using Nanopore sequencing [110]. The SV, a 7.9 kb deletion was found to be responsible for a novel *BCO2/enjSRV* hybrid protein that was found to cause undesirable yellow discolouration of adipose tissue [110]. A genetic test for this trait will allow on-farm management decisions before the trait is detected at time of slaughter. A number of SVs linked to coat colour have also been characterised in cattle [111,112], while of arguably less economic significance the phenotypes caused by these SVs may be of interest to stud breeders. Two complex SVs involving the *KIT* locus (*B. taurus* chromosome 6) have been found to cause colour sidedness in a number of European breeds [111,113]. A quadruplication of a 6 kb region upstream of the *TWIST2* gene in cattle is strongly associated with a second coat-related phenotype: an area of unpigmented skin and coat around the animals' mid-section [112,114]. These studies highlight both the additional predictive power that may be leveraged using larger genomic variants but also the suitability of the MinION sequencer to this application [115]. This application could be harnessed further by using the MinION to incorporate causative mutation (particularly SV) calling into genomic selection.

2.4. DNA/RNA Modification

Another benefit of ONT sequencing is the ability to directly sequence both DNA and RNA molecules to identify molecule modification. Over 17 and 160 post-transcriptional modifications have been characterised in DNA and RNA, respectively [116], with the most common being methylation [117]. These modifications influence DNA-protein interactions and play a crucial role in biological development [118] and ageing [119,120]. Bisulfite sequencing is generally regarded as the gold standard for DNA methylation profiling [121,122]. However, bisulfite DNA reads often have a lower mapping rate, uneven genome coverage and sequence biases [116]. Alternatively, methylated DNA immunoprecipitation can also be used for methylation profiling. However, this method does not provide base resolution information [123]. RNA modifications are generally also captured using immunoprecipitation alongside reverse transcription and sequencing [123]. However, these processes often only interrogate a specific group of modifications at any one time [116]. The MinION's ability to sequence DNA and native RNA means it can deliver sequence information as well as post-transcriptional modification profiling. In the case of DNA sequencing, this means methylation profiling can be performed alongside genotyping, which could be incorporated into pipelines for predicting predisposition to premature ageing [124,125], disease [126] or desirable phenotypic traits [127]. Methylated CpG sites have also been demonstrated to be a significant contributor to Nanopore SNP errors and masking of methylation regions has been shown to decrease the error rate of Nanopore sequencing [25], using this method SNP calling accuracy using the MinION could be increased.

One example where understanding the importance of methylation on a trait for economic benefit might be ossification. Ossification is a measure of physiological age of a beef carcass and extensive

ossification is an undesirable quality in beef cattle. Ossification is associated with increased dark cutting which is the discolouration of meat as a result of high pH post-mortem [128]. Dark cutting results in a shorter shelf life and bland flavour [129] and therefore producers are often penalised \$0.5 AUD/kg of carcass weight [128]. With a strong correlation between ageing and methylation already established in humans [130–132], methylation markers for pre-mature ageing and ossification could be developed and called using the MinION. Evidence also suggests that methylation may play a role in stress-related psychiatric disorders and stress-induced depressive behaviour [133,134]. Therefore, methylation markers for livestock welfare and temperament could also be called using the MinION.

2.5. Genomic Prediction and Crush-Side Genotyping

Genomic prediction, initially proposed by Meuwissen, et al. [135], has been extensively used within a number of livestock industries such as poultry and dairy. Current genomic prediction typically uses SNP arrays to genotype a large number of markers throughout the genome. These markers are used to calculate the genomic relatedness of animals in a reference population with known phenotypes [135]. By comparing the genomic relatedness of animals with unknown phenotypes to those in the reference population, a genomic estimated breeding value (GEBV) can be calculated [135]. This has allowed some intensive agricultural industries to transition from purely phenotypic trait selection to genomic selection (provided large reference populations of genotyped and phenotyped individuals are established [136]).

A number of limitations to using SNP arrays for genomic prediction exist [136]. A key economic limitation of SNP arrays is that the larger or denser a SNP array, the more expensive it becomes [137]. A further limitation to SNP array genotyping is the inability to accurately genotype SNPs whose flanking regions align to multiple regions in the genome. As an example, it is common practice to remove between 30 and 40 k SNPs from the Brassica 60 k SNP array [138]. Furthermore, SNPs on SNP arrays are typically only selected if both alleles are common and therefore cannot be in complete linkage disequilibrium (LD) with a causal variant with one rare allele [79]. Therefore, SNPs discovered using SNP markers in a GWAS are unlikely to account for all genetic variation as rare variants likely will not be tagged by the genotyped SNPs [79]. The MinION's ability to generate whole genome sequence data could mean, with sufficient computing power, both common and rare SNPs as well as structural variants could be called, decreasing the missing heritability in genomic prediction. However, the achievable accuracy of genomic predictions given the high error rate of Nanopore sequencing must be investigated before this application can be adopted.

Outside the biological limitations of SNP genotyping by array, this method is also impractical in some industries. For example, in Australia's northern beef industry, cattle are often only handled once a year, in a mechanical restraint commonly referred to as a 'crush'. Additionally, northern Australian cattle enterprises will typically only make a mail 'run' once or twice a fortnight. This makes proximity to laboratories and turnaround time significant limitations to implementing genomic selection in practice. To address this need, we propose the future application of the MinION sequencer to rapidly on-farm genotype cattle for genomic selection in Australia's northern beef industry, which we term "crush-side genotyping" (Figure 3). Such a method would allow point of management decisions such as culling and young bull selection to be informed by GEBV, in industries where laboratory turnaround time has limited the use of SNP array genotyping. For example, groups of cattle held in a holding yard could be genotyped overnight. This would allow for management decisions to be made the following morning, rather than weeks later, by which time the cattle would be released and not seen until the following year. Using the MinION, whole genome sequence data could be imputed from SNPs, INDELS or copy number variants (CNVs) which are reliably called for each animal. This would require additional computation, but could be highly parallelised. The concept is similar to that proposed by Ros-Freixedes, et al. [139], where very low sequence coverage is used to genotype animals. A method similar to hybrid peeling, a technique for pedigree-based imputation [140,141], could also be incorporated for fast imputation, phasing and calling. Li, et al. [142] demonstrated

that low-coverage ($0.5\text{--}1\times$) short-read, second-generation sequencing, can increase the power of GWAS and decrease the measurement error of polygenic risk scores for coronary artery disease in humans when compared to SNP arrays. Although these results were reported using highly accurate second-generation sequencing, the ability for long reads to map more accurately would likely increase the accuracy of phasing haplotypes at low coverage. The use of the MinION would additionally allow the incorporation of larger genomic variant markers or causative mutations into genomic selection for increased accuracy.

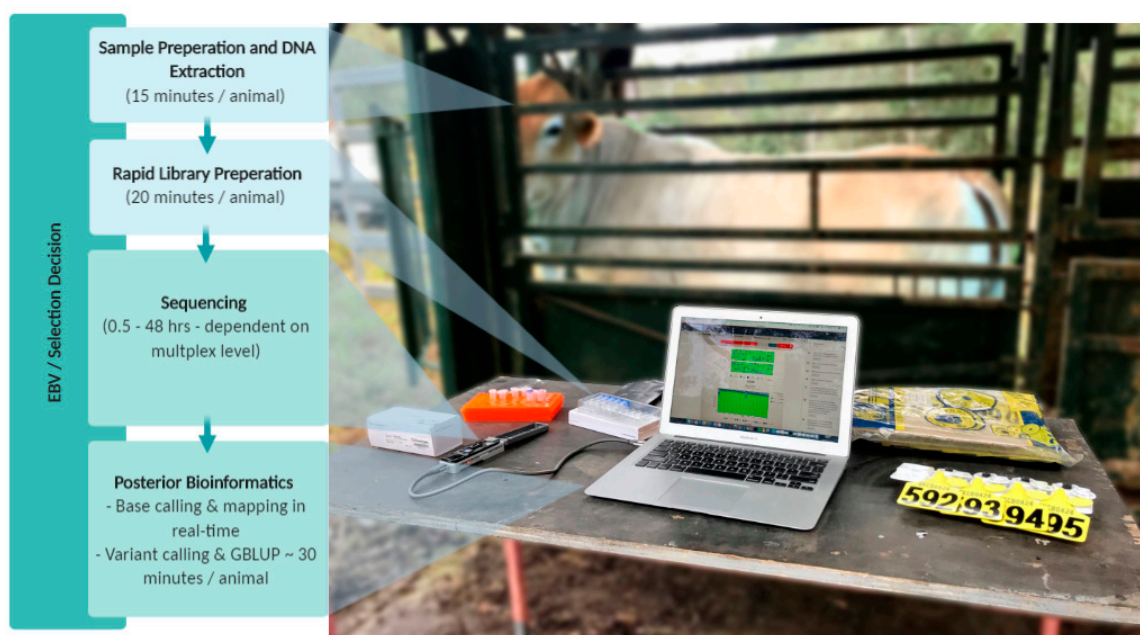


Figure 3. On-site genotyping of a brahman heifer in Queensland using the MinION. With multiplexing a mob of cattle could be genotyped overnight using this pipeline and selection decisions made the following morning using genomic breeding values.

To date, true real-time crush-side genotyping is not feasible as two major limitations exist: cost and the development of a rapid wet-lab pipeline. The pipeline turnaround time is being actively addressed by a number of groups globally. The application of the MinION in disease outbreaks has driven development of rapid protocols for DNA extraction [143], in-field computing [144] and library preparation [145]. For example, Boykin, et al. [146] have demonstrated the PDQeX nucleic acid extractor (MicroGEM, Dunedin, New Zealand), which is capable of extracting DNA from up to 24 different tissue samples in less than 20 min, can rapidly extract DNA for use with the MinION. Additionally, Zou, et al. [143] have developed a cellulose paper-based DNA purification technique that is able to extract and purify DNA in 30 s, another example of available rapid DNA extraction methods. Another development that could increase the speed of the wet-lab turnaround time is ONT's VolTRAX [12], which completely automates the library preparation and flow cell loading steps. Ideally the VolTRAX could be combined with Nanopore's field sequencing library preparation kit. The field sequencing kit is capable of library preparation in 10 min and requires no cold storage and limited laboratory equipment [12], although this protocol currently produces a lower sequencing yield. The VolTRAX and rapid sequencing kit will increase the reproducibility of sequencing runs and decrease the library preparation requirements and turnaround time. The release of the MinION Mk1c has also addressed the hurdle of infield sequence acquisition with its GPU base calling ability and 1 TB solid state drive (SSD) built in with the sequencer. The transition towards GPU-compatible base calling packages has shortened the analysis turnaround time by an order of magnitude [147]. Base calling speeds have increased from 120,000 bases/s using a central processing unit (CPU) in 2017

to over 1,000,000 bases/second using ONT's GPU basecaller Guppy [148]. A lack of internet connection in remote locations presents an obvious hurdle to crush-side genotyping analysis. However, an offline metagenomics analysis tool, MINDS (MinION Detection Software), has already demonstrated that offline analysis of MinION data is currently possible [149]. From 173,000 reads, MINDS correctly identified 19 of 20 species from the MSA-2002 mock metagenomic community in 15 min, without internet connection [149]. Although, this clearly demonstrates the current ability to efficiently analyse MinION data while offline, the significant increase in data throughput necessary in the context of crush-side genotyping means that considerable optimisation would be required.

The second key limitation, cost, will likely be addressed by parallelising the sequencing. Currently, SNP genotyping is available in a variety of densities, from ultra-low ~ 200 SNP parentage arrays to high density 777 k, where density refers to the number of SNPs genotyped and therefore the degree of separation between markers. Traditionally, the beef industry, in particular Australia's Northern beef industry, use commercial 50 k SNP arrays to achieve acceptable genomic prediction accuracies. These 50 k SNP arrays today cost \$35–55 AUD [150], a significant drop from \$400–500 AUD per animal for 10,000 SNPs in 2006 [151]. For crush-side genotyping to be feasible, the cost per sequencing run will not necessarily have to approach the cost of SNP array genotyping. Instead, in industries where SNP array genotyping has not been used due to the limitations discussed, it will only need to cost less than the increase in profitability resulting from genomic selection. Peñagaricano [152] estimated genomic selection has resulted in the doubling of annual genetic gain for economically important traits in dairy cattle since its implementation. Most notably, a 3-fold to 4-fold increase in genetic gain for traits that can only be measured late in life, such as female fertility, have been reported [152]. Currently, a single MinION sequencing run costs ~\$700 USD [12] and can consistently yield over 20 Gb using R9.4 flow cells [12]. This yield has increased dramatically from less than 1 Gb in 2014 using ONT's R6.0 flow cells [153]. Alongside the increase in sequencing yield, in the last six years, the single-pass error rate from Nanopore sequencing has decreased almost 4-fold from 38.2% using the R6.0 flow cells [154] to 9–10% using R9.4 flow cells [155].

We see the possible introduction of adaptive sequencing as a key link in the chain to significantly reduce the cost of sequencing and the required whole genome coverage. A number of markers spread evenly throughout the genome, similar to SNP array genotyping, could theoretically be enriched in real time using adaptive sequencing. This would produce accurate marker genotypes at comparatively ultra-low whole genome coverages. Another method that could be used for rapid selective enrichment is LamPORE's loop-mediated isothermal amplification (LAMP). LamPORE uses a 35 min LAMP to amplify three highly conserved genes in SARS-CoV-2 for rapid diagnosis [12]. The LAMP method has been optimised to be highly parallelised (9000 samples can be barcoded in 24 h) [12], which is very attractive for crush-side genotyping. However, the additional wet-lab steps are currently also a limitation of this technique. The affordable characterisation of rumen microbiome and rapid diagnosis of pathogens in livestock could be more suitable applications of this method. Alternatively, CRISPR-Cas9-targeted enrichment is another method in ONT's toolkit which could be used in place of adaptive sequencing [156]. Currently, CRISPR-Cas9-targeted enrichment provides greater enrichment of desired loci than adaptive sequencing [56,156,157], but introduces additional wet-lab steps [156]. CRISPR-Cas9 could potentially be used where relatively few enriched loci are required, for example in parentage verification or the targeted introduction/eradication of specific genotypes (e.g., deleterious mutations) in a herd.

This method of enrichment shares similarities with genotype by sequencing (GBS), a cost-effective method for low-coverage genotyping [158–160]. GBS uses restriction enzyme cleavage to reduce genome complexity for sequencing on short-read sequencing platforms [158–160]. This method of low-coverage sequencing has the advantage of not requiring a priori: a reference genome in the case of adaptive sequencing and CRISPR-Cas9 enrichment [161]. This means that GBS can simultaneously be used for variant discovery as well as genotyping [158]. Additionally this method is currently more cost effective than the proposed crush-side genotyping. However, a number of limitations to GBS exist.

Most importantly, the second-generation sequencing platforms used for GBS are currently not portable, which is the major drawback crush-side genotyping aims to solve. The real-time nature of adaptive sequencing also means that the chosen enriched loci can be changed during the sequencing run [56]. Enrichment could also be turned off once sufficient sequencing coverage is achieved to capture genome wide variants.

Targeted sequencing enrichment on the MinION would allow for the sequencing of multiple samples per flow cell. This could be combined with ONT's barcoding library kits. Up to 96 samples can currently be multiplexed using ONT's PCR barcoding kit. However, this requires an additional 60 min library preparation and PCR. The rapid barcoding kit (10 min library preparation) currently only uses 12 barcodes, but would likely be more suitable to ensure the pipeline remains rapid and portable [12]. It is likely that the number of samples that can be barcoded with the rapid kit will increase over time. If $6 \times$ coverage is assumed to be the minimum requirement for accurate genotype calling on the majority of loci using the MinION and an average of 50,000 SNP markers are needed for accurate genomic prediction then—given reported sequencing enrichment rates [56,156,157]—we estimate that 1.8 Gb will be necessary per animal. We assume the maximum theoretical yield from a single MinION flow cell is 20 Gb. MinION sequencing of cattle tail hair DNA has yielded up to 45 Gb per flow cell using non-targeted sequencing. Based on these assumptions, multiplexing 11 animals per flow cell is possible. At this multiplexing level, the sequencing cost per animal is approximately \$90 USD. Additionally, minimum coverages for accurate genotyping may decrease with improvements in sequencing accuracy and software advances. Furthermore, accurate genomic prediction is regularly achieved with as little as 4000 SNP markers in cattle, which would decrease the price further. Crush-side genotyping, may not yet be feasible and relies heavily on emerging technology. However, we believe the rapid development of the technology means crush-side genotyping will be feasible in the short to mid-term.

One limitation to the current application of crush-side genotyping is reduced flow cell yields when using enrichment methods. Payne, et al. [56] reported a significant drop in sequencing efficiency while using adaptive sequencing: markedly lower enrichment was achieved towards the end of the sequencing run. This was attributed to the degradation and blocking of pores [56]. Payne, et al. [56] overcame this issue by washing the flow cells with nuclease. Further optimisation of these recently published methods: adaptive sequencing, Cas9 enrichment and LAMP sequencing, will see the feasibility of crush-side genotyping increase shortly.

The ability of the MinION to call SNPs at known loci at low coverage in *Brassica napus* L. has already been assessed [138]. Despite having higher error rates than other sequencing platforms, the MinION achieved an accuracy of 96% in calling SNPs at four million known locations [138]. However, a predominantly homozygous canola population was used in this study, which would result in a higher accuracy than would be expected in outbred populations [138]. However, the ability for the MinION to accurately genotype SNPs in a heterozygous population at low coverage has not been investigated. Therefore, we conducted a preliminary investigation into the feasibility of crush-side genotyping using simulated Nanopore read data to assess the accuracy of SNP calling in heterozygous populations.

3. Materials and Methods

We conducted a simulation study to test the hypothesis that SNP genotypes can be accurately identified in a diploid species using MinION data; and that changing the filtering and calling parameters used to call SNPs will affect the accuracy of the identified SNP.

An artificial diploid of chromosome 28 was generated by randomly inserting 10,000 SNPs into ARS-UCD1.2 chromosome 28 using the package SimuG [162]. A deep learning neural network based Nanopore read simulator, DeepSim [163], was used to simulate the Oxford Nanopore reads. The package is able to simulate Nanopore reads using a pore model which models the expected current signal when given a reference genome input [163]. Both the reference and mutated chromosome 28 were used as the reference genome input for the pore model in order to simulate a heterozygous diploid. The default base caller Albacore (version 2.3.1) was used to base call the raw current signal in the FAST5 files and

generate the final reads in fastq format. An exponential read length distribution, used to simulate reads in humans [163], was used. The error profile of the simulated reads closely resembled that of reads generated with the MinION from cattle (Table S1, Figures S1–S3) and summary statistics were generated using NanoStat and NanoPlot [164]. To simulate a realistic on-farm sequencing preparation a mean read length of 2000 bp was used and various coverages ($2\times$, $4\times$, $6\times$, $8\times$, $10\times$, $12\times$ and $20\times$) were simulated. The simulated Nanopore reads were mapped back to the original ARS-UCD1.2 chromosome 28. Samtools Mpileup [165] and BCFtools call [166] were used to produce VCF files with allele counts at each SNP loci. The individual base quality threshold (Q) was varied between 5, 7, 10 and the default 13 in the mpileup command. The probability of SNP discovery (P) in the BCFtools call command was set to 1 to allow discovery of all variants. An R script was used to call genotypes based on observations of reference or alternate alleles from the VCF file (Figure 4; Table S2). All heterozygous calls were categorised as either accurate or false positives (Figure 4) based on the observed allele matching the known substituted allele. This accounted for false positive heterozygous calls, caused by sequencing errors at the loci of interest.

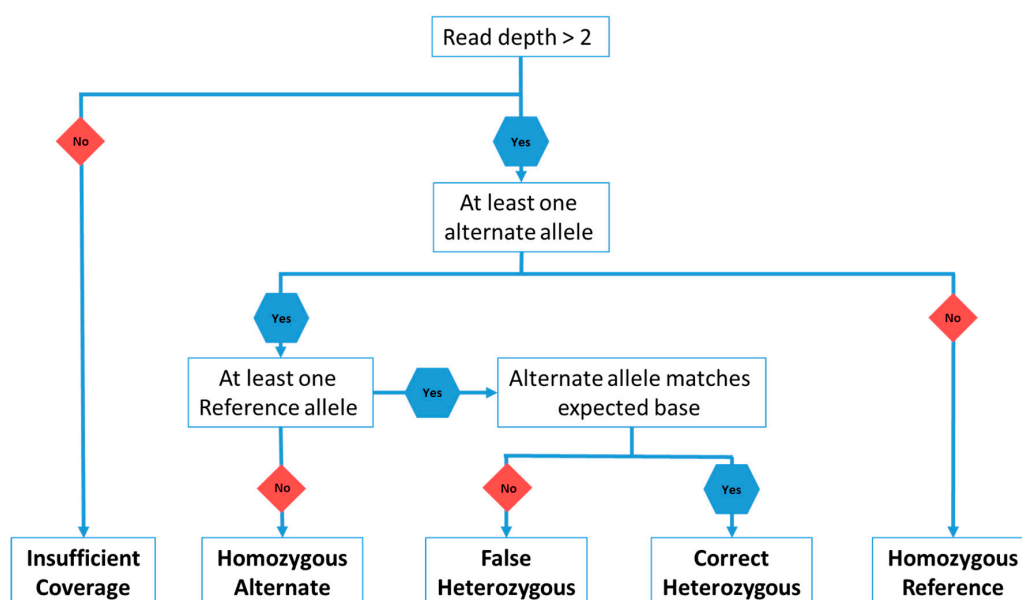


Figure 4. Flow chart for the genotype calling script that was used with the VCF files generated by Samtools and BCFtools.

An additional 10,000 base positions on chromosome 28 were also genotype called from each alignment as controls. All of these 10,000 positions were homozygous for the ARS-UCD1.2 reference allele. These positions were examined to determine the effect of sequencing errors on homozygous genotypes. The same genotype calling methods and parameters were used.

4. Results and Discussion

4.1. Heterozygous Positions

The Q score threshold did not have a significant effect ($p > 0.01$) on the overall number of SNPs called (Figure 5), but did have a significant effect on the percentage of loci called correctly as heterozygous ($p < 0.01$), while sequencing coverage had a significant effect on both the overall number of SNPs called and the number of SNPs called correctly ($p < 0.01$). Accuracies of 85% were achieved with coverage as low as $6\times$ and a Q score threshold of 5. The number of incorrect homozygous reference genotype calls was consistently higher than incorrect homozygous alternate calls (Figures 5 and 6). This was thought to be an artifact of training and mapping to the reference genome ARS-UCD 1.2. Therefore, the same reads were mapped to the mutant chromosome where the proportion of

homozygous reference to homozygous alternate alleles was more even (Figure 7 and Figure S4). This suggests that the bias between homozygous reference and homozygous alternate calls is related to the mapping of reads to a particular reference genome.

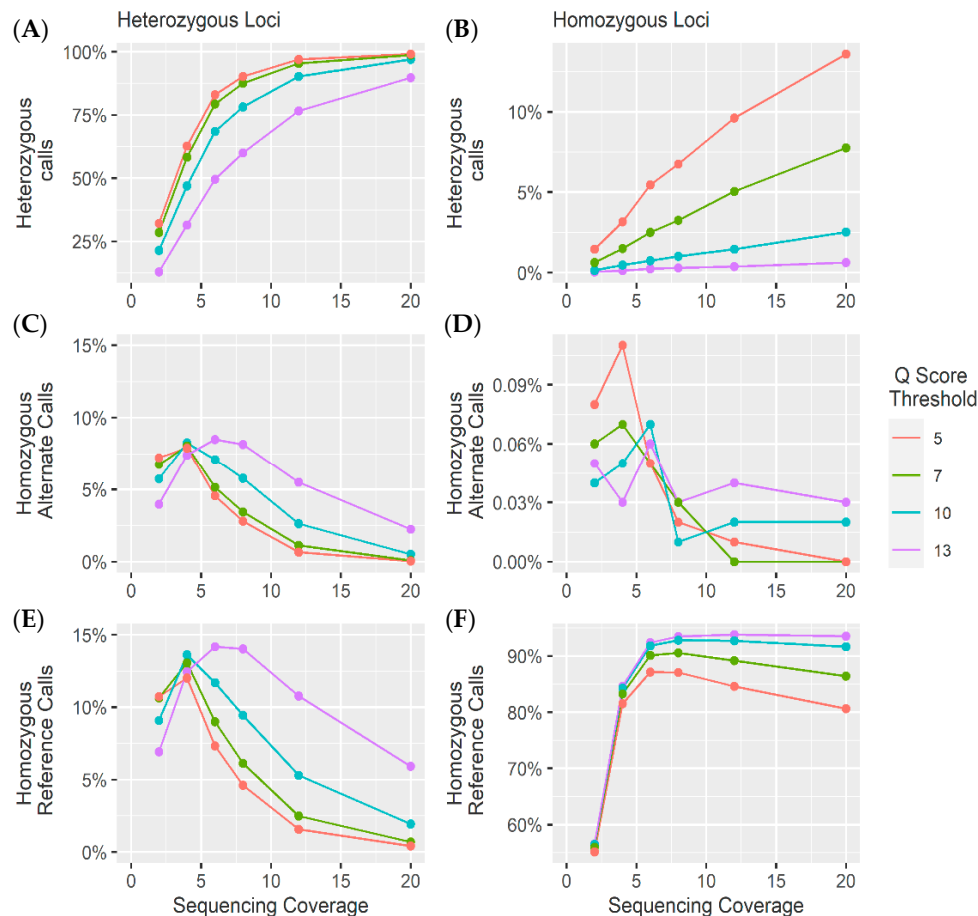


Figure 5. Genotype calling results from the 10,000 heterozygous loci (left) and 10,000 homozygous sites (right) at different sequencing coverages. (A) The percentage of SNPs correctly identified (heterozygous). (B) The percentage of false heterozygous alternate calls. (C) The percentage of false homozygous reference calls. (D) The percentage of false homozygous alternate calls. (E) The percentage of false homozygous reference calls. (F) The percentage of accurate homozygous reference calls.

4.2. Homozygous Positions

Analysis of the genotype calls at the 10,000 homozygous loci demonstrated greater accuracy at lower coverages. However, at higher coverages, the method used to call heterozygotes meant there was an increase in heterozygous calls as a result of sequencing errors (Figures 5–8). As coverage increased, the likelihood of at least one observation of an alternate allele from a sequencing error increased, consequently increasing the number of false heterozygous calls. The Q score threshold also had a significant effect ($p < 0.01$) on the number of false heterozygous calls, with a Q score threshold of 5 having far more false heterozygous calls than other Q scores. This is likely a result of the sequencing errors having a low base quality and therefore more sequencing errors being called with more relaxed Q score cut offs. Accuracies of almost 95% were achieved for these loci at coverages as low as 6x, very similar to results reported by Malmberg, et al. [138]. The percentage of homozygous alternate calls remained extremely low across all sequencing coverages with a maximum of just over 0.1% at 4x sequencing coverage. These results suggest that a low Q threshold is favourable for heterozygous calling. However, a more stringent Q threshold is necessary to prevent false heterozygous

calls at homozygous loci. False homozygous calls at low-coverage heterozygous loci are a considerable limitation of low-coverage genotyping [167]. A number of studies have proposed the maximum likelihood (ML) method as a solution [166,168]. The use of allele dosage/probability based genotyping, as is often implemented when using Illumina data [169] would likely decrease the number of false homozygous genotypes. Wang, et al. [167] demonstrated that this method for genotype correction in GBS data increased the rate of correct genotypes as well as the accuracy of genomic prediction. However, SNP array genotyping was more accurate than the corrected GBS genotypes for coverages below $10 \times$ [166]. Alternatively, Dodds, et al. [161] describe a method for calculating unbiased estimates of relatedness by using only SNPs with genotype calls in both individuals. They refer to this method as kinship using GBS with depth adjustment (KGD) [161]. Despite being developed specifically for use in GBS data, the similarities between GBS and the crush-side genotyping method proposed here suggest KGD could be adapted for crush-side genotyping.

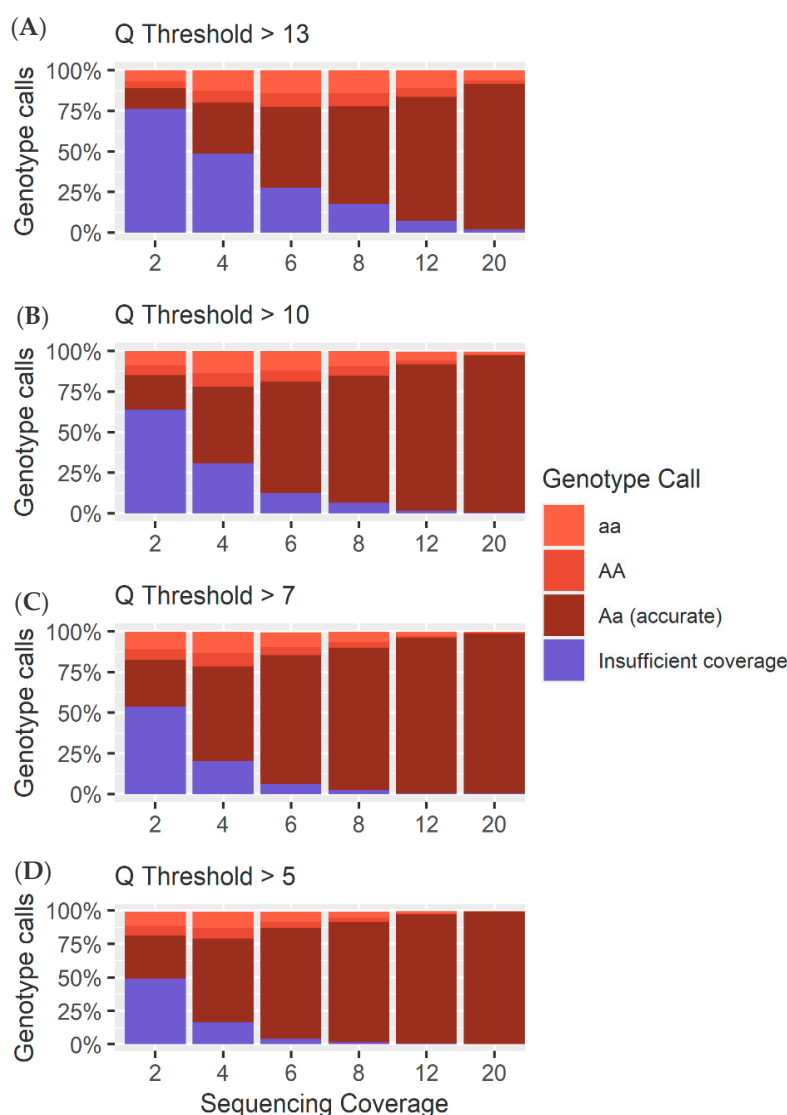


Figure 6. (A) A breakdown of the different calls (Table S2) for each sequencing coverage at a Q threshold of 13 for the 10,000 heterozygous loci. (B) A breakdown of the different calls (Table S2) for each sequencing coverage at a Q threshold of 10 for the 10,000 heterozygous loci. (C) A breakdown of the different calls (Table S2) for each sequencing coverage at a Q threshold of 7 for the 10,000 heterozygous loci. (D) A breakdown of the different calls (Table S2) for each sequencing coverage at a Q threshold of 5 for the 10,000 heterozygous loci.

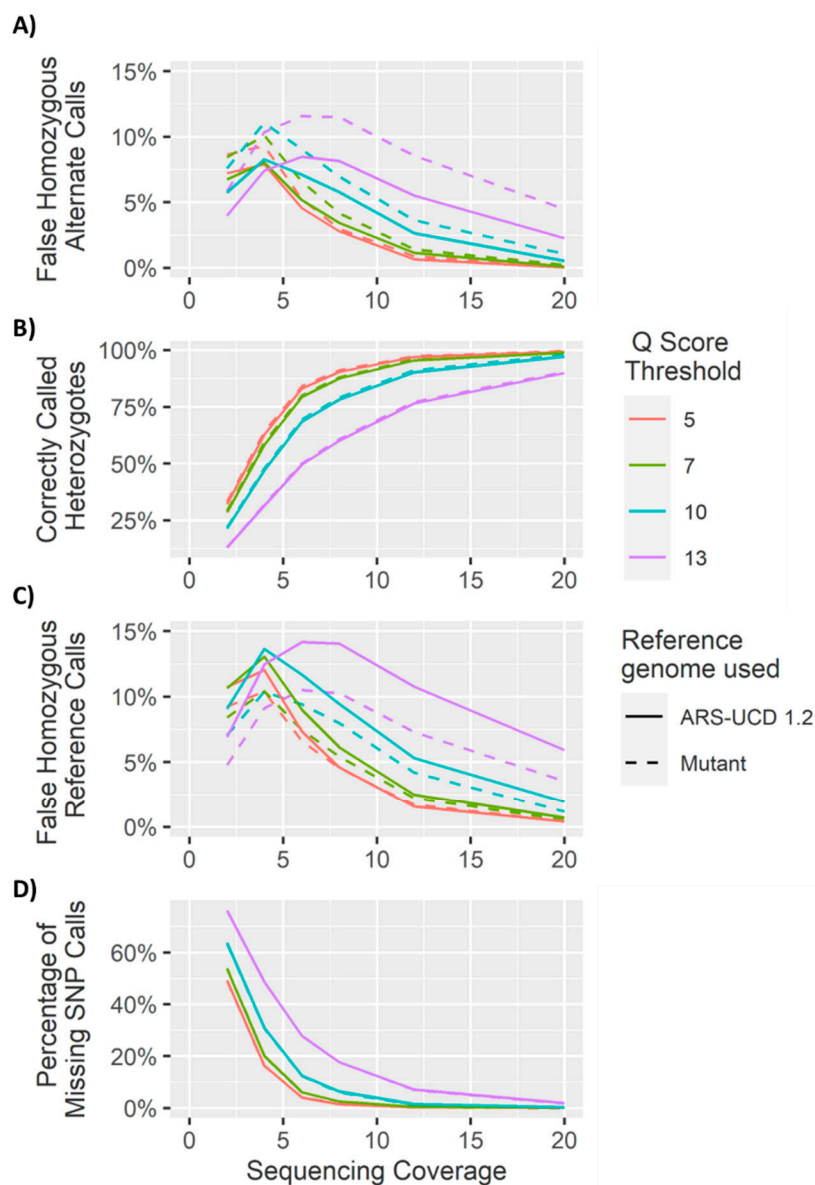


Figure 7. Genotype calling results from the alignments to chromosome 28 of ARS-UCD 1.2 and to the mutant chromosome at the 10,000 heterozygous sites. **(A)** The percentage of SNPs correctly identified (heterozygous). **(B)** The percentage of false homozygous alternate calls. **(C)** The percentage of false homozygous reference calls. **(D)** The percentage of SNP loci with coverage less than 2 $\times$.

Overall, these simulations have highlighted genotype calling using Nanopore data is possible in both homozygous and heterozygous populations. However, a more refined method for genotype calling is required to increase the accuracy of genotype calling in highly heterozygous populations. At high coverage ($>6\times$) the error rates reported here for heterozygous and homozygous positions appear to be within the tolerable threshold of genomic prediction given imputation error rates of 5.5–0.9% are commonly reported [161,170–174]. Validation of the results reported here on real datasets is still required. However these results provide preliminary evidence that Oxford Nanopore sequencing has potential for on-farm SNP genotyping for current GBLUP predictions. Importantly, even read lengths that are relatively short for MinION data did not result in decreased SNP calling accuracy.

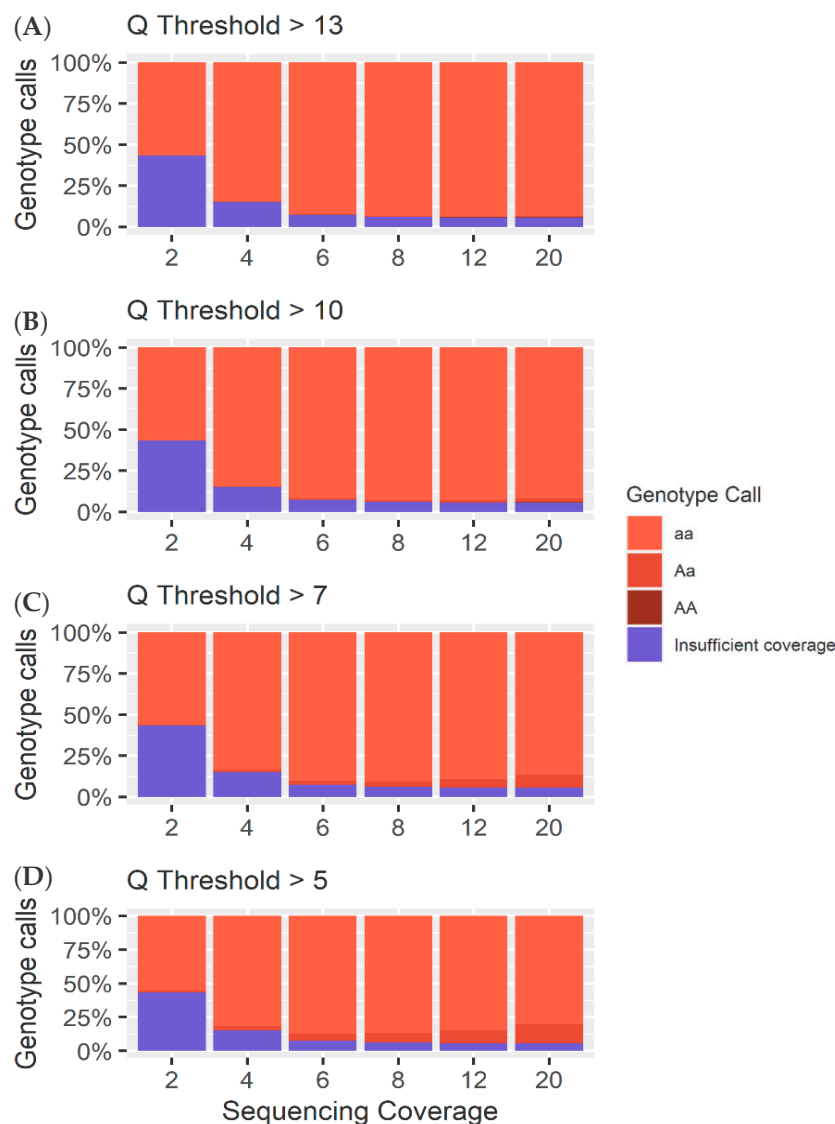


Figure 8. (A) A breakdown of the different calls (Table S2) for each sequencing coverage at a Q threshold of 13 for the 10,000 homozygous loci. (B) A breakdown of the different calls (Table S2) for each sequencing coverage at a Q threshold of 10 for the 10,000 homozygous loci. (C) A breakdown of the different calls (Table S2) for each sequencing coverage at a Q threshold of 7 for the 10,000 homozygous loci. (D) A breakdown of the different calls (Table S2) for each sequencing coverage at a Q threshold of 5 for the 10,000 homozygous loci.

5. Conclusions

The MinION has opened new applications for sequencing technology as well as providing benefits to current sequencing applications. Its portability, read length and real-time base calling capabilities have cemented it as a new rapid diagnostic tool, particularly in humans, despite the high error rates. A number of studies have also now demonstrated potential applications in agriculture as a diagnostic tool for disease and antibiotic resistance [38,42,53]. In particular, O'Donnell, et al. [38] demonstrated the effective use of the MinION to monitor the 2019 outbreak of ASF and developed a real-time rapid diagnostic tool, ASF-FAST. Furthermore, the MinION provides enormous benefits in characterising complex genomic regions for genome assemblies. The snow sheep, giant redwood and giant sequoia genome assemblies have all demonstrated increased contiguity as a result of using long Nanopore reads [5–7]. More accurate genome assemblies will also allow the MinION to be used effectively

in the characterisation and calling of structural variants to increase our understanding of genetic variation [110], which is of particular interest in livestock. This may help quantify some of the missing heritability in a number of common diseases and economically desirable traits. Additionally, the ability to directly capture base modifications in real time could help to one day incorporate methylation and epigenetics into genomic predictions. With strong evidence suggesting a correlation between premature ageing and methylation [130–132], we see carcass ossification as an obvious application of this.

With the ability to rapidly genotype both SNPs and SVs, we propose the use of the MinION for crush-side genotyping in Australia's Northern cattle industry to deliver on-farm genomic prediction. The capacity to sequence, map and process reads in real time means that crush-side genotyping can provide rapid genomic prediction to industries where laboratory turnaround time is a major limitation. The power to call causative mutations rather than SNPs sharing LD and rapidly incorporate new markers using a script rather than developing entirely new SNP arrays will have significant merit as new causal mutations are discovered.

Our preliminary in silico investigation assessed the feasibility of low-coverage SNP genotyping in a heterozygous population using simulated Nanopore data. An artificial diploid of *B. taurus* chromosome 28 with 10,000 SNPs was created using the package SimuG and filtering and calling parameters were altered to investigate genotyping accuracies. The results suggested that coverages as low as 6× can produce genotyping accuracies greater than 85%. These results provide preliminary evidence that MinION sequencing has the potential to be used for rapid on-farm crush-side genotyping for genomic prediction in Australia's northern beef industry.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4425/11/12/1478/s1>, Figure S1: Read length histogram for real and simulated Nanopore data, Figure S2: Average read quality versus aligned length for real and simulated Nanopore data, Figure S3: Average read quality versus read length for real and simulated Nanopore data, Figure S4: Comparison of genotype calls for heterozygous positions aligned to different reference genomes, Table S1: Comparison between the simulated and real Nanopore read statistics, Table S2: Definitions of Genotype Calls.

Author Contributions: Conceptualization, E.M.R. and B.J.H.; methodology, H.J.L.; validation, H.J.L.; formal analysis, H.J.L.; writing—original draft preparation, H.J.L.; writing—review and editing, L.T.N., E.M.R., and B.J.H.; supervision, B.J.H. and E.M.R.; funding acquisition, B.J.H. and H.J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Meat and Livestock Australia, grant number L.GEN.1808, as well as Meat and Livestock Australia's post graduate top-up scholarship B.STU.2001.

Acknowledgments: The authors would like to acknowledge contributions from all members of the Centre for Animal Science at the Queensland Alliance for Agriculture and Food Innovation, in particular the Hayes' research lab; contributions and discussions with Bailey Engle, Christie Warburton, James Copley, Roy Costilla Monteagudo, Simon Lee, Kai Voss-Fels and Hamish Buchorn. The authors would also like to acknowledge Professor Tim Mahony and Sandy Jarrett from the Department of Agriculture and Fisheries for discussions regarding bovine disease. The authors would also like to thank the reviewers for their significant contributions which helped to strengthen the manuscript overall. Finally, the authors would like to acknowledge BioRender (BioRender.com) which was used to create Figures 1 and 3.

Conflicts of Interest: The authors declare no actual or potential conflicts of interest that affect their ability to objectively present or review research or data.

References

1. Pareek, C.S.; Smoczynski, R.; Tretyn, A. Sequencing technologies and genome sequencing. *J. Appl. Genet.* **2011**, *52*, 413–435. [[CrossRef](#)] [[PubMed](#)]
2. Heather, J.M.; Chain, B. The sequence of sequencers: The history of sequencing DNA. *Genom.* **2016**, *107*, 1–8. [[CrossRef](#)]
3. Acinas, S.G.; Sarma-Rupavtarm, R.; Klepac-Ceraj, V.; Polz, M.F. PCR-induced sequence artifacts and bias: Insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *J. Appl. Environ. Microbiol.* **2005**, *71*, 8966–8969. [[CrossRef](#)] [[PubMed](#)]

4. Kebschull, J.M.; Zador, A.M. Sources of PCR-induced distortions in high-throughput sequencing data sets. *Nucleic Acids Res.* **2015**, *43*. [CrossRef] [PubMed]
5. Jansen, H.J.; Liem, M.; Jong-Raadsen, S.A.; Dufour, S.; Weltzien, F.A.; Swinkels, W.; Koelewijn, A.; Palstra, A.P.; Pelster, B.; Spink, H.P.; et al. Author Correction: Rapid de novo assembly of the European eel genome from nanopore sequencing reads. *Sci. Rep.* **2019**, *9*, 7911. [CrossRef] [PubMed]
6. Michael, T.P.; Jupe, F.; Bemm, F.; Motley, S.T.; Sandoval, J.P.; Lanz, C.; Loudet, O.; Weigel, D.; Ecker, J.R. High contiguity Arabidopsis thaliana genome assembly with a single nanopore flow cell. *Nat. Commun.* **2018**, *9*, 541. [CrossRef]
7. Diaz-Virague, F.; Pita, S.; Greif, G.; De Souza, R.D.M.; Iraola, G.; Robello, C. Nanopore Sequencing Significantly Improves Genome Assembly of the Protozoan Parasite Trypanosoma cruzi. *Genome Biol. Evol.* **2019**, *11*, 1952–1957. [CrossRef]
8. Hunt, S.P.; Jarvis, D.E.; Larsen, D.J.; Mosyakin, S.L.; Kolano, B.A.; Jackson, E.W.; Martin, S.L.; Jellen, E.N.; Maughan, P.J. A Chromosome-Scale Assembly of the Garden Orach (*Atriplex hortensis* L.) Genome Using Oxford Nanopore Sequencing. *Front. Plant. Sci.* **2020**, *11*. [CrossRef]
9. Liu, L.; Li, Y.; Li, S.; Hu, N.; He, Y.; Pong, R.; Lin, D.; Lu, L.; Law, M. Comparison of next-generation sequencing systems. *J. Biomed. Biotechnol.* **2012**, *2012*, 251364. [CrossRef]
10. Leggett, R.M.; Clark, M.D. A world of opportunities with nanopore sequencing. *J. Exp. Bot.* **2017**, *68*, 5419–5429. [CrossRef]
11. Weirather, J.L.; de Cesare, M.; Wang, Y.; Piazza, P.; Sebastiano, V.; Wang, X.J.; Buck, D.; Au, K.F. Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis. *F1000Research* **2017**, *6*, 100. [CrossRef] [PubMed]
12. Oxford Nanopore Technologies. Products. Available online: <https://nanoporetech.com/products> (accessed on 19 June 2020).
13. Ashton, P.M.; Nair, S.; Dallman, T.; Rubino, S.; Rabsch, W.; Mwaigwisya, S.; Wain, J.; O’Grady, J. MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. *Nat. Biotechnol.* **2015**, *33*, 296–300. [CrossRef] [PubMed]
14. De Santis, D.; Truong, L.; Martinez, P.; D’Orsogna, L. Rapid high-resolution HLA genotyping by MinION Oxford nanopore sequencing for deceased donor organ allocation. *HLA* **2020**, *96*, 141–162. [CrossRef] [PubMed]
15. Wang, X.V.; Blades, N.; Ding, J.; Sultana, R.; Parmigiani, G. Estimation of sequencing error rates in short reads. *BMC Bioinform.* **2012**, *13*, 185. [CrossRef]
16. Minoche, A.E.; Dohm, J.C.; Himmelbauer, H. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. *Genome Biol.* **2011**, *12*, R112. [CrossRef]
17. Rang, F.J.; Kloosterman, W.P.; de Ridder, J. From squiggle to basepair: Computational approaches for improving nanopore sequencing read accuracy. *Genome Biol.* **2018**, *19*, 90. [CrossRef]
18. Weischenfeldt, J.; Symmons, O.; Spitz, F.; Korbel, J.O. Phenotypic impact of genomic structural variation: Insights from and for human disease. *Nat. Rev. Genet.* **2013**, *14*, 125–138. [CrossRef]
19. Shafin, K.; Pesout, T.; Lorig-Roach, R.; Haukness, M.; Olsen, H.E.; Bosworth, C.; Armstrong, J.; Tigyi, K.; Maurer, N.; Koren, S.; et al. Nanopore sequencing and the Shasta toolkit enable efficient de novo assembly of eleven human genomes. *Nat. Biotechnol.* **2020**, *38*, 1044–1053. [CrossRef]
20. James, P.; Stoddart, D.; Harrington, E.D.; Beaulaurier, J.; Ly, L.; Reid, S.; Turner, D.J.; Juul, S. LamPORE: Rapid, accurate and highly scalable molecular screening for SARS-CoV-2 infection, based on nanopore sequencing. *medRxiv* **2020**. [CrossRef]
21. Samarakoon, H.; Punchihewa, S.; Senanayake, A.; Ragel, R.; Gamaarachchi, H. F5N: Nanopore Sequence Analysis Toolkit for Android Smartphones. *bioRxiv* **2020**. [CrossRef]
22. Mitsuhashi, S.; Kryukov, K.; Nakagawa, S.; Takeuchi, J.S.; Shiraishi, Y.; Asano, K.; Imanishi, T. A portable system for rapid bacterial composition analysis using a nanopore-based sequencer and laptop computer. *Sci. Rep.-UK* **2017**, *7*, 1–9. [CrossRef] [PubMed]
23. Quick, J.; Ashton, P.; Calus, S.; Chatt, C.; Gossain, S.; Hawker, J.; Nair, S.; Neal, K.; Nye, K.; Peters, T.; et al. Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of Salmonella. *Genome Biol.* **2015**, *16*. [CrossRef] [PubMed]

24. Greninger, A.L.; Naccache, S.N.; Federman, S.; Yu, G.X.; Mbala, P.; Bres, V.; Stryke, D.; Bouquet, J.; Somasekar, S.; Linnen, J.M.; et al. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med.* **2015**, *7*. [CrossRef] [PubMed]
25. Greig, D.R.; Jenkins, C.; Gharbia, S.; Dallman, T.J. Comparison of single-nucleotide variants identified by Illumina and Oxford Nanopore technologies in the context of a potential outbreak of Shiga toxin-producing *Escherichia coli*. *Gigascience* **2019**, *8*. [CrossRef] [PubMed]
26. Bialasiewicz, S.; Duarte, T.P.S.; Nguyen, S.H.; Sukumaran, V.; Stewart, A.; Appleton, S.; Pitt, M.E.; Bainomugisa, A.; Jennison, A.V.; Graham, R.; et al. Rapid diagnosis of *Capnocytophaga canimorsus* septic shock in an immunocompetent individual using real-time Nanopore sequencing: A case report. *BMC Infect. Dis.* **2019**, *19*. [CrossRef]
27. Quick, J.; Loman, N.J.; Duraffour, S.; Simpson, J.T.; Ettore, S.; Cowley, L.; Bore, J.A.; Koundouno, R.; Dudas, G.; Mikhail, A.; et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature* **2016**, *530*, 228. [CrossRef]
28. Giovanetti, M.; Faria, N.R.; Lourenco, J.; de Jesus, J.G.; Xavier, J.; Claro, I.M.; Kraemer, M.U.G.; Fonseca, V.; Dellicour, S.; Theze, J.; et al. Genomic and Epidemiological Surveillance of Zika Virus in the Amazon Region. *Cell Rep.* **2020**, *30*, 2275–2283. [CrossRef]
29. Taiaaroa, G.; Rawlinson, D.; Featherstone, L.; Pitt, M.; Caly, L.; Druce, J.; Purcell, D.; Harty, L.; Tran, T.; Roberts, J.; et al. Direct RNA sequencing and early evolution of SARS-CoV-2. *bioRxiv* **2020**. [CrossRef]
30. Kim, D.; Lee, J.Y.; Yang, J.S.; Kim, J.W.; Kim, V.N.; Chang, H. The Architecture of SARS-CoV-2 Transcriptome. *Cell* **2020**, *181*, 914–921 e910. [CrossRef]
31. Xu, Y.F.; Lewandowski, K.; Jeffery, K.; Downs, L.O.; Foster, D.; Sanderson, N.D.; Kavanagh, J.; Vaughan, A.; Salvagno, C.; Vipond, R.; et al. Nanopore metagenomic sequencing to investigate nosocomial transmission of human metapneumovirus from a unique genetic group among haematology patients in the United Kingdom. *J. Infect.* **2020**, *80*, 571–577. [CrossRef]
32. Rhodes, J.; Abdolrasouli, A.; Farrer, R.A.; Cuomo, C.A.; Aanensen, D.M.; Armstrong-James, D.; Fisher, M.C.; Schelenz, S. Genomic epidemiology of the UK outbreak of the emerging human fungal pathogen *Candida auris* (vol 7, pg 43, 2018). *Emerg. Microbes Infect.* **2018**, *7*. [CrossRef]
33. Costard, S.; Wieland, B.; de Glanville, W.; Jori, F.; Rowlands, R.; Vosloo, W.; Roger, F.; Pfeiffer, D.U.; Dixon, L.K. African swine fever: How can global spread be prevented? *Philos. Trans. R. Soc. B* **2009**, *364*, 2683–2696. [CrossRef]
34. Galindo, I.; Alonso, C. African Swine Fever Virus: A Review. *Viruses* **2017**, *9*, 103. [CrossRef]
35. Mason-D'Croz, D.; Bogard, J.R.; Herrero, M.; Robinson, S.; Sulser, T.B.; Wiebe, K.; Willenbockel, D.; Godfray, H.C.J. Modelling the global economic consequences of a major African swine fever outbreak in China. *Nat. Food* **2020**, *1*, 221–228. [CrossRef]
36. Dixon, L.K.; Islam, M.; Nash, R.; Reis, A.L. African swine fever virus evasion of host defences. *Virus Res.* **2019**, *266*, 25–33. [CrossRef] [PubMed]
37. Liu, L.H.; Atim, S.; LeBlanc, N.; Rauh, R.; Esau, M.; Chenais, E.; Mwebe, R.; Nelson, W.M.; Masembe, C.; Nantima, N.; et al. Overcoming the challenges of pen-side molecular diagnosis of African swine fever to support outbreak investigations under field conditions. *Transbound Emerg. Dis.* **2019**, *66*, 908–914. [CrossRef]
38. O'Donnell, V.K.; Grau, F.R.; Mayr, G.A.; Samayoa, T.L.S.; Dodd, K.A.; Barrette, R.W. 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. *J. Clin. Microbiol.* **2020**, *58*. [CrossRef]
39. Ryan, W.; Jeremy, C.; Nick, L. PoreChop. Available online: <https://github.com/rrwick/Porechop> (accessed on 8 October 2020).
40. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef]
41. Gallardo, C.; Fernandez-Pinero, J.; Pelayo, V.; Gazeau, I.; Markowska-Daniel, I.; Pridotkas, G.; Nieto, R.; Fernandez-Pacheco, P.; Bokhan, S.; Nevolko, O.; et al. Genetic Variation among African Swine Fever Genotype II Viruses, Eastern and Central Europe. *Emerg. Infect. Dis.* **2014**, *20*, 1544–1547. [CrossRef]
42. Theuns, S.; Vanmechelen, B.; Bernaert, Q.; Deboutte, W.; Vandenhoe, M.; Beller, L.; Matthijssens, J.; Maes, P.; Nauwynck, H.J. Nanopore sequencing as a revolutionary diagnostic tool for porcine viral enteric disease complexes identifies porcine kobuvirus as an important enteric virus. *Sci. Rep.-UK* **2018**, *8*, 9830. [CrossRef]

43. Omaleki, L.; Blackall, P.J.; Cuddihy, T.; Beatson, S.A.; Forde, B.M.; Turni, C. Using genomics to understand inter- and intra- outbreak diversity of *Pasteurella multocida* isolates associated with fowl cholera in meat chickens. *Microb. Genom.* **2020**, *6*, e000346. [CrossRef] [PubMed]
44. Baksi, S.; Rao, N.; Khan, M.; Chauhan, P.; Chauhan, A. Efficacy of Inactivated Fowl Cholera Vaccine in Chickens. *Psm. Vet. Res.* **2018**, *3*, 32–35.
45. Snowden, G.D.; Van Vleck, L.D.; Cundiff, L.V.; Bennett, G.L. Bovine respiratory disease in feedlot cattle: Environmental, genetic, and economic factors. *J. Anim. Sci.* **2006**, *84*, 1999–2008. [CrossRef] [PubMed]
46. Noyes, N.R.; Benedict, K.M.; Gow, S.P.; Booker, C.W.; Hannon, S.J.; McAllister, T.A.; Morley, P.S. Mannheimia haemolytica in Feedlot Cattle: Prevalence of Recovery and Associations with Antimicrobial Use, Resistance, and Health Outcomes. *J. Vet. Intern. Med.* **2015**, *29*, 705–713. [CrossRef]
47. Griffin, D.; Chengappa, M.M.; Kuszak, J.; McVey, D.S. Bacterial Pathogens of the Bovine Respiratory Disease Complex. *Vet. Clin. N. Am.-Food A* **2010**, *26*, 381–394. [CrossRef]
48. Taylor, J.D.; Fulton, R.W.; Lehenbauer, T.W.; Step, D.L.; Confer, A.W. The epidemiology of bovine respiratory disease: What is the evidence for predisposing factors? *Can. Vet. J.* **2010**, *51*, 1095–1102.
49. Edwards, T.A. Control methods for bovine respiratory disease for feedlot cattle. *Vet. Clin. N. Am. Food Anim Pract.* **2010**, *26*, 273–284. [CrossRef]
50. McCabe, M.S.; Cormican, P.; Johnston, D.; Earley, B. Simultaneous detection of DNA and RNA virus species involved in bovine respiratory disease by PCR-free rapid tagmentation-based library preparation and MinION nanopore sequencing. *bioRxiv* **2018**. [CrossRef]
51. Zhang, M.; Huang, Y.; Godson, D.L.; Fernando, C.; Alexander, T.W.; Hill, J.E. Assessment of metagenomic sequencing and qPCR for detection of influenza D virus in bovine respiratory tract samples. *bioRxiv* **2020**. [CrossRef]
52. Shaffer, L. Inner Workings: Portable DNA sequencer helps farmers stymie devastating viruses. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 3351–3353. [CrossRef]
53. Lim, A.; Naidenova, B.; Bates, H.; Willyerd, K.; Snider, T.; Couger, M.B.; Chen, C.; Ramachandran, A. Nanopore ultra-long read sequencing technology for antimicrobial resistance detection in Mannheimia haemolytica. *J. Microbiol. Meth.* **2019**, *159*, 138–147. [CrossRef] [PubMed]
54. Andrew, R.; James, H.; Nick, L.; Aine, O.t.; Chris, S.; Roger, P.; Botond, S. RAMPART. Available online: <https://github.com/artic-network/rampart> (accessed on 8 October 2020).
55. Nick, L.; Will, R.; Andrew, R. nCoV-2019 Novel Coronavirus Bioinformatics Protocol. Available online: <https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html> (accessed on 8 October 2020).
56. Payne, A.; Holmes, N.; Clarke, T.; Munro, R.; Debebe, B.; Loose, M. Nanopore adaptive sequencing for mixed samples, whole exome capture and targeted panels. *bioRxiv* **2020**. [CrossRef]
57. Kovaka, S.; Fan, Y.; Ni, B.; Timp, W.; Schatz, M.C. Targeted nanopore sequencing by real-time mapping of raw electrical signal with UNCALLED. *bioRxiv* **2020**. [CrossRef]
58. Houe, H. Epidemiological features and economical importance of bovine virus diarrhoea virus (BVDV) infections. *Vet. Microbiol.* **1999**, *64*, 89–107. [CrossRef]
59. Alexandersen, S.; Zhang, Z.; Donaldson, A.I.; Garland, A.J. The pathogenesis and diagnosis of foot-and-mouth disease. *J. Comp. Pathol.* **2003**, *129*, 1–36. [CrossRef]
60. Thornton, P.K. Livestock production: Recent trends, future prospects. *Philos. Trans. R. Soc. Lond B Biol. Sci.* **2010**, *365*, 2853–2867. [CrossRef] [PubMed]
61. Bickhart, D.M.; Liu, G.E. The challenges and importance of structural variation detection in livestock. *Front. Genet.* **2014**, *5*, 37. [CrossRef]
62. Zimin, A.V.; Delcher, A.L.; Florea, L.; Kelley, D.R.; Schatz, M.C.; Puiu, D.; Hanrahan, F.; Pertea, G.; Van Tassel, C.P.; Sonstegard, T.S.; et al. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.* **2009**, *10*, R42. [CrossRef]
63. International Chicken Genome Sequencing, C. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **2004**, *432*, 695–716. [CrossRef]
64. International Sheep Genomics, C.; Archibald, A.L.; Cockett, N.E.; Dalrymple, B.P.; Faraut, T.; Kijas, J.W.; Maddox, J.F.; McEwan, J.C.; Hutton Oddy, V.; Raadsma, H.W.; et al. The sheep genome reference sequence: A work in progress. *Anim. Genet.* **2010**, *41*, 449–453. [CrossRef]
65. Fang, X.; Mou, Y.; Huang, Z.; Li, Y.; Han, L.; Zhang, Y.; Feng, Y.; Chen, Y.; Jiang, X.; Zhao, W.; et al. The sequence and analysis of a Chinese pig genome. *Gigascience* **2012**, *1*, 16. [CrossRef] [PubMed]

66. Zimin, A.V.; Kelley, D.R.; Roberts, M.; Marçais, G.; Salzberg, S.L.; Yorke, J.A. Mis-assembled “segmental duplications” in two versions of the *Bos taurus* genome. *PLoS ONE* **2012**, *7*, e42680. [[CrossRef](#)] [[PubMed](#)]
67. Lang, D.; Zhang, S.; Ren, P.; Liang, F.; Sun, Z.; Meng, G.; Tan, Y.; Hu, J.; Li, X.; Lai, Q.; et al. Comparison of the two up-to-date sequencing technologies for genome assembly: HiFi reads of Pacbio Sequel II system and ultralong reads of Oxford Nanopore. *bioRxiv* **2020**. [[CrossRef](#)]
68. Upadhyay, M.; Hauser, A.; Kunz, E.; Krebs, S.; Blum, H.; Dotsev, A.; Okhlopkov, I.; Bagirov, V.; Brem, G.; Zinovieva, N.; et al. The first draft genome assembly of Snow sheep (*Ovis nivicola*). *Genome Biol. Evol.* **2020**. [[CrossRef](#)]
69. Scott, A.D.; Zimin, A.V.; Puiu, D.; Workman, R.; Britton, M.; Zaman, S.; Caballero, M.; Read, A.C.; Bogdanove, A.J.; Burns, E.; et al. The giant sequoia genome and proliferation of disease resistance genes. *bioRxiv* **2020**. [[CrossRef](#)]
70. Pillai, S.; Gopalan, V.; Lam, A.K.-Y. Review of sequencing platforms and their applications in phaeochromocytoma and paragangliomas. *Crit. Rev. Oncol. Hematol.* **2017**, *116*, 58–67. [[CrossRef](#)]
71. Rhoads, A.; Au, K.F. PacBio Sequencing and Its Applications. *Genom. Proteom. Bioinform.* **2015**, *13*, 278–289. [[CrossRef](#)]
72. Loit, K.; Adamson, K.; Bahram, M.; Puusepp, R.; Anslan, S.; Kiiker, R.; Drenkhan, R.; Tedersoo, L. Relative Performance of MinION (Oxford Nanopore Technologies) versus Sequel (Pacific Biosciences) Third-Generation Sequencing Instruments in Identification of Agricultural and Forest Fungal Pathogens. *Appl. Environ. Microbiol.* **2019**, *85*, e01368-19. [[CrossRef](#)]
73. Hon, T.; Mars, K.; Young, G.; Tsai, Y.-C.; Karalius, J.W.; Landolin, J.M.; Maurer, N.; Kudrna, D.; Hardigan, M.A.; Steiner, C.C.; et al. Highly accurate long-read HiFi sequencing data for five complex genomes. *bioRxiv* **2020**. [[CrossRef](#)]
74. Kono, N.; Arakawa, K. Nanopore sequencing: Review of potential applications in functional genomics. *Dev. Growth Differ.* **2019**, *61*, 316–326. [[CrossRef](#)]
75. Song, B.; Mott, R.; Gan, X. Recovery of novel association loci in *Arabidopsis thaliana* and *Drosophila melanogaster* through leveraging INDELs association and integrated burden test. *PLoS Genet.* **2018**, *14*, e1007699. [[CrossRef](#)]
76. Rosen, B.D.; Bickhart, D.M.; Schnabel, R.D.; Koren, S.; Elsik, C.G.; Tseng, E.; Rowan, T.N.; Low, W.Y.; Zimin, A.; Couldrey, C.; et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* **2020**, *9*. [[CrossRef](#)]
77. Cretu Stancu, M.; van Roosmalen, M.J.; Renkens, I.; Nieboer, M.M.; Middelkamp, S.; de Ligt, J.; Pregno, G.; Giachino, D.; Mandrile, G.; Espejo Valle-Inclan, J.; et al. Mapping and phasing of structural variation in patient genomes using nanopore sequencing. *Nat. Commun.* **2017**, *8*, 1326. [[CrossRef](#)]
78. Norris, A.L.; Workman, R.E.; Fan, Y.; Eshleman, J.R.; Timp, W. Nanopore sequencing detects structural variants in cancer. *Cancer Biol. Ther.* **2016**, *17*, 246–253. [[CrossRef](#)]
79. Wray, N.R.; Yang, J.; Hayes, B.J.; Price, A.L.; Goddard, M.E.; Visscher, P.M. Pitfalls of predicting complex traits from SNPs. *Nat. Rev. Genet.* **2013**, *14*, 507–515. [[CrossRef](#)]
80. Haile-Mariam, M.; Nieuwhof, G.J.; Beard, K.T.; Konstatinov, K.V.; Hayes, B.J. Comparison of heritabilities of dairy traits in Australian Holstein-Friesian cattle from genomic and pedigree data and implications for genomic evaluations. *J. Anim. Breed. Genet.* **2013**, *130*, 20–31. [[CrossRef](#)]
81. Theunissen, F.; Flynn, L.L.; Anderton, R.S.; Mastaglia, F.; Pytte, J.; Jiang, L.; Hodgetts, S.; Burns, D.K.; Saunders, A.; Fletcher, S.; et al. Structural Variants May Be a Source of Missing Heritability in sALS. *Front. Neurosci-Switz* **2020**, *14*. [[CrossRef](#)]
82. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorff, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A.; et al. Finding the missing heritability of complex diseases. *Nature* **2009**, *461*, 747–753. [[CrossRef](#)]
83. Ravenhall, M.; Benavente, E.D.; Sutherland, C.J.; Baker, D.A.; Campino, S.; Clark, T.G. An analysis of large structural variation in global *Plasmodium falciparum* isolates identifies a novel duplication of the chloroquine resistance associated gene. *Sci. Rep.* **2019**, *9*, 8287. [[CrossRef](#)]
84. Rodríguez-Santiago, B.; Malats, N.; Rothman, N.; Armengol, L.; Garcia-Closas, M.; Kogevinas, M.; Villa, O.; Hutchinson, A.; Earl, J.; Marenne, G.; et al. Mosaic uniparental disomies and aneuploidies as large structural variants of the human genome. *Am. J. Hum. Genet.* **2010**, *87*, 129–138. [[CrossRef](#)]

85. Walsh, T.; McClellan, J.M.; McCarthy, S.E.; Addington, A.M.; Pierce, S.B.; Cooper, G.M.; Nord, A.S.; Kusenda, M.; Malhotra, D.; Bhandari, A.; et al. Rare Structural Variants Disrupt Multiple Genes in Neurodevelopmental Pathways in Schizophrenia. *Science* **2008**, *320*, 539–543. [\[CrossRef\]](#) [\[PubMed\]](#)
86. Harewood, L.; Chaignat, E.; Reymond, A. Structural variation and its effect on expression. *Methods Mol. Biol.* **2012**, *838*, 173–186. [\[CrossRef\]](#) [\[PubMed\]](#)
87. Lupski, J.R.; de Oca-Luna, R.M.; Slaugenhaupt, S.; Pentao, L.; Guzzetta, V.; Trask, B.J.; Saucedo-Cardenas, O.; Barker, D.F.; Killian, J.M.; Garcia, C.A.; et al. DNA duplication associated with Charcot-Marie-Tooth disease type 1A. *Cell* **1991**, *66*, 219–232. [\[CrossRef\]](#)
88. Gonzalez, E.; Kulkarni, H.; Bolivar, H.; Mangano, A.; Sanchez, R.; Catano, G.; Nibbs, R.J.; Freedman, B.I.; Quinones, M.P.; Bamshad, M.J.; et al. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* **2005**, *307*, 1434–1440. [\[CrossRef\]](#) [\[PubMed\]](#)
89. Bayes, M.; Magano, L.F.; Rivera, N.; Flores, R.; Perez Jurado, L.A. Mutational mechanisms of Williams-Beuren syndrome deletions. *Am. J. Hum. Genet.* **2003**, *73*, 131–151. [\[CrossRef\]](#) [\[PubMed\]](#)
90. Bi, W.; Yan, J.; Stankiewicz, P.; Park, S.S.; Walz, K.; Boerkoel, C.F.; Potocki, L.; Shaffer, L.G.; Devriendt, K.; Nowaczyk, M.J.; et al. Genes in a refined Smith-Magenis syndrome critical deletion interval on chromosome 17p11.2 and the syntenic region of the mouse. *Genome Res.* **2002**, *12*, 713–728. [\[CrossRef\]](#)
91. Yang, Y.; Chung, E.K.; Wu, Y.L.; Savelli, S.L.; Nagaraja, H.N.; Zhou, B.; Hebert, M.; Jones, K.N.; Shu, Y.; Kitzmiller, K.; et al. Gene copy-number variation and associated polymorphisms of complement component C4 in human systemic lupus erythematosus (SLE): Low copy number is a risk factor for and high copy number is a protective factor against SLE susceptibility in European Americans. *Am. J. Hum. Genet.* **2007**, *80*, 1037–1054. [\[CrossRef\]](#)
92. Shanta, O.; Noor, A.; Sebat, J.; Chaisson, M.J.P.; Sanders, A.D.; Zhao, X.F.; Malhotra, A.; Porubsky, D.; Rausch, T.; Gardner, E.J.; et al. The effects of common structural variants on 3D chromatin structure. *BMC Genom.* **2020**, *21*. [\[CrossRef\]](#)
93. Bianco, S.; Lupianez, D.G.; Chiariello, A.M.; Annunziatella, C.; Kraft, K.; Schopflin, R.; Wittler, L.; Andrey, G.; Vingron, M.; Pombo, A.; et al. Polymer physics predicts the effects of structural variants on chromatin architecture. *Nat. Genet.* **2018**, *50*, 662–667. [\[CrossRef\]](#)
94. Sanders, S.J.; Ercan-Sencicek, A.G.; Hus, V.; Luo, R.; Murtha, M.T.; Moreno-De-Luca, D.; Chu, S.H.; Moreau, M.P.; Gupta, A.R.; Thomson, S.A.; et al. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* **2011**, *70*, 863–885. [\[CrossRef\]](#)
95. Craddock, N.; Hurles, M.E.; Cardin, N.; Pearson, R.D.; Plagnol, V.; Robson, S.; Vukcevic, D.; Barnes, C.; Conrad, D.F.; Giannoulatou, E.; et al. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature* **2010**, *464*, 713–U786. [\[CrossRef\]](#) [\[PubMed\]](#)
96. Williams, N.M.; Zaharieva, I.; Martin, A.; Langley, K.; Mantripragada, K.; Fossdal, R.; Stefansson, H.; Stefansson, K.; Magnusson, P.; Gudmundsson, O.O.; et al. Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: A genome-wide analysis. *Lancet* **2010**, *376*, 1401–1408. [\[CrossRef\]](#)
97. Roses, A.D.; Akkari, P.A.; Chiba-Falek, O.; Lutz, M.W.; Gottschalk, W.K.; Saunders, A.M.; Saul, B.; Sundseth, S.; Burns, D. Structural variants can be more informative for disease diagnostics, prognostics and translation than current SNP mapping and exon sequencing. *Expert Opin. Drug Met.* **2016**, *12*, 135–147. [\[CrossRef\]](#) [\[PubMed\]](#)
98. Roses, A.D.; Lutz, M.W.; Amrine-Madsen, H.; Saunders, A.M.; Crenshaw, D.G.; Sundseth, S.S.; Huentelman, M.J.; Welsh-Bohmer, K.A.; Reiman, E.M. A TOMM40 variable-length polymorphism predicts the age of late-onset Alzheimer's disease. *Pharm. J.* **2010**, *10*, 375–384. [\[CrossRef\]](#)
99. Chakraborty, M.; Emerson, J.J.; Macdonald, S.J.; Long, A.D. Structural variants exhibit widespread allelic heterogeneity and shape variation in complex traits. *Nat. Commun.* **2019**, *10*. [\[CrossRef\]](#)
100. Voichek, Y.; Weigel, D. Identifying genetic variants underlying phenotypic variation in plants without complete genomes. *Nat. Genet.* **2020**, *52*, 534–540. [\[CrossRef\]](#)
101. Mitsuhashi, S.; Nakagawa, S.; Ueda, M.T.; Imanishi, T.; Frith, M.C.; Mitsuhashi, H. Nanopore-based single molecule sequencing of the D4Z4 array responsible for facioscapulohumeral muscular dystrophy. *Sci. Rep. UK* **2017**, *7*. [\[CrossRef\]](#)

102. Jin, Y.; Cai, H.; Liu, J.; Lin, F.; Qi, X.; Bai, Y.; Lei, C.; Chen, H.; Lan, X. The 10 bp duplication insertion/deletion in the promoter region within paired box 7 gene is associated with growth traits in cattle. *Arch. Anim. Breed.* **2016**, *59*, 469–476. [\[CrossRef\]](#)
103. Mariasegaram, M.; Harrison, B.E.; Bolton, J.A.; Tier, B.; Henshall, J.M.; Barendse, W.; Prayaga, K.C. Fine-mapping the POLL locus in Brahman cattle yields the diagnostic marker CSAFG29. *Anim. Genet.* **2012**, *43*, 683–688. [\[CrossRef\]](#)
104. Allais-Bonnet, A.; Grohs, C.; Medugorac, I.; Krebs, S.; Djari, A.; Graf, A.; Fritz, S.; Seichter, D.; Baur, A.; Russ, I.; et al. Novel Insights into the Bovine Polled Phenotype and Horn Ontogenesis in Bovidae. *PLoS ONE* **2013**, *8*. [\[CrossRef\]](#)
105. Drogemuller, C.; Wohlke, A.; Momke, S.; Distl, O. Fine mapping of the polled locus to a 1-Mb region on bovine chromosome 1q12. *Mamm Genome* **2005**, *16*, 613–620. [\[CrossRef\]](#) [\[PubMed\]](#)
106. Long, C.R.; Gregory, K.E. Inheritance of the Horned, Scurred, and Polled Condition in Cattle. *J. Hered* **1978**, *69*, 395–400. [\[CrossRef\]](#)
107. Crystal Ketel, M.A.-C. Investigating candidate scur genes in Bos taurus breeds. *J. Anim. Sci.* **2019**, *97*, 226–227. [\[CrossRef\]](#)
108. Medugorac, I.; Seichter, D.; Graf, A.; Russ, I.; Blum, H.; Gopel, K.H.; Rothhammer, S.; Forster, M.; Krebs, S. Bovine polledness—an autosomal dominant trait with allelic heterogeneity. *PLoS ONE* **2012**, *7*, e39477. [\[CrossRef\]](#) [\[PubMed\]](#)
109. Wiedemar, N.; Tetens, J.; Jagannathan, V.; Menoud, A.; Neuenschwander, S.; Bruggmann, R.; Thaller, G.; Drogemuller, C. Independent polled mutations leading to complex gene expression differences in cattle. *PLoS ONE* **2014**, *9*, e93435. [\[CrossRef\]](#)
110. Kent, M.; Moser, M.; Boman, I.A.; Lindtveit, K.; Árnýasi, M.; Sundsaasen, K.K.; Våge, D.I. Insertion of an endogenous Jaagsiekte Sheep Retrovirus element into the BCO2-gene abolishes its function and leads to yellow discoloration of adipose tissue in Norwegian Spælsau (*Ovis aries*). *bioRxiv* **2020**. [\[CrossRef\]](#)
111. Durkin, K.; Coppieters, W.; Drögemüller, C.; Ahariz, N.; Cambisano, N.; Druet, T.; Fasquelle, C.; Haile, A.; Horin, P.; Huang, L.; et al. Serial translocation by means of circular intermediates underlies colour sidedness in cattle. *Nature* **2012**, *482*, 81–84. [\[CrossRef\]](#)
112. Awasthi Mishra, N.; Drögemüller, C.; Jagannathan, V.; Keller, I.; Wüthrich, D.; Bruggmann, R.; Beck, J.; Schütz, E.; Brenig, B.; Demmel, S.; et al. A structural variant in the 5'-flanking region of the TWIST2 gene affects melanocyte development in belted cattle. *PLoS ONE* **2017**, *12*, e0180170. [\[CrossRef\]](#)
113. Brenig, B.; Beck, J.; Floren, C.; Bornemann-Kolatzki, K.; Wiedemann, I.; Hennecke, S.; Swalve, H.; Schütz, E. Molecular genetics of coat colour variations in White Galloway and White Park cattle. *Anim. Genet.* **2013**, *44*, 450–453. [\[CrossRef\]](#)
114. Rothhammer, S.; Kunz, E.; Krebs, S.; Bitzer, F.; Hauser, A.; Zinovieva, N.; Klymiuk, N.; Medugorac, I. Remapping of the belted phenotype in cattle on BTA3 identifies a multiplication event as the candidate causal mutation. *Genet. Sel. Evol.* **2018**, *50*. [\[CrossRef\]](#)
115. Lamb, H.J.; Ross, E.M.; Nguyen, L.T.; Lyons, R.E.; Moore, S.S.; Hayes, B.J. Characterization of the poll allele in Brahman cattle using long-read Oxford Nanopore sequencing. *J. Anim. Sci.* **2020**, *98*. [\[CrossRef\]](#) [\[PubMed\]](#)
116. Zhao, L.Y.; Song, J.; Liu, Y.; Song, C.X.; Yi, C. Mapping the epigenetic modifications of DNA and RNA. *Protein Cell* **2020**, *11*, 792–808. [\[CrossRef\]](#) [\[PubMed\]](#)
117. Leger, A.; Amaral, P.P.; Pandolfini, L.; Capitanchik, C.; Capraro, F.; Barbieri, I.; Migliori, V.; Luscombe, N.M.; Enright, A.J.; Tzelepis, K.; et al. RNA modifications detection by comparative Nanopore direct RNA sequencing. *bioRxiv* **2019**. [\[CrossRef\]](#)
118. Greenberg, M.V.C.; Bourc'his, D. The diverse roles of DNA methylation in mammalian development and disease. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 590–607. [\[CrossRef\]](#) [\[PubMed\]](#)
119. Unnikrishnan, A.; Freeman, W.M.; Jackson, J.; Wren, J.D.; Porter, H.; Richardson, A. The role of DNA methylation in epigenetics of aging. *Pharmacology* **2019**, *195*, 172–185. [\[CrossRef\]](#)
120. Gentilini, D.; Mari, D.; Castaldi, D.; Remondini, D.; Ogliari, G.; Ostan, R.; Bucci, L.; Sirchia, S.M.; Tabano, S.; Cavagnini, F.; et al. Role of epigenetics in human aging and longevity: Genome-wide DNA methylation profile in centenarians and centenarians' offspring. *Age (Dordr)* **2013**, *35*, 1961–1973. [\[CrossRef\]](#)
121. Adey, A.; Shendure, J. Ultra-low-input, tagmentation-based whole-genome bisulfite sequencing. *Genome Res.* **2012**, *22*, 1139–1143. [\[CrossRef\]](#)

122. Yamaguchi, S.; Hong, K.; Liu, R.; Shen, L.; Inoue, A.; Diep, D.; Zhang, K.; Zhang, Y. Tet1 controls meiosis by regulating meiotic gene expression. *Nature* **2012**, *492*, 443–447. [\[CrossRef\]](#)
123. Dominissini, D.; Nachtergaele, S.; Moshitch-Moshkovitz, S.; Peer, E.; Kol, N.; Ben-Haim, M.S.; Dai, Q.; Di Segni, A.; Salmon-Divon, M.; Clark, W.C.; et al. The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. *Nature* **2016**, *530*, 441–446. [\[CrossRef\]](#)
124. Marioni, R.E.; Belsky, D.W.; Deary, I.J.; Wagner, W. Association of facial ageing with DNA methylation and epigenetic age predictions. *Clin. Epigenet.* **2018**, *10*, 140. [\[CrossRef\]](#)
125. Tharakan, R.; Ubaida-Mohien, C.; Moore, A.Z.; Hernandez, D.; Tanaka, T.; Ferrucci, L. Blood DNA Methylation and Aging: A Cross-Sectional Analysis and Longitudinal Validation in the InCHIANTI Study. *J. Gerontol. A Biol. Sci. Med. Sci.* **2020**, *75*, 2051–2055. [\[CrossRef\]](#) [\[PubMed\]](#)
126. Barres, R.; Zierath, J.R. DNA methylation in metabolic disorders. *Am. J. Clin. Nutr.* **2011**, *93*, 897S–900S. [\[CrossRef\]](#)
127. Wang, X.; Kadarmideen, H.N. An Epigenome-Wide DNA Methylation Map of Testis in Pigs for Study of Complex Traits. *Front. Genet.* **2019**, *10*, 405. [\[CrossRef\]](#) [\[PubMed\]](#)
128. McGilchrist, P.; Alston, C.L.; Gardner, G.E.; Thomson, K.L.; Pethick, D.W. Beef carcasses with larger eye muscle areas, lower ossification scores and improved nutrition have a lower incidence of dark cutting. *Meat Sci.* **2012**, *92*, 474–480. [\[CrossRef\]](#) [\[PubMed\]](#)
129. Ferguson, D.M.; Bruce, H.L.; Thompson, J.M.; Egan, A.F.; Perry, D.; Shorthose, W.R. Factors affecting beef palatability—Farmgate to chilled carcass. *Aust. J. Exp. Agric.* **2001**, *41*, 879–891. [\[CrossRef\]](#)
130. Johnson, A.A.; Akman, K.; Calimport, S.R.; Wuttke, D.; Stolzing, A.; de Magalhaes, J.P. The role of DNA methylation in aging, rejuvenation, and age-related disease. *Rejuvenation Res.* **2012**, *15*, 483–494. [\[CrossRef\]](#)
131. Ciccarone, F.; Tagliatesta, S.; Caiafa, P.; Zampieri, M. DNA methylation dynamics in aging: How far are we from understanding the mechanisms? *Mech. Ageing Dev.* **2018**, *174*, 3–17. [\[CrossRef\]](#)
132. Shinozaki, G.; Braun, P.R.; Hing, B.W.Q.; Ratanatharathorn, A.; Klisares, M.J.; Duncan, G.N.; Jellison, S.S.; Heinzman, J.T.; Nagahama, Y.; Close, L.; et al. Epigenetics of Delirium and Aging: Potential Role of DNA Methylation Change on Cytokine Genes in Glia and Blood Along With Aging. *Front. Aging Neurosci.* **2018**, *10*, 311. [\[CrossRef\]](#)
133. Kosten, T.A.; Huang, W.; Nielsen, D.A. Sex and litter effects on anxiety and DNA methylation levels of stress and neurotrophin genes in adolescent rats. *Dev. Psychobiol.* **2014**, *56*, 392–406. [\[CrossRef\]](#)
134. Klengel, T.; Pape, J.; Binder, E.B.; Mehta, D. The role of DNA methylation in stress-related psychiatric disorders. *Neuropharmacology* **2014**, *80*, 115–132. [\[CrossRef\]](#)
135. Meuwissen, T.H.E.; Hayes, B.J.; Goddard, M.E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **2001**, *157*, 1819–1829. [\[PubMed\]](#)
136. Georges, M.; Charlier, C.; Hayes, B. Harnessing genomic information for livestock improvement. *Nat. Rev. Genet.* **2019**, *20*, 135–156. [\[CrossRef\]](#) [\[PubMed\]](#)
137. Ganai, M.W.; Polley, A.; Graner, E.M.; Plieske, J.; Wieseke, R.; Luerssen, H.; Durstewitz, G. Large SNP arrays for genotyping in crop plants. *J. Biosci.* **2012**, *37*, 821–828. [\[CrossRef\]](#) [\[PubMed\]](#)
138. Malmberg, M.M.; Spangenberg, G.C.; Daetwyler, H.D.; Cogan, N.O.I. Assessment of low-coverage nanopore long read sequencing for SNP genotyping in doubled haploid canola (*Brassica napus* L.). *Sci. Rep.-UK* **2019**, *9*. [\[CrossRef\]](#) [\[PubMed\]](#)
139. Ros-Freixedes, R.; Gonen, S.; Gorjanc, G.; Hickey, J.M. A method for allocating low-coverage sequencing resources by targeting haplotypes rather than individuals. *Genet. Sel. Evol.* **2017**, *49*. [\[CrossRef\]](#)
140. Ros-Freixedes, R.; Whalen, A.; Gorjanc, G.; Mileham, A.J.; Hickey, J.M. Evaluation of sequencing strategies for whole-genome imputation with hybrid peeling. *Genet. Sel. Evol.* **2020**, *52*, 18. [\[CrossRef\]](#)
141. Whalen, A.; Ros-Freixedes, R.; Wilson, D.L.; Gorjanc, G.; Hickey, J.M. Hybrid peeling for fast and accurate calling, phasing, and imputation with sequence data of any coverage in pedigrees. *Genet. Sel. Evol.* **2018**, *50*, 67. [\[CrossRef\]](#)
142. Li, J.H.; Mazur, C.A.; Berisa, T.; Pickrell, J.K. Low-pass sequencing increases the power of GWAS and decreases measurement error of polygenic risk scores compared to genotyping arrays. *bioRxiv* **2020**. [\[CrossRef\]](#)
143. Zou, Y.; Mason, M.G.; Wang, Y.; Wee, E.; Turni, C.; Blackall, P.J.; Trau, M.; Botella, J.R. Correction: Nucleic acid purification from plants, animals and microbes in under 30 seconds. *PLoS Biol.* **2018**, *16*, e1002630. [\[CrossRef\]](#)

144. Xu, Y.; Yang-Turner, F.; Volk, D.; Crook, D. NanoSPC: A scalable, portable, cloud compatible viral nanopore metagenomic data processing pipeline. *Nucleic Acids Res.* **2020**, *48*, W366–W371. [\[CrossRef\]](#)
145. Grumaz, C.; Hoffmann, A.; Vainshtein, Y.; Kopp, M.; Grumaz, S.; Stevens, P.; Decker, S.O.; Weigand, M.A.; Hofer, S.; Brenner, T.; et al. Rapid Next-Generation Sequencing-Based Diagnostics of Bacteremia in Septic Patients. *J. Mol. Diagn.* **2020**, *22*, 405–418. [\[CrossRef\]](#)
146. Boykin, L.M.; Sseruwagi, P.; Alicai, T.; Ateka, E.; Mohammed, I.U.; Stanton, J.A.L.; Kayuki, C.; Mark, D.; Fute, T.; Erasto, J.; et al. Tree Lab: Portable Genomics for Early Detection of Plant Viruses and Pests in Sub-Saharan Africa. *Genes-Basel* **2019**, *10*, 632. [\[CrossRef\]](#)
147. Sanderson, N.D.; Street, T.L.; Foster, D.; Swann, J.; Atkins, B.L.; Brent, A.J.; McNally, M.A.; Oakley, S.; Taylor, A.; Peto, T.E.A.; et al. Real-time analysis of nanopore-based metagenomic sequencing from infected orthopaedic devices. *BMC Genom.* **2018**, *19*, 714. [\[CrossRef\]](#) [\[PubMed\]](#)
148. Wick, R.R.; Judd, L.M.; Holt, K.E. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol.* **2019**, *20*. [\[CrossRef\]](#) [\[PubMed\]](#)
149. Deshpande, S.V.; Reed, T.M.; Sullivan, R.E.; Kerkhof, L.J.; Beigel, K.M.; Wade, M.M. Offline Next Generation Metagenomics Sequence Analysis Using MinION Detection Software (MINDS). *Genes-Basel* **2019**, *10*, 578. [\[CrossRef\]](#) [\[PubMed\]](#)
150. Copley Pastoral Company. *Crows Nest*; Copley Pastoral Company: Winya, QLD, Australia, 2020.
151. Schaeffer, L.R. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* **2006**, *123*, 218–223. [\[CrossRef\]](#)
152. Peñagaricano, F. Chapter 6—Genetics and genomics of dairy cattle. In *Animal Agriculture*; Bazer, F.W., Lamb, G.C., Wu, G., Eds.; Academic Press: Cambridge, MA, USA, 2020; pp. 101–119.
153. Mikheyev, A.S.; Tin, M.M.Y. A first look at the Oxford Nanopore MinION sequencer. *Mol. Ecol. Resour.* **2014**, *14*, 1097–1102. [\[CrossRef\]](#)
154. Laver, T.; Harrison, J.; O'Neill, P.A.; Moore, K.; Farbos, A.; Paszkiewicz, K.; Studholme, D.J. Assessing the performance of the Oxford Nanopore Technologies MinION. *Biomol. Detect. Quantif* **2015**, *3*, 1–8. [\[CrossRef\]](#)
155. Zeng, J.; Cai, H.; Peng, H.; Wang, H.; Zhang, Y.; Akutsu, T. Causalcall: Nanopore Basecalling Using a Temporal Convolutional Network. *Front. Genet.* **2020**, *10*. [\[CrossRef\]](#)
156. Gilpatrick, T.; Lee, I.; Graham, J.E.; Raimondeau, E.; Bowen, R.; Heron, A.; Sedlazeck, F.J.; Timp, W. Targeted Nanopore Sequencing with Cas9 for studies of methylation, structural variants and mutations. *bioRxiv* **2019**. [\[CrossRef\]](#)
157. López-Girona, E.; Davy, M.W.; Albert, N.W.; Hilario, E.; Smart, M.E.M.; Kirk, C.; Thomson, S.J.; Chagné, D. CRISPR-Cas9 enrichment and long read sequencing for fine mapping in plants. *Plant Methods* **2020**, *16*, 121. [\[CrossRef\]](#) [\[PubMed\]](#)
158. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* **2011**, *6*. [\[CrossRef\]](#) [\[PubMed\]](#)
159. Romay, M.C.; Millard, M.J.; Glaubitz, J.C.; Peiffer, J.A.; Swarts, K.L.; Casstevens, T.M.; Elshire, R.J.; Acharya, C.B.; Mitchell, S.E.; Flint-Garcia, S.A.; et al. Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol.* **2013**, *14*. [\[CrossRef\]](#) [\[PubMed\]](#)
160. Crossa, J.; Beyene, Y.; Kassa, S.; Perez, P.; Hickey, J.M.; Chen, C.; de los Campos, G.; Burgueno, J.; Windhausen, V.S.; Buckler, E.; et al. Genomic Prediction in Maize Breeding Populations with Genotyping-by-Sequencing. *G3-Genes Genom Genet.* **2013**, *3*, 1903–1926. [\[CrossRef\]](#) [\[PubMed\]](#)
161. Dodds, K.G.; McEwan, J.C.; Brauning, R.; Anderson, R.M.; van Stijn, T.C.; Kristjansson, T.; Clarke, S.M. Construction of relatedness matrices using genotyping-by-sequencing data. *BMC Genom.* **2015**, *16*. [\[CrossRef\]](#)
162. Yue, J.X.; Liti, G.N. simuG: A general-purpose genome simulator. *Bioinformatics* **2019**, *35*, 4442–4444. [\[CrossRef\]](#) [\[PubMed\]](#)
163. Li, Y.; Han, R.; Bi, C.; Li, M.; Wang, S.; Gao, X. DeepSimulator: A deep simulator for Nanopore sequencing. *Bioinformatics* **2018**, *34*, 2899–2908. [\[CrossRef\]](#)
164. De Coster, W.; D'Hert, S.; Schultz, D.T.; Cruys, M.; Van Broeckhoven, C. NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics* **2018**, *34*, 2666–2669. [\[CrossRef\]](#)
165. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [\[CrossRef\]](#)

166. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **2011**, *27*, 2987–2993. [[CrossRef](#)]
167. Wang, X.; Lund, M.S.; Ma, P.; Janss, L.; Kadarmideen, H.N.; Su, G. Improving genomic predictions by correction of genotypes from genotyping by sequencing in livestock populations. *J. Anim. Sci. Biotechnol.* **2019**, *10*, 8. [[CrossRef](#)] [[PubMed](#)]
168. Maruki, T.; Lynch, M. Genotype Calling from Population-Genomic Sequencing Data. *G3: Genes|Genomes|Genet.* **2017**, *7*, 1393–1404. [[CrossRef](#)] [[PubMed](#)]
169. Martin, E.R.; Kinnamon, D.D.; Schmidt, M.A.; Powell, E.H.; Zuchner, S.; Morris, R.W. SeqEM: An adaptive genotype-calling approach for next-generation sequencing studies. *Bioinformatics* **2010**, *26*, 2803–2810. [[CrossRef](#)]
170. Dassonneville, R.; Brondum, R.F.; Druet, T.; Fritz, S.; Guillaume, F.; Guldbrandsen, B.; Lund, M.S.; Ducrocq, V.; Su, G. Effect of imputing markers from a low-density chip on the reliability of genomic breeding values in Holstein populations. *J. Dairy Sci* **2011**, *94*, 3679–3686. [[CrossRef](#)]
171. Howie, B.N.; Donnelly, P.; Marchini, J. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *PLoS Genet.* **2009**, *5*. [[CrossRef](#)]
172. Mulder, H.A.; Calus, M.P.L.; Druet, T.; Schrooten, C. Imputation of genotypes with low-density chips and its effect on reliability of direct genomic values in Dutch Holstein cattle. *J. Dairy Sci.* **2012**, *95*, 876–889. [[CrossRef](#)] [[PubMed](#)]
173. Dassonneville, R.; Fritz, S.; Ducrocq, V.; Boichard, D. Short communication: Imputation performances of 3 low-density marker panels in beef and dairy cattle. *J. Dairy Sci* **2012**, *95*, 4136–4140. [[CrossRef](#)] [[PubMed](#)]
174. Schrooten, C.; Dassonneville, R.; Ducrocq, V.; Brondum, R.F.; Lund, M.S.; Chen, J.; Liu, Z.T.; Gonzalez-Recio, O.; Pena, J.; Druet, T. Error rate for imputation from the Illumina BovineSNP50 chip to the Illumina BovineHD chip. *Genet. Sel. Evol.* **2014**, *46*. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).