*Review*

# Suitability of GWAS as a Tool to Discover SNPs Associated with Tick Resistance in Cattle: A Review

**Nelisiwe Mkize [1,2,*]** , **Azwihangwisi Maiwashe [1]**, **Kennedy Dzama [2]**, **Bekezela Dube [1]** and **Ntanganedzeni Mapholi [3]**

[1] Agricultural Research Council-Animal Production Campus, Private Bag X2, Irene 0062, South Africa; norman@arc.agric.za (A.M.); dubeb@arc.agric.za (B.D.)

[2] Department of Animal Sciences, University of Stellenbosch, Private Bag X1, Matieland, Stellenbosch 7602, South Africa; kdzama@sun.ac.za

[3] Department of Agriculture and Animal Health, University of South Africa, Private Bag X6, Florida 1710, South Africa; maphon@unisa.ac.za

* Correspondence: nemkize@outlook.com

**Abstract:** Understanding the biological mechanisms underlying tick resistance in cattle holds the potential to facilitate genetic improvement through selective breeding. Genome wide association studies (GWAS) are popular in research on unraveling genetic determinants underlying complex traits such as tick resistance. To date, various studies have been published on single nucleotide polymorphisms (SNPs) associated with tick resistance in cattle. The discovery of SNPs related to tick resistance has led to the mapping of associated candidate genes. Despite the success of these studies, information on genetic determinants associated with tick resistance in cattle is still limited. This warrants the need for more studies to be conducted. In Africa, the cost of genotyping is still relatively expensive; thus, conducting GWAS is a challenge, as the minimum number of animals recommended cannot be genotyped. These population size and genotype cost challenges may be overcome through the establishment of collaborations. Thus, the current review discusses GWAS as a tool to uncover SNPs associated with tick resistance, by focusing on the study design, association analysis, factors influencing the success of GWAS, and the progress on cattle tick resistance studies.

**Keywords:** tick control; genotyping technology; quality control; association test

## 1. Introduction

Traditionally, animal improvement programs were limited to phenotypic information only, which may be ineffective for traits that are costly to measure, such as tick resistance. In cattle production, the presence of bovine ticks is considered as one of the main sources of diseases, which are detrimental to animal health [1]. Ticks and tick-borne diseases (TTBDs) have substantial effects on animal health and welfare wellbeing, as well as a serious economic impact in both developed and developing countries [2]. To mitigate the bovine tick burden, a wide range of tick control strategies have been adopted; however, they are ineffective in completely eradicating ticks. Currently, farmers use acaricides, which were efficient when they were introduced; however, they later developed limitations which are detrimental to the animal production economy and to consumers. Various researchers have reported that the prolonged utilization of acaricides on food-producing animals creates the development of acaricide resistance by ticks, subsequently reducing the efficacy of chemicals [3,4]. Additionally, acaricide residues have been traced in milk and meat products [5]. The presence of these residues in food products and in the environment poses a health threat to human beings. This underlines the need for alternative tick control measures, which are chemically-free and environmentally-friendly [6]. A potential alternative approach to control ticks would be the use of genomic information, which entails the exploitation of genetic variation in host resistance to tick infestation. The success

of this approach depends on the discovery of genetic determinants associated with low tick load in cattle.

The development of high throughput genotyping technologies has provided an opportunity to identify novel genetic variants, such as single nucleotide polymorphisms (SNPs), associated with economic traits in cattle. SNPs are genetic markers of choice because they are heritable and abundantly distributed across the genome. Genome-wide association studies (GWAS) are increasingly becoming the common experimental approach to investigate SNP markers associated with various economic traits in animal production. This approach operates by associating the phenotype with the genotype data to investigate the causal genetic variants for traits of interest using statistical models. The use of SNP markers in breeding for tick resistance (low tick load in cattle) has the potential to assist breeders in making informed decisions to improve host resistance to ticks in cattle [7–9]. Through the use of GWAS, a number of studies have been conducted to investigate genetic variants for tick resistance in cattle, in different breeds and regions [10–14].

Studies that have been conducted to date have presented evidence of the association of various genomic regions with low tick load in cattle and recommended the validation of the discovered regions. Some of the challenges associated with GWAS include different phenotyping methods and genotyping strategies. These challenges could be overcome by the standardization of phenotyping procedures for tick count [15]. Genotype imputation has been identified and recommended as a cost-effective approach to account for the missingness of genotyped data and facilitate the improvement of GWAS power. Additionally, the establishment of collaborations holds the potential to solve issues associated with small discovery populations and running costs for GWAS for tick resistance.

Despite the number of tick resistance GWAS studies that have been conducted to date, the availability of data is still a challenge, which is a global challenge hindering the success of improving tick resistance traits through genomic selection. Several researchers noted that in depth knowledge of genome variation for tick resistance in cattle is required [7,16]. The generation of genomic information for host resistance to ticks is currently gaining more attention because there is a need for the knowledge on genetic determinants influencing this trait. This paper reviews GWAS as a genetic tool to identify genetic variants associated with resistance to ticks in cattle.

## 2. GWAS Overview

The advent of high density SNP genotyping platforms has provided opportunities to detect quantitative trait loci (QTLs) and uncover the genetic architecture of quantitative traits. This development has stimulated interest among researchers to explore genetic variabilities associated with various diseases affecting animals using GWAS. Genome-wide association analysis relies on recombination to rearrange the genome. Its underlying principle is to seek correlation between phenotype and genotype based on a non-random association of alleles at two or more loci [17,18].

This GWAS approach has successfully uncovered genetic determinants associated with disease susceptibility and resistance in humans, animals, and plants [19–21] However, using this approach to uncover genetic determinants associated with traits which follow polygenic patterns of inheritance, such as tick resistance, is not straight forward, since such traits are controlled by multiple genes. Despite this challenge, the approach is used to search for marker variants indirectly associated with certain diseases or traits of interest by assuming that a marker is in linkage disequilibrium (LD) with the underlying causal variant [22]. Linkage disequilibrium refers to the non-random association of alleles at different loci in a given population [23]. Currently, the GWAS approach is gaining popularity in mapping QTLs associated with traits of economic importance or complex traits. This is because GWAS is able to detect variants that can be in LD with the causal variant, and this information could be used to narrow genomic regions that harbour causal variants [24,25], providing genetic determinant information that could be useful for the genetic selection of economic traits, such as tick resistance in cattle. The continued

success of GWAS depends on careful population selection and collaborative analytical approaches. Work by [26] reviewed the guidelines for successful GWAS analysis intensively and presented a useful workflow, which might be of value when conducting GWAS.

Bovine GWAS have successfully discovered genetic determinants associated with distinctive disease resistance or susceptibility, such as tuberculosis [27], resistance to ticks [8,12], mastitis [28]; and foot and mouth disease [29]. Moreover, this approach has been used in the successful mapping of genetic variants associated with meat quality [30,31] and milk production [32,33]. Such studies assist in providing information on the genetic architecture of QTL, generate biological knowledge about the expression of economic traits and facilitate the improvement of genomic selection.

### 3. Computer Software for GWAS and Genomic Public Databases

The most commonly used computer programs for GWAS are presented in Table 1. They perform the same activities and their access is generally free. The effective use of these programs requires the user to have operating skills.

**Table 1.** Common publicly available computer programs for GWAS.

| Software | Focus | Website | Reference |
|---|---|---|---|
| PLINK | Stratification, LD and structured association mapping | http://pngu.mgh.harvard.edu/purcell/plink, accessed on 20 May 2019 | [34] |
| R (GenABEL) | Stratification, LD and structured association mapping | https://cran.r-project.org/src/contrib/Archive/GenABEL, accessed on 20 May 2019 | [35] |
| SVS | Stratification, LD, haplotype blocs and structured association mapping | http://www.goldenhelix.com, accessed on 1 June 2019 | [36] |
| GenAMap | Stratification, LD and structured association mapping | http://cogito-b.ml.cmu.edu/genamap, accessed on 1 June 2019 | [37] |
| GEMMA | Stratification, Fits LMM and BSLM models, IBD analysis, estimation of chip heritability, and association mapping. | http://www.xzlab.org/software.html, accessed on 1 June 2019 | [38] |
| Blupf90 | Data conditioning, estimate variances using several methods, and use SNP information for improved accuracy of breeding values + for genome-wide association studies (GWAS) | http://nce.ads.uga.edu/wiki/doku.php?id=documentation, accessed on 17 November 2021 | [39] |

GEMMA—genome-wide efficient mixed model association; LLM—linear mixed model; BSLM- Bayesian sparse linear mixed model; SNPs—single nucleotide polymorphisms; LD—linkage disequilibrium; IBD—identical by descent; SVS—SNP and variation suite.

The information generated from genomic studies is housed on different web databases for public access. Such databases include NCBI, EMBL-EBI, Ensembl, Animal QTLdb and NAGRP. Some of these databases are not specific to any organism, while some are specific to livestock genomics (Table 2). To date, a number of QTLs and associations related to tick resistance have been identified in different cattle breeds using different research approaches. The information of the QTLs and associations is recorded in a database known as Animal QTLdb, which is currently updated whenever there is an availability of new information [40]. Based on the available information, it is noted that most of the QTLs and associations have been uncovered mainly on chromosome 10, followed by BTA23, BTA14, BTA11, BTA2 and others. This is depicted in Figure 1, sourced from the Animals QTLdb database. Information on African indigenous breeds is lacking on these databases. The limited information contained on these databases concerning tick resistance in cattle breeds shows that more research still needs to be conducted, especially in African cattle breeds.

**Table 2.** Some web databases that house genomic information associated with economic traits.

| Genomic Database | Description | URL |
|---|---|---|
| NCBI (Genbank) | Repository for biomedical and genomic information | https://www.ncbi.nlm.nih.gov/ accessed on 6 December 2019 |
| Ensembel | Genome browser | https://www.ensembl.org/index.html accessed on 6 December 2019 |
| Animal QTLdb | Animal QTL database | https://www.animalgenome.org/cgi-bin/QTLdb/index accessed on 4 September 2020 |
| NAGRP | Genomic information browser | https://www.animalgenome.org/ accessed on 4 June 2020 |
| EMBL-EBI | Genomic information database | https://www.ebi.ac.uk/ accessed on 15 June 2020 |
| DDBJ | Genomic information browser | https://www.ddbj.nig.ac.jp/index-e.html accessed on 15 June 2020 |
| UCSC | Genome browser | https://genome.ucsc.edu/ accessed on 15 June 2020 |
| Refseq | Reference sequence database | https://www.ncbi.nlm.nih.gov/refseq/ accessed on 15 June 2020 |
| VEGA | Genome browser | http://vega.archive.ensembl.org/index.html accessed on 15 June 2020 |

Animal QTLdb—Animal quantitative trait loci database; NAGRP—national animal genome research program; EMBI-EBI—European molecular biology laboratory-European bioinformatics institute; DDBJ—DNA data bank of Japan; UCSC—University of California Santa Cruz; Refseq—Reference sequence; VEGA—vertebra genome annotation.
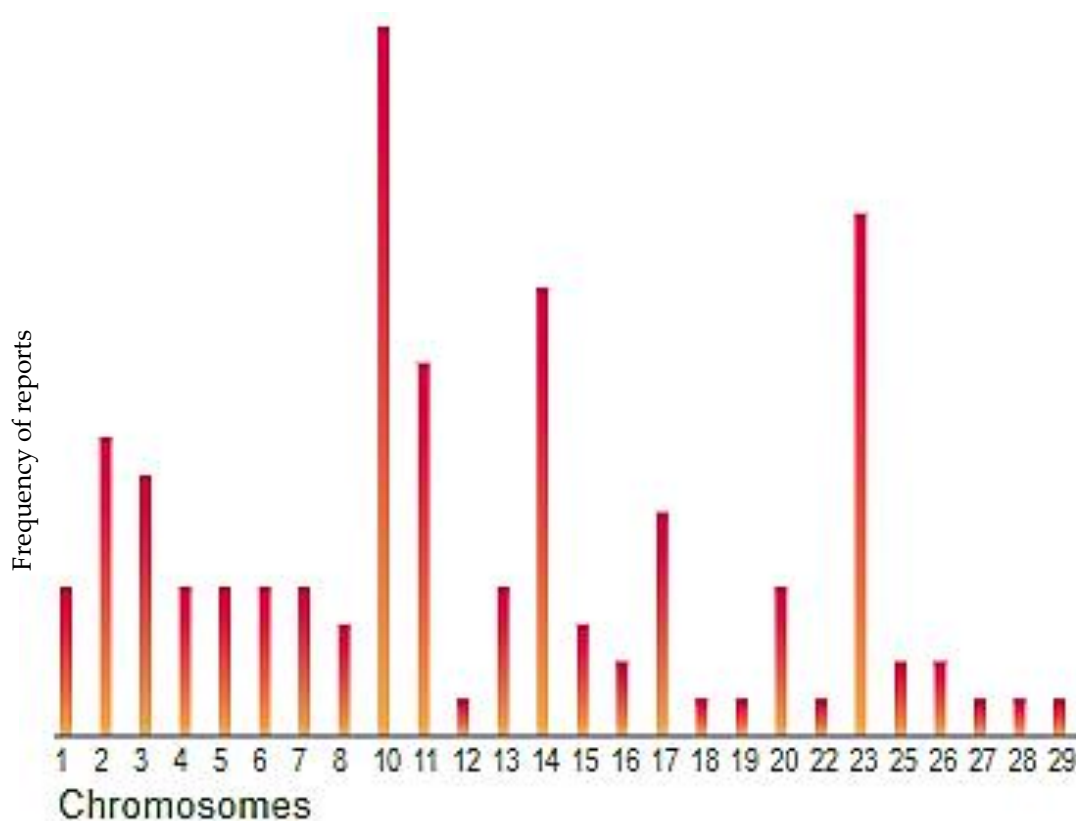


**Figure 1.** Distribution of tick resistance related QTL/associations in bovine, based on count of report data (sourced from: https://www.animalgenome.org/cgi-bin/QTLdb/BT/index (accessed on 4 September 2020).

## 4. Available Genotyping Platforms and Coverage

The abundance of SNPs in the genome and its ability to be amenable to high throughput automated analysis make SNP genotyping the most preferred approach to studying genetic variation in animals, humans, and plants [41–43]. Moreover, SNPs are heritable and allow single base resolution, making the identification of causal markers easy [44].

Initially, SNP genotyping arrays were developed for human studies [45], and then, the technology was adopted in animal and plant research. In cattle research, three well known commercial companies that produce commercial Bovine SNP arrays are Illumina, Neo-Geneseek® and Affymetrix. These companies have developed SNP genotyping platforms with different densities, which are used for GWAS, identification of selection sweeps, and investigating genome-wide genetic diversity and relationships in cattle. The summary of SNP genotyping platforms available for Bovine is provided by [46]. The SNP genotyping platforms include different densities of SNPs, ranging from as low as Golden Gate Bovine 3K (2900 SNPs) to Bovine HD (777,962 SNPs) [47]. The density of an array plays an important role in the success of a GWAS study. It was recommended by [48] that denser SNP arrays should be used for crossbreed GWAS studies, due to its large hypothetical effective population size. GWAS experiments performed using data from denser SNP arrays are provided with enough marker density to dissect the genetic architecture of the trait of interest. In instances where low density SNP arrays have been used, genotype imputation is advisable. This phenomenon will be briefly discussed later in this article.

The development of SNP arrays is advancing rapidly, making genomic selection feasible [47]. They noted that the development of genotyping platforms has not been regulated under a standardized system. This results in difficulties in making comparisons and merging data genotyped by different commercial companies. Therefore, there is need for standardization during SNP array development to minimize downstream research challenges [47]. Despite this success, studies on local African breeds are still facing the challenge of the SNP arrays being originally created using exotic breeds. This underscores the need for the development of SNP arrays that will incorporate ancestry data from African breeds. The development of such SNP arrays could potentially extend insights on the architecture of traits being investigated using African cattle populations. Although the decreasing cost of genotyping potentially makes GWAS possible and affordable, in Africa the cost of genotyping is still relatively expensive. This is a challenge for GWAS because it is difficult to genotype the minimum number of animals recommended [48]. The minimum number of animals required is determined by conducting a statistical power test, to ensure the low rate of discovering false positive results in GWAS [49]. The issue of attaining a better number of samples for GWASs can be solved through the establishment of collaborations.

## 5. Testing for an Association

The main goal for GWAS is to test a null hypothesis, stating that there is no association between the genotype and the expression of the trait of interest. This is facilitated by choosing the appropriate association test approach, generally influenced by covariates, population structure, study population and pedigree structure [50].A single locus statistical test and multiple locus tests are the two approaches that are currently being used, depending on the focus of the study. A single locus statistical test compares the genotype and the phenotype by focusing on one SNP at a time [51]. This test uses a regression model, assuming that the trait being studied is normally distributed and the variance is the same within a population. Testing one SNP at a time results in multiple tests, which may produce false positive and false negative results [48,52,53]. Therefore, it is necessary to correct for multiple testing to prevent spurious associations, making it impossible to conduct follow up studies. This can be performed using the false discovery rate (FDR) and Bonferroni correction. The adjustment through the FDR approach corrects for the expected rate of false discoveries, and also gives the investigator an insight into the proportion of true associations in the study [54,55] However, this adjustment is considered less stringent compared to the Bonferroni correction [55,56].

After the adjustment, a SNP is considered statistically significant if its p value is less than or equals the adjusted genome-wide cut off [55]. A single SNP association approach worsens the missing heritability problem, which is the gap between the heritability measured using pedigree information and that measured through GWAS [57,58]. Biological

mechanisms, such as epistasis, epigenetics and others, are attributors of missing heritability. However, in a human-based study, none of these mechanisms accounted for the missing heritability [59] The study included human microbiome information to understand the heritability of a given trait in humans. They reported that microbiome is associated with many important traits and encodes for extra genes which interact with human genes. The interaction can be a source of variation and phenotypic plasticity [59]. Thus, the inclusion of microbiome in GWAS for cattle can be used to solve the issue of missing heritability. Alternatively, the multilocus association approach can be adopted, because it examines nonlinear relationships genome-wide. Multilocus models do not require Bonferroni correction; this is beneficial because it reduces the high chances of losing many loci associated with the targeted trait through failing to meet the stringent requirement for the significant test, as it happens in single locus models [60]. Multilocus models are more advantageous compared to single locus models, because of their ability to allow the estimation of three variance components, a high power of QTL detection, and the SNP effect is random [61]. Table 3 shows some examples of single and multilocus models that are currently being used in GWAS across different species.

**Table 3.** Models that can be used for GWAS analysis.

| Model Type | Model | Reference |
|---|---|---|
| Single locus | General linear model(GLM) | [62] |
| | Mixed lieanr model (MLM) | [63] |
| | Logistic mixed model(LMM) | [64] |
| | Compressed mixed linear model (CMLM) | [63] |
| Multi-locus | Multilocus random SNP effect mixed linear models(mrMLM) | [65,66] |
| | Fast multilocus random SNP effect effiecient mixed model association (FASTmrEMMA) | [67] |

During the association analysis, population stratification is regarded as one of the confounding effects that can inflate the variance of the usual statistics [65]. The inflation of test statistics may potentially attribute high false positive discoveries; hence, it should be accounted for. There are two ways to account for population stratification in GWAS: the variance is adjusted using genetic control or principal component analysis. It was pointed out by [65] that genomic control adjusts the variance by calculating the statistics on data from null loci. Figure 2 shows the quartile–quantile plots before correction (A) and after (B). Figure 2A shows a departure from the diagonal observed, which indicates a high inflation rate. Figure 2B shows the improvement after correction using genomic control, where the inflation rate has decreased.

At the end of the association analysis, an independent association test replication is recommended on the SNPs to validate the findings [66]. This is, however, still not applicable for most studies, because of the associated cost implications, time, and other factors associated with the study design. When a replication study is conducted, GWAS is performed on statistically significant SNPs. This process makes use of a small sample size, attributing to a low GWAS power and making it difficult to confirm the initial findings [68]. Lack of replications does not necessarily mean findings are not valid, but the study should be properly designed, and all confounding factors accounted for, to ensure the validity of the findings.
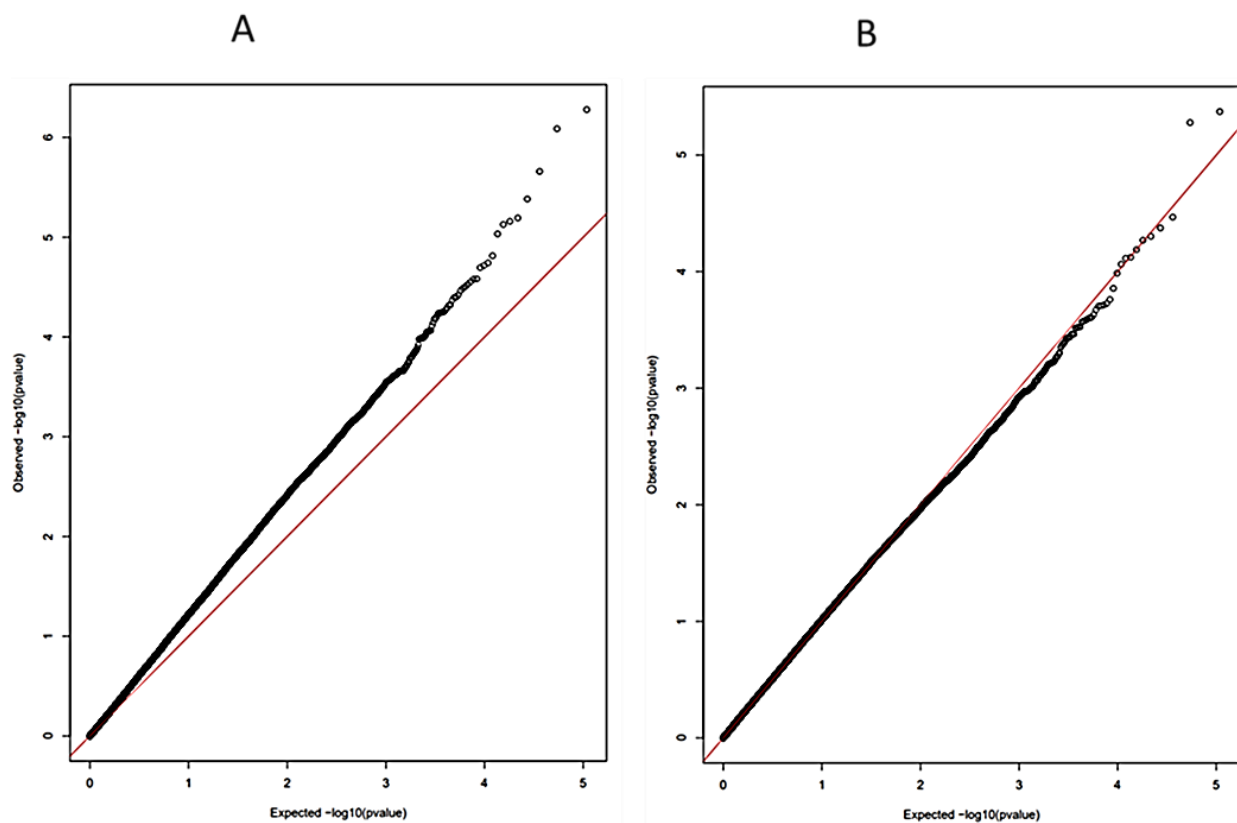
**Figure 2.** (**A**) Quartile–quantile plots before correction. (**B**) Quartile–quantile plots after correction.

The preferable mode to summarize and present GWAS findings visually is the use of a Manhattan plot (Figure 3). Figure 3 presents y-axis $-\log_{10}$ (*p*-value) versus x-axis (the chromosome position for each SNP tested), where each circle signifies a SNP. Generally, the SNPs are stacked together to form a signal that is influenced by the level of LD amongst the SNPs with relation to the causal marker. Several studies have used Manhattan plots to present their GWAS findings in cattle [8,14]. In South Africa, [8] conducted a GWAS study on the indigenous Nguni and produced the plots presented in Figure 3. The two dotted horizontal lines depict the suggestive (red) and the actual genome-wide cut off line [grey], which represents the level of significance. If a SNP passes the grey line, it is considered significant. On this study, a significant SNP associated with tick resistance in Nguni cattle was observed in BTA 10. There are many free tools that can be used to graphically plot a GWAS Manhattan diagram; examples include SNPEVG [69], R package (qqman) [70], Stata [71], and Manhattan ++ [72]. When a GWAS association has been detected, tools such as Manhattan harvester [73] and Locus Zoom [74] provide opportunities to study the detected region in depth. Both tools focus on the physical position of the chromosome of interest; however, LocusZoom is more informative because it allows the visualization of LD levels, recombination rates and genes [74]. The LocusZoom tool has mostly been used in human research. Figure 4 depicts an example of a LocusZoom plot, sourced from a human GWAS study [75], showing the architecture of a region of interest in chromosome 19 (significant SNPs, associated recombination rate, and genes). The use of tools such as LocusZoom in bovine research will provide an in-depth insight into the landscape of genetic contribution associated with the expression of economic traits, such as tick resistance, growth traits, milk production and others.
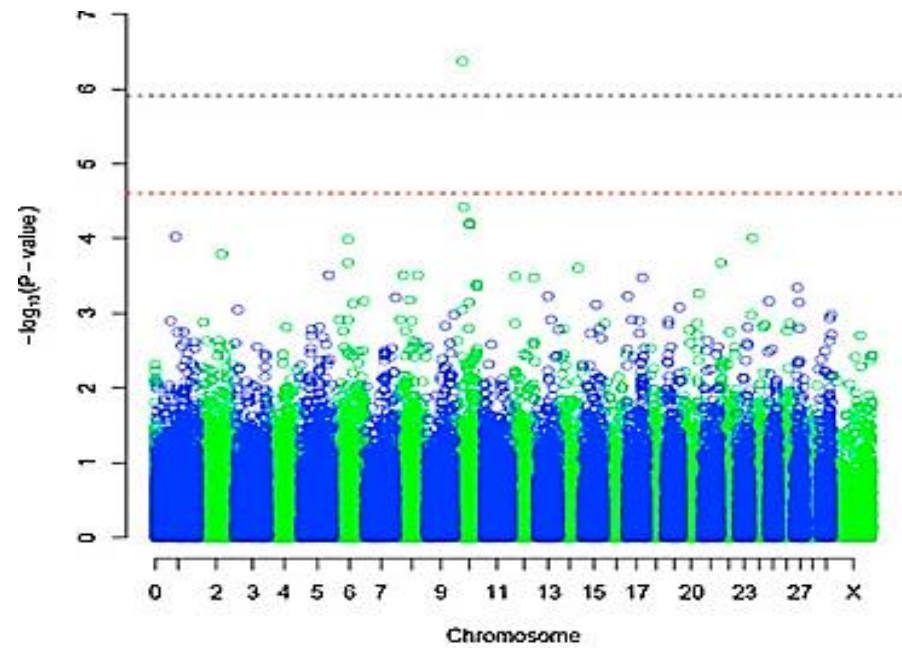
**Figure 3.** Manhattan plot showing findings for a single marker GWAS where the association of low tick load (total *A. hebraeum* ticks) and genotype was assessed in Nguni breed, using a genome-wide *p* value < 0.05 as a cut-off. The redline indicates suggestive threshold and the grey indicates the genome-wide cut off (taken from [8]).
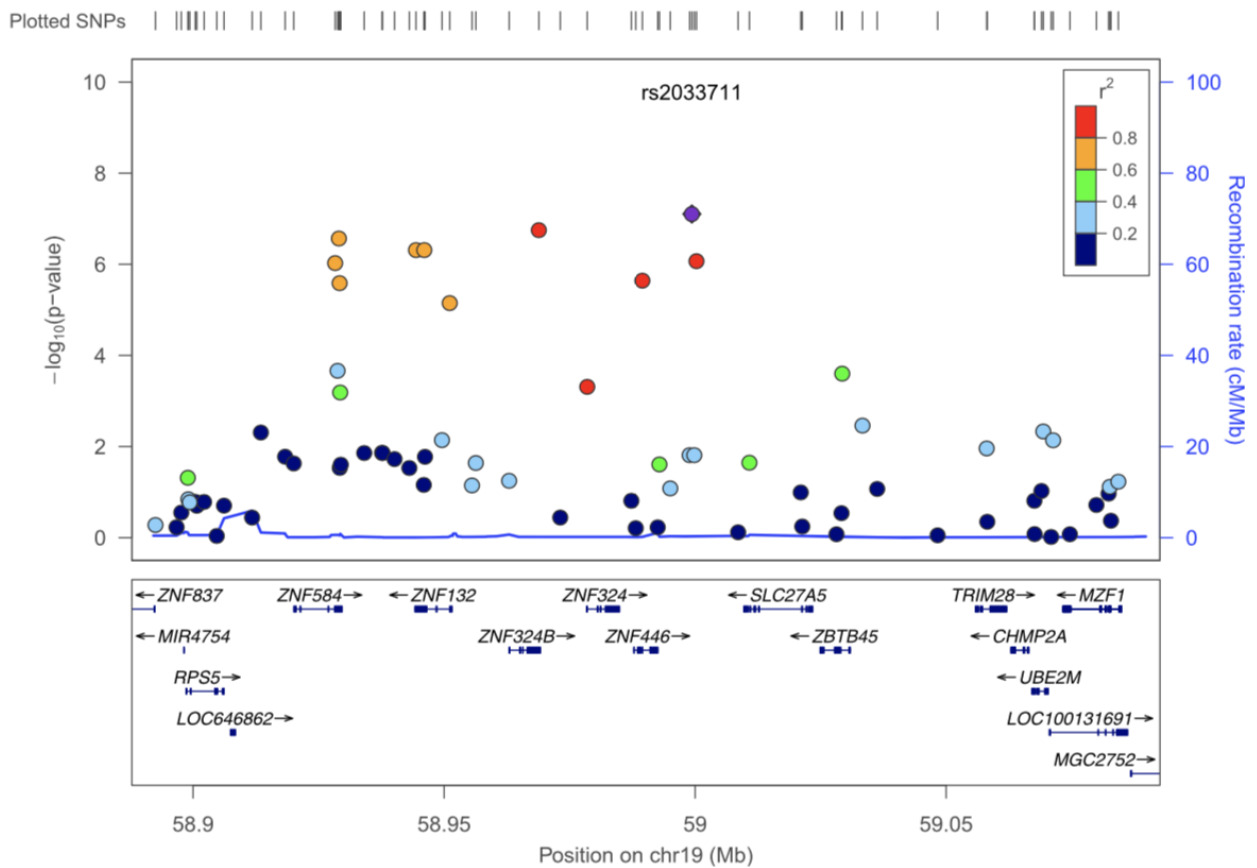


**Figure 4.** LocusZoom plot showing in depth findings for most significant SNPs in Chromosome 19, from a human GWAS that was focused to study genetic variation underlying renal uric acid excretion in Hispanic Children [75].

## 6. Post GWAS Analysis

The discovered GWAS hits are used to physically map candidate genes underlying the trait being studied. The National Centre of Biotechnology Information (https://www.ncbi.nlm.nih.gov/ accessed on 6 December 2019) and ENSEMBL(http://www.ensembl.org/index.html accessed on 6 December 2019) databases are generally used for gene annotation to identify candidate genes associated with identified SNP markers [76]. When the genes have been identified, their biological functionality or relevance can be verified through the use of functional annotation databases such as DAVID and KEGG [58,77,78]. Furthermore, it is also possible to create gene networks using open source software such as Cytoscape [79]. The gene network provides a better biological understanding of the interaction of genes underlying the trait of interest. A recent GWAS by [14] identified SNPs and candidate genes (*TREM1, TREM2, CD83, MYO5A, TREML1, PRSS16*) associated with tick resistance, and then used the information to create a gene network. Tick resistance is a complex trait that is influenced by various determinants. Therefore, gene networks are very important to give insight on the interconnection of genes responsible for the expression of the tick resistance trait. Despite the success that has been made, the inferring of true causal genes and biological mechanisms from GWAS results in tick resistance studies is still a challenge. This is due to the difficulties associated with the interpretation of GWAS findings and the limitation of available data.

## 7. Factors Influencing the Success of GWAS

### 7.1. GWAS Experimental Design

The success of GWAS requires proper study experimental design, in addition to factors such as population of interest, sample size and standardized data collection. In addition, concise pipelines to execute the actual analysis should be taken into consideration. Phenotyping data should be properly collected to reduce high rates of outliers, which may potentially create noisy data. The data is often tested for normality, to assess the distribution and the presence of outliers so that necessary steps can be taken to address the violation of normal distribution assumptions and the removal of outliers. Generally, the tick resistance trait is known to be not normally distributed and literature shows that some studies addressed this issue by transforming the data to confer normality. It has been noted that, although removal of outliers is important, it affects the size of the population tested. Selection of population for association analysis generates a structure that leads to specific genetic variation and an effect on the end use of association analysis [49]. In cattle, within family pure breeds and crossbreeds have been used for dairy and beef GWAS experiments to identify genomic regions associated with phenotypic variation in economically important traits. Crossbred lines, specifically the F2 design, have been identified as an appropriate design because they exhibit a high level of LD compared to pure lines. High LD potentially increases the power of GWAS [80]. Studies have been conducted to study the LD patterns within the crossbred F2 population [48,81,82].

In relation to cattle tick resistance studies, GWAS studies have been conducted using the crossbred F2 design in different regions [12,14]. However, it can be noted that some previous studies were conducted using data that was genotyped using microsatellite markers. The limitation of microsatellite markers involves the lack of an adequate number of informative markers, while on the other hand, SNPs are very abundant in the genome [83]. A study by [14] is currently the only known study that used the F2 crossbred design to discover tick resistance genetic determinants using SNP genotyped data. Although all these studies have been conducted, and information generated, the development of crossbred F2 populations remains a challenge. More time is required to build the required population, and this is associated with high costs of SNP genotyping a large sample size, especially in developing countries. However, there is need for this population design to be used for the genomic improvement of the tick resistance trait, regardless of the associated shortcomings.

### 7.2. Phenotyping

Phenotyping for tick resistance in cattle is often assessed through natural or artificial infestation. The assessment of tick count under the approach of interest is one of the baseline indicators for tick resistance differences. Tick count is the commonly used mode of phenotyping in studies that assess tick resistance differences in a population. Counting ticks for GWAS phenotyping where a large population size is the requirement becomes a bottleneck. A study by [84] noted that counting ticks on every animal in a population is labor intensive, time consuming, and requires trained technicians and expensive infrastructure to constrain the animal. This stresses the animals through handling and the use of a trained technician does not rule out the possibilities of bias tendencies when counting. More challenges are faced under natural infestation because of multiple tick species and the need to categorize them accordingly when counting. All these shortcomings play a part in the reduced success rate of recording a tick count for big study populations. A study by [85] on Nguni cattle estimated the correlation between whole body tick count and tick count in different body parts. They also assessed correlations between tick species and different seasons, and observed that some tick species were prevalent in certain seasons, indicating that seasons are influential on tick distribution. The approach for this study is very informative because it gave insight into the distribution of ticks in different parts of the animals. The challenge is that the process of counting ticks is labor intensive, and this will be very difficult for studies with large population size.

Artificial tick infestation has been used an alternative approach; however, this approach does not represent the true reflection of natural infestation in regions where various tick species inhabit [84]. To solve this problem, ref. [86] introduced a scoring method that overcame some disadvantages associated with tick count. However, ref. [87] reported that the scoring method provides a low heritability, as compared to tick count. Low heritability constitutes a reduced power to detect association, since heritability is a good indicator that explains the genetic variance contribution towards the expressed phenotype [49]. Therefore, since the scoring method is associated with low heritability estimates, tick count remains a better method of phenotyping for tick resistance GWASs. However, there is need for high throughput phenotyping methods that will eliminate the issues of biasness when counting ticks and reduce associated costs and stress imposed to animals.

### 7.3. Population Size

Population size is amongst the limiting factors affecting the statistical power of a GWAS. A small population size results in a reduced statistical power to detect a causal variant. This poses a major challenge in understanding the biological mechanism underlying economic traits such as tick resistance. Generally, a population comprising 1000 individuals is regarded as a better population size, constituting about 80% GWAS statistical power [88]. In certain research set ups, it is difficult to attain a minimum of 1000 samples to conduct a GWAS, because of various factors at play associated with infrastructure and costs. These factors are not limiting researchers' drive to find information that could be used to improve animal welfare and facilitate sustainable food security in the current trying times. According to [89], a population of individuals ranging from 100 to 500 can be suitably used to perform GWAS. However, such a population needs to be carefully selected based on the status of the genetic variation of the trait being studied, and other factors influencing phenotype such as environment. Moreover, the selection criterion must take population stratification into consideration to avoid false positive results. The selected GWAS population should result in low stratification [90] and large genetic variation to potentially detect true variants that can be used in breeding schemes to improve the trait of interest.

GWAS statistical power is defined as the probability of rejecting a null hypothesis under an assumed alternative hypothesis [88]. It depends, mainly, on a clear study design and the number of samples used for the analysis. Generally, one thousand samples, equating to 80% power, are considered the normal number in human and animal studies [88]. The same applies to livestock studies. However, the issue of sample numbers remains a

problem in developing countries. Hence, many GWAS livestock studies are conducted on samples below one thousand. Factors playing a role in this regard include the lack of recorded phenotypes for certain traits and the availability of resources to cover genotyping costs. Although such costs are gradually decreasing, researchers in Africa are still battling to afford genotyping on cattle numbers near the normal one thousand mark. The situation is even more difficult when high density SNP genotyping chips are considered. GWAS analysis requires a large sample size to achieve sufficient statistical power [91,92]. Attaining a large population size for tick resistance studies is a challenge in developing countries, because of lack of funds, resources, well trained data collectors and infrastructure. To tackle this problem, the establishment of collaborations could be beneficial, since it will promote the sharing of information, insights and funds. This will improve the development of GWAS and its quality.

### 7.4. Data Quality Control for GWAS

Prior to actual genome-wide association analysis, the genotyped data is subjected to quality control (QC) to decrease the chances of discovering false positive and false negative associations [93]. False positive association is defined as the occurrence of identifying a SNP association that is not truly influencing the trait of interest in the study. On the other hand, false negative association is defined as an incidence where a SNP that is influencing the trait of interest is not associated with the trait in the study. The genotyped data is subjected to stringent filters that are performed on a sample and SNP level. The quality control filters include missing call rate, minor allele frequency (MAF) and Hardy Weinberg equilibrium (HWE). The samples with a missing call rate higher than 1–5% could be a result of poor DNA quality potentially influencing genotypic errors [56]. The removal of SNPs with high missing genotypes may increase SNPs with accurate genotype calls for downstream analysis. However, in a study with a small population size, it is not good practice to lose samples because this negatively affects the power of the study. This can be overcome through genotype imputation to replace the missing SNP markers, explained later.

The other aspect of QC is to assess the format of genotyped data, where sometimes there is a mixing of the AB and ACTG formats, which needs to be corrected to form one uniform format. Additionally, data is assessed for MAF, which removes SNPs not complying with a given threshold for a particular study. Very low allele frequencies are less informative [94] and can result in the discovery of fake associations [95]. GWAS capitalizes on the LD that exists between the markers; thus, it is very important to assess the deviation of SNPs from the HWE. The deviation of SNPs from HWE is set using Chi square test [96], where the significance level for rejecting is based on P values ranging from $10^{-5}$–$10^{-7}$ [50,56]. The SNPs that are not in compliance with a stipulated criterion are removed from the dataset that is used in the downstream analysis. Overall, statistical software such as PLINK, and R environment are used for quality control. However, PLINK is the most preferred because it is free and flexible to accommodate large scale data management [56]. It is noted there is no universal criterion to perform QC and there is no perfect QC pipeline that is able to capture all the problematic SNPs in a population being studied. Therefore, it is important to view the clustering intensity plots for SNPs, to ensure that there are no obvious clustering problems.

### 7.5. The Extent of LD Measures $r^2$ in GWAS

GWAS relies on the extent of LD between markers across the genome, where its measure ($r^2$) is bounded between 0 and 1, with 1 considered as the perfect association [48]. Linkage disequilibrium in a population can be affected by population structure, genetic drift, selection, recombination rate, migration and mutations [97]. The development of LD maps and haplotype block structures at the population level are useful parameters for guiding GWAS. In association studies, the presence of LD creates two possible outcomes, namely, direct and indirect association [98]. Direct association occurs when the SNP influencing a biological system is directly genotyped in the study and found to be

statistically associated with the trait. The indirect association outcome occurs when the SNP is indirectly linked with the trait. The phenomenon is termed taq SNP and is graphically explained well in a review by [98]. The feasibility of GWAS strongly depends on the extent of LD, as the latter determines the required SNP markers and mapping resolution [99]. Therefore, it is important to study the extent of LD in a population of interest before the association analysis is performed.

### 7.6. The Effect of Genotype−Environment Interaction

According to [100], genotype by environment (GxE) interaction is defined as the different responses of genotypes under different environmental conditions. GxE affects the ranking of animal performance under different environmental conditions. Therefore, it is important to closely monitor economical traits that are largely influenced by environmental factors [101]. In cattle production, tick resistance is among the traits that are highly influenced by environmental factors. Climate change influences the distribution and density of tick populations and, because it has adverse influences on the life cycle of ticks, it increases the chances of tick−host interactions.

Currently, genetic evaluations for tick tolerance in countries such as Australia and Brazil are performed routinely. Despite the success of these evaluations, ref. [102] noted that GxE is not taken into consideration in these evaluations. They also pointed out that failure to consider GxE interaction in genetic evaluations can potentially affect genetic gain, as the selection of candidate comparisons is environmentally dependent. Thus, the cost implications of not accounting for GxE can be high [103]. This is because animals observed as top performers in one environment will not necessarily perform the same in a different environment, a phenomenon that is associated with the loss of genetic progress. Few studies have provided evidence that resistance to ticks in cattle can be influenced by various environmental effects [104]. A recent study investigated the existence of GxE using different models in Hereford and Braford cattle [102]. Their findings showed that the estimates of repeatability varied along the environmental gradient (range 0.18–0.45), indicating that resistance to ticks is environmentally influenced. Additionally, the posterior means of the genetic correlations across the environmental tick infestation surface plot demonstrated a large plateau above 0.80. This indicates that there will be re-ranking of performance for a trait of interest between environments, which necessitates the separation of breeding programs for each environment [105,106]. GxE interaction contributes to the genetic architecture of complex traits and it affects the chances of discovering a true association between phenotype and genotype in GWASs. According to [107], failure to adjust for environmental effects results in the reduced chances of predicting an association. This has been proven through assimilation studies, intensively in human epidemiological studies [108–110] and plant studies [111,112].

### 7.7. Batch Effect

For most study designs, samples are not genotyped at once, instead, they are handled in batches. One of the reasons behind this can be the use of large number of samples, which makes it impossible to genotype the samples at the same time. In addition, some studies collect samples at different time intervals, prompting the genotyping to be also conducted at different time intervals. For cattle tick studies, sometimes data is collected from different environments that are geographically spaced, leading to data being treated in batches. Additionally, for cattle ticks studies conducted on hybrids, the data is partitioned according to the development of the hybrids. It is known that the development of hybrid populations can never be carried out at one go. Instead it is conducted in batches, for example, the development of an F2 population. The partitioning of samples gives rise to the batch effect, which results in apparent associations confounded by the batches. It is therefore necessary to assess the dataset for a potential batch effect, since it has potential to yield spurious association if this is not accounted for. Before the actual GWAS association test, the batches'

genotyped data should be handled independently to assess the presence of confounding effects [113].

*7.8. Genotype Imputation as a Cost Effective Approach to Improve the Power of GWASs*

The main purpose of imputation is to infer missing genotypes of the SNPs that are not directly genotyped in the study, using *in silico* haplotype information from reference samples with genotypes from denser genotyping arrays [114,115]. Genotype imputation holds the potential to improve the statistical power to detect association by reducing the number of missing genotypes, thereby increasing the overall number of genotypes available for association analysis [114]. Genotype imputation has the potential to boost GWAS statistical power by up to 10% over testing only genotyped SNPs [116]. The performance of association tests on typed SNPs may not lead to a significant association, especially when the sample size is small [114]. Their findings were different when the association test was conducted on imputed genotyped data, where there was also a more detailed view of the association region.

Various tools are available for imputation and some are presented in Table 4. For example, BEAGLE has been built to handle genotype intensity data so that genotypes can be called using LD information between the SNPs, offering an improvement in genotyping error rates. Imputation is a cost effective tool to generate genomic variants at denser platforms [46,117]. Imputation allows high density genotypes to be imputed reliably from low density SNP arrays [117]. This potentially solves the affordability issue in developing countries, as more animals can be genotyped at low cost. However, the accuracy of imputation and the factors that affect it should be taken into consideration because they determine the reliability of the tool.

**Table 4.** Some available software packages for genotype imputation.

| Software | Usage | Website |
|---|---|---|
| BEAGLE | Prephases haplotypes infers missing genotypes, and identifies IBD in related samples | https://Faculty.washington.edu/browning/beagle/old.beagle.html accessed on 9 July 2020 |
| GIGI | Imputes missing genotypes on a pedigree | https://faculty.washington.edu/wijsman/progdists/gigi/software/GIGI/GIGI.html accessed on 9 July 2020 |
| IMPUTE2 | Prephases haplotypes, infers missing genotypes | https://mathgen.stats.ox.ac.uk/impute/impute_v2.html accessed on 9 July 2020 |
| MaCH/minimac3 | Prephases haplotypes, infers missing genotypes | https://github.com/statgen/Minimac4 accessed on 9 July 2020 |

IBD—identical by descent; GIGI—Genotype imputation given inheritance.

In humans, whole genome sequencing and imputation based GWAS strategies were used to refine the association signals and recover novel association signals for complex traits [118,119]. Sequencing and imputation GWAS is powerful and cost effective, and can also be applied on non-European populations [120]. In cattle, the whole genome sequencing and imputation GWAS strategies have been applied to study the genetic architecture of quantitative traits in beef cattle [50]. In relation to tick resistance, most GWAS were carried out using 50K genotyping platforms (Table 5). Thus, some studies used imputation to ensure better statistical analysis effectiveness [13,14,121]. Imputation boosted the number of SNPs that were tested for association and subsequently improved the power. Imputation makes it applicable for cattle tick resistance-focused studies to be conducted using samples that were initially genotyped at low density platforms, then imputed to high density. This has been proposed as a cost-effective approach than can solve the current problems associated with generating genomic data in cattle production, especially in developing countries. Currently, in different regions, there are ongoing studies that are investigating the feasibility of the tool for cattle research. The outcomes of these studies will give

insights that could be used to properly apply this tool in GWAS aimed to study biological mechanisms underlying tick resistance.

## 8. Progress on Tick Resistance GWAS in Cattle

Quantitative trait loci studies using microsatellites and SNPs have been inconsistent, with a very low percentage relating phenotypic variation to tick infestation [122]. Most studies were conducted in subtropical regions such as Brazil, Australia and Mexico (Table 5). Brazilian studies have successfully mapped genomic regions associated with resistance to ticks on $F_2$ Gyr × Holstein and on Hereford and Braford [13]. Similar study were conducted in Australia by [123]. In some instances, different studies identified QTLs on similar chromosomes, regardless of the differences in the breeds and tick species used. This underlines the need for the validation of the role of these chromosomes in cattle tick resistance. Validation could be pursued through GWAS meta-analysis, which can be achieved through collaborations.

**Table 5.** Previous GWAS studies on genomic regions associated with tick resistance in different regions of the world.

| Region | Breed | Sample Size | Mode of Infestation | Genotyping Platform | Findings | Reference |
|---|---|---|---|---|---|---|
| Brazil | $F_2$ *B. taurus* × *B. indicus* | 382 | Artificial | Microsatellite | Identified significant genomic regions on chromosomes 5, 7 and 14 | [11] |
| Brazil | $F_2$ Gyr × Holstein | 376 | Artificial | Microsatellite markers | Identified dry season specific QTL on BTA 2 and 10, rainy season specific QTL on BTA 5, 11 and 27 and BTA 23 for both seasons | [12] |
| Australia | Brown-Swiss, Holstein-Friesian, mixed taurine | 189 | Natural | MegAllele genotyping bovine10K SNP | Identified genes associated with tick burden, namely *TNFSF8 [CD30]*, and *SIRPA* | [123] |
| South Africa | Nguni | 586 | Natural | Illumina BovineSNP50 BeadChip | Identified significant genomic regions on chromosomes 1, 3, 6, 7, 8, 10, 11, 12, 14, 15, 17, 19 and 26 | [8] |
| Brazil | Braford and Hereford | 3455 | Natural | Illumina BovineSNP50 BeadChip | Identified 48 tag SNPs associated with tick resistance | [13] |
| Brazil | $F_2$ Gir × Holstein | 46 | Artificial | Illumina BovineSNP50 BeadChip | Identified genes associated with immune system function, namely, *TREM1*, *TREM2*, *CD83*, *MYO5A*, *TREML1*, and *PRSS16* | [14] |

$F_2$—Second filial generation; QTL—quantitative trait loci; BTA—Bos taurus; SNPs—single nucleotide polyimorphisms.

To date, only one study in Africa used SNP genotyping and GWAS as an approach to investigate genetic variants associated with tick resistance in cattle [8]. This study was conducted on South African Nguni cattle and identified several genomic regions

harboring QTLs associated with tick count traits. Despite the studies conducted previously, information on genetic determinants associated with cattle resistance to ticks is still limited. Further investigations focusing on unravelling genomic determinants associated with tick resistance will identify and provide understanding on biological mechanisms associated with TTBDs in cattle production. The information from the investigations will present a great opportunity to improve breeding programs to produce animals that are more resistant to tick infestation, while enhancing productivity [7,12].

## 9. Breeding Cattle for Tick Resistance

Breeding for genetic resistance is a potentially promising strategy to control ticks [124]. Sufficient genetic variation is one of the key factors determining the success of breeding schemes in livestock production. Bovine quantitative genetics studies have demonstrated low to high heritability for resistance to ticks depending on the breed [125,126]. Such findings hold the potential for tick resistance to be included as a goal in breeding schemes. *Bos taurus* breeds are known to be highly productive; however, they are also known to be highly susceptible to ticks and this makes their use in tropical production systems unsustainable, especially by resource-poor farmers. On the other hand, *Bos indicus* breeds are known for being resilient toward ticks as compared to *B. taurus*, but they yield lower production.

These differences motivate the development of crossbreeding programs between *Bos indicus* and *Bos taurus* cattle. These programs aim to produce crossbred animals, to facilitate the improvement of production whilst controlling TTBDs through the use of genomic selection. The success of genomic selection for tick resistance depends on the availability of proper genetic evaluation programs for the trait. The existence of genetic evaluation programs for tick resistance has the potential to generate information and facilitate the improvement of tick resistance. In countries such as Australia and Brazil, there are ongoing genetic evaluation programs for tick resistance. These develop crossbred animals that are productive under environmental conditions that have a high prevalence of ticks [127]. In Brazil, there is a genetic evaluation program known as Delta G Connection, which involves the Hereford and Braford cattle [102]. The Australians have developed the Australian Friesian Sahiwal, which produces acceptable levels of milk in an environment with ticks. The lack of information of such programs from other regions may suggest less enthusiasm from other regions, especially Africa and Asia [127]. There are concerns about the selection potential for resistance to ticks and the tradeoffs with other traits of economic significance. In this regard, studies have demonstrated that there is very low genetic correlation between tick count and various productive, adaptive and pubertal traits [105].

The existence of genetic components of variation in host resistance to ticks in cattle is currently being studied to discover tick resistance molecular markers that can be used in marker assisted selection (MAS). This information from locally adopted breeds can be used through crossbreeding to upgrade local breeds using highly productive exotic breeds. The challenges hindering the application of MAS for tick resistance are costs, resource populations, requirements of technical skills and the validation of discovered molecular markers for each population. All these challenges translate to the lack of molecular markers associated with tick resistance in cattle. However, despite the limitations, the development of studies that will investigate and generate genomic information on tick resistance and production traits holds the potential to increase the accuracy of selection. Therefore, the use of molecular genetics techniques, together with conventional breeding tools, is important in balancing the process of selection for tick resistance.

Host resistance to ticks is potentially an alternative tick control strategy that could solve the current TTBDs problems affecting the beef and dairy industry. The development of this alternative strategy requires the generation of knowledge that will broaden the understanding of biological mechanisms underlying tick resistance in cattle. Generally, the first step to control for a certain trait of interest requires the studying of genetic determinants influencing the expression of the trait. The advancement of technology has made it possible to use a genome-wide association approach to gain understanding on

mechanisms underlying tick resistance in cattle and to generate knowledge. However, information is still limited, and this is one of the hurdles preventing the facilitation of breeding for tick resistance through genomic selection. The information gap warrants the need for more GWASs to be conducted, to provide an understanding on biological mechanisms underlying tick resistance.

The costs associated with TTBDs are a major constrain for beef and dairy production. Therefore, the use of genomic selection as a tool to breed for tick resistant cattle will reduce the costs associated with the intensive use of vaccines and acaricides. This will further ensure animal welfare, facilitate increased production and subsequently increase profit margins. The implementation of genomic selection for tick resistance in developing countries is currently hindered by the high costs associated with the generation of a large training population, phenotyping and genotyping. The success of genomic selection for tick resistance is possible through the use of cost-effective genotype imputation methods to increase the power of GWASs and the accuracy of GEBVs estimates. There are countries that have included tick resistance trait in their breeding goals. However, from those countries, there are no studies that have been put in place to investigate the economic aspects of including the tick resistance trait in a genomic selection breeding goal. Therefore, there is a need for studies that will investigate the prospect of cost-effective genomic selection for tick resistance.

## 10. Limitations for GWAS to Uncover Tick Resistance Causal Variants in Cattle

The discovery of common causal variants associated with tick resistance is limited by various attributes, which include the nature of the trait, methodological challenges and the lack of financial resources. Tick resistance is a polygenic trait, influenced by more than one gene. This means the discovery of major genes through the use of GWAS is impossible. Factors such as epistasis, epigenetics, microbiome, and environment attribute to missing heritability [128,129]. This results in GWAS not being able to capture all the genetic determinates underlying the expression of tick resistance trait [14]. The lack of biological understanding on how these factors influence missing heritability in cattle is part of the obstacles hindering the discovery of true loci associated with tick resistance. The failure of GWAS to capture common variants is not a limitation to GWAS only. To date, there is no perfect genomic technology available to capture all the genetic information underlying the expression of complex traits [130]. Despite this drawback, there is a pressing need to use the available technologies to discover information that could be used to improve the trait through selection. Methodological challenges such as the GWAS study design of choice, poor structure of the data, phenotyping uncertainties, genotyping errors, improper data analysis, and the poor interpretation of results plays a huge role in limiting the success of GWAS for the trait of interest. The use of a poor data structure and failure to ensure proper data analysis increase the discovery of type I and type II errors. Therefore, ensuring proper data analysis and a clear interpretation of results is crucial for the generation of information that will inform alternative controlling strategies for ticks. The fluctuating exchange rate makes it difficult for most researchers to conduct GWAS studies focused on studying the genetic architecture of tick resistance, since such studies require a large sample size that is expensive to develop, sustain and genotype.

## 11. Conclusions

When conducting GWAS, it is necessary to account for factors that affect the rate of discovering an association, and control the rate of discovering spurious associations. GWASs have been very successful in the discovery of SNPs and candidate genes associated with tick resistance in various cattle breeds of different origins. Despite the success, more information is needed given that most GWASs on tick resistance in cattle are underpowered. This underlines the need for continuous data collection to enable larger and more powerful studies. For the studies that have been conducted using low density markers, genotype imputation is the most appropriate cost-effective approach for GWAS for tick resistance

in cattle. The availability of modest research populations, tools and funds are the current limitations of GWASs in developing countries. GWASs in African countries are performed using SNP genotyping arrays developed using exotic breeds. Thus, ascertainment bias effect leads to a low discovery rate of variants that influence the expression of the phenotype. This underscores the need for the inclusion of information from local breeds of different regions in the development of SNP genotyping arrays for cattle. This will facilitate a better understanding of variation in the breeds that are naturally adapted to the African production environments.

**Author Contributions:** A.M., N.M. (Ntanganedzeni Mapholi) and K.D. Conceptualized, designed and sourced funding for the research. N.M. (Nelisiwe Mkize) and B.D. coordinated the research. N.M. (Nelisiwe Mkize) collect the data for the manuscript and initially drafted the manuscript. A.M., K.D., B.D., N.M. (Ntanganedzeni Mapholi) critically revised the manuscript. N.M. (Nelisiwe Mkize) and B.D. prepared the final draft of the manuscript. All authors have read and agreed to the published version of the manuscript.

**Conflicts of Interest:** The authors declare that there is no conflict of interest.

# References

1.  Cardoso, F.F.; Gomes, C.C.G.; Sollero, B.P.; Oliveira, M.; Roso, V.M.; Piccoli, M.L.; Higa, R.H.; Yokoo, M.J.; Caetano, A.R.; Aguilar, I. Genomic prediction for tick resistance in Braford and Hereford cattle. *J. Anim. Sci.* **2015**, *93*, 2693–2705. [CrossRef] [PubMed]
2.  Muchenje, V.; Dzama, K.; Chimonyo, M.; Raats, J.G.; Strydom, P.E. Tick susceptibility and its effects on growth performance and carcass characteristics of Nguni, Bonsmara and Angus steers raised on natural pasture. *Animal* **2008**, *2*, 298–304. Available online: https://www.cambridge.org/core/journals/animal/article/abs/tick-susceptibility-and-its-effects-on-growth-performance-and-carcass-characteristics-of-nguni-bonsmara-and-angus-steers-raised-on-natural-pasture/F17959F52851D2B2B159FF9A9C30BE40 (accessed on 20 May 2019). [CrossRef]
3.  Ibelli, A.; Ribeiro, A.; Giglioti, R.; Regitano, L.; Alencar, M.; Chagas, A.; Paço, A.; Oliveira, H.; Duarte, J.; Oliveira, M. Resistance of cattle of various genetic groups to the tick *Rhipicephalus microplus* and the relationship with coat traits. *Veter. Parasitol.* **2012**, *186*, 425–430. Available online: https://www.sciencedirect.com/science/article/pii/S0304401711007515?via%3Dihub (accessed on 20 May 2019). [CrossRef]
4.  Guerrero, F.D.; Miller, R.J.; de León, A.A.P. Cattle tick vaccines: Many candidate antigens, but will a commercially viable product emerge? *Int. J. Parasitol.* **2012**, *42*, 421–427. [CrossRef]
5.  Castro-Janer, E.; Martins, J.R.; Mendes, M.C.; Namindome, A.; Klafke, G.M.; Schumaker, T.T.S. Diagnoses of fipronil resistance in Brazilian cattle ticks (*Rhipicephalus* (*Boophilus*) *microplus*) using in vitro larval bioassays. *Veter. Parasitol.* **2010**, *173*, 300–306. [CrossRef]
6.  Rajput, Z.I.; Hu, S.; Chen, W.; Arijo, A.G.; Xiao, C. Importance of ticks and their chemical and immunological control in livestock. *J. Zhejiang Univ. Sci. B* **2006**, *7*, 912–921. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1635821/ (accessed on 20 May 2019). [CrossRef]
7.  Mapholi, N.O.; Marufu, M.C.; Maiwashe, A.; Banga, C.B.; Muchenje, V.; MacNeil, M.D.; Chimonyo, M.; Dzama, K. Towards a genomics approach to tick (Acari: Ixodidae) control in cattle: A review. *Ticks Tick-Borne Dis.* **2014**, *5*, 475–483. [CrossRef]
8.  Mapholi, N.O.; Maiwashe, A.; Matika, O.; Riggio, V.; Bishop, S.C.; MacNeil, M.D.; Banga, C.; Taylor, J.F.; Dzama, K. Genome-wide association study of tick resistance in South African Nguni cattle. *Ticks Tick-Borne Dis.* **2016**, *7*, 487–497. [CrossRef] [PubMed]
9.  Nyangiwe, N.; Yawa, M.; Muchenje, V. Driving forces for changes in geographic range of cattle ticks (Acari: Ixodidae) in Africa: A review. *S. Afr. J. Anim. Sci.* **2018**, *48*, 829–841. Available online: http://www.scielo.org.za/scielo.php?script=sci_abstract&pid=S0375-15892018000600004&lng=en&nrm=iso&tlng=en (accessed on 20 May 2019). [CrossRef]
10. Acosta-Rodríguez, R.; Alonso-Morales, R.; Balladares, S.; Flores-Aguilar, H.; García-Vazquez, Z.; Gorodezky, C. Analysis of BoLA class II microsatellites in cattle infested with *Boophilus microplus* ticks: Class II is probably associated with susceptibility. *Veter. Parasitol.* **2005**, *127*, 313–321. [CrossRef]
11. Gasparin, G.; Miyata, M.; Coutinho, L.L.; Martinez, M.L.; Teodoro, R.L.; Furlong, J.; Machado, M.A.; Sliva, M.V.G.B.; Sonstegard, T.S.; Regitano, L.C.A. Mapping of quantitative trait loci controlling tick [*Rhipicephalus* (*Boophilus*) *microplus*] resistance on bovine chromosomes 5, 7 and 14. *Anim. Genet.* **2007**, *38*, 453–459. [CrossRef]

12. Machado, M.A.; SAzevedo, A.L.; Teodoro, R.L.; Pires, M.A.; CDPeixoto, M.G.; de Freitas, C.; Prata, M.C.A.; Furlong, J.; da Silva, M.V.G.; Guimaraes, S.F.E.; et al. Genome wide scan for quantitative trait loci affecting tick resistance in cattle (*Bos taurus × Bos indicus*). *BMC Genom.* **2010**, *11*, 280. Available online: https://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-11-280 (accessed on 20 May 2019). [CrossRef]

13. Sollero, B.P.; Junqueira, V.S.; Gomes, C.C.G.; Caetano, A.R.; Cardoso, F.F. Tag SNP selection for prediction of tick resistance in Brazilian Braford and Hereford cattle breeds using Bayesian methods. *Genet. Sel. Evol.* **2017**, *49*, 49. [CrossRef] [PubMed]

14. Otto, P.I.; Guimarães, S.E.F.; Verardo, L.L.; Azevedo, A.L.S.; Vandenplas, J.; Sevillano, C.A.; Marques, D.B.D.; Pires, M.F.A.; Freitas, C.; Verneque, R.S.; et al. Genome-wide association studies for tick resistance in Bos taurus × Bos indicus crossbred cattle: A deeper look into this intricate mechanism. *J. Dairy Sci.* **2018**, *101*, 11020–11032. [CrossRef] [PubMed]

15. Marufu, M.; Qokweni, L.; Chimonyo, M.; Dzama, K. Relationships between tick counts and coat characteristics in Nguni and Bonsmara cattle reared on semiarid rangelands in South Africa. *Ticks Tick-Borne Dis.* **2011**, *2*, 172–177. [CrossRef]

16. Ibeagha-Awemu, E.M.; Kgwatalala, P.; Ibeagha, A.E.; Zhao, X. A critical analysis of disease-associated DNA polymorphisms in the genes of cattle, goat, sheep, and pig. *Mamm. Genome* **2008**, *19*, 226–245. [CrossRef]

17. Myles, S.; Peiffer, J.; Brown, P.J.; Ersoz, E.S.; Zhang, Z.; Costich, D.E.; Buckler, E.S. Association mapping: Critical considerations shift from genotyping to experimental design. *Plant Cell* **2009**, *21*, 2194–2202. [CrossRef]

18. Visscher, P.M.; Wray, N.R.; Zhang, Q.; Sklar, P.; McCarthy, M.I.; Brown, M.A.; Yang, J. 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am. J. Hum. Genet.* **2017**, *101*, 5–22. [CrossRef]

19. Sherva, R.; Farrer, L.A. Power and Pitfalls of the Genome-Wide Association Study Approach to Identify Genes for Alzheimer's Disease. *Curr. Psychiatry Rep.* **2011**, *13*, 138–146. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3154249/ (accessed on 20 May 2019). [CrossRef]

20. Psifidi, A.; Banos, G.; Matika, O.; Desta, T.T.; Bettridge, J.; Hume, D.A.; Dessie, T.; Christely, R.; Wigely, R.; Hanotte, O.; et al. Genome-wide association studies of immune, disease and production traits in indigenous chicken ecotypes. *Genet. Sel. Evol.* **2016**, *48*, 74. [CrossRef]

21. Patron, J.; Serra-Cayuela, A.; Han, B.; Li, C.; Wishart, D.S. Assessing the performance of genome-wide association studies for predicting disease risk. *PLoS ONE* **2019**, *14*, e0220215. Available online: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0220215 (accessed on 20 May 2019).

22. Wang, Z.; Chatterjee, N. Increasing mapping precision of genomewide association studies: To genotype and impute, sequence, or both? *Genome Biol.* **2017**, *18*, 118. [CrossRef]

23. Paes, G.P.; Viana, J.M.S.; Silva, F.F.E.; Mundim, G.B. Linkage disequilibrium, SNP frequency change due to selection, and association mapping in popcorn chromosome regions containing QTLs for quality traits. *Genet. Mol. Biol.* **2016**, *39*, 97–110. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4807383/ (accessed on 20 May 2019). [CrossRef] [PubMed]

24. Amos, C.I. Successful design and conduct of genome-wide association studies. *Hum. Mol. Genet.* **2007**, *16*, R220–R225. [CrossRef]

25. Greely, H.T. The uneasy ethical and legal underpinnings of large-scale genomic biobanks. *Annu. Rev. Genom. Hum. Genet.* **2007**, *8*, 343–364. [CrossRef]

26. Reed, E.; Nunez, S.; Kulp, D.; Qian, J.; Reilly, M.P.; Foulkes, A.S. A guide to genome-wide association analysis and post-analytic interrogation. *Stat. Med.* **2015**, *34*, 3769–3792. [CrossRef]

27. Bermingham, M.L.; Bishop, S.C.; Woolliams, J.A.; Pong-Wong, R.; Allen, A.R.; McBride, S.H.; Ryder, J.J.; Wright, D.M.; Skuce, R.A.; McDowell, S.W.J.; et al. Genome-wide association study identifies novel loci associated with resistance to bovine tuberculosis. *Heredity* **2014**, *112*, 543–551. [CrossRef]

28. Sodeland, M.; Kent, M.P.; Olsen, H.G.; Opsal, M.A.; Svendsen, M.; Sehested, E.; Hayes, B.J.; Lien, S. Quantitative trait loci for clinical mastitis on chromosomes 2, 6, 14 and 20 in Norwegian Red cattle. *Anim. Genet.* **2011**, *42*, 457–465. [CrossRef]

29. Lee, B.-Y.; Lee, K.-N.; Lee, T.; Park, J.-H.; Kim, S.-M.; Lee, H.-S.; Chung, D.-S.; Shim, H.-S.; Lee, H.-K.; Kim, H. Bovine Genome-wide Association Study for Genetic Elements to Resist the Infection of Foot-and-mouth Disease in the Field. *Asian-Australas. J. Anim. Sci.* **2015**, *28*, 166–170. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4283160/ (accessed on 20 May 2019). [CrossRef] [PubMed]

30. Xia, J.; Fan, H.; Chang, T.; Xu, L.; Zhang, W.; Song, Y.; Bo, Z.; Zhang, L.; Gao, X.; Chen, Y.; et al. Searching for new loci and candidate genes for economically important traits through gene-based association analysis of Simmental cattle. *Sci. Rep.* **2017**, *7*, 42048. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5294460/ (accessed on 20 May 2019). [CrossRef]

31. Santiago, G.G.; Siqueira, F.; Cardoso, F.F.; Regitano, L.C.A.; Ventura, R.; Sollero, B.P.; Souza, M.D.; Monkry, F.B.; Ferreria, A.B.R.; Torres, R.A.A. Genomewide association study for production and meat quality traits in Canchim beef cattle. *J. Anim. Sci.* **2017**, *95*, 3381–3390. [CrossRef]

32. Yue, S.J.; Zhao, Y.Q.; Gu, X.R.; Yin, B.; Jiang, Y.L.; Wang, Z.H.; Shi, K.R. A genome-wide association study suggests new candidate genes for milk production traits in Chinese Holstein cattle. *Anim. Genet.* **2017**, *48*, 677–681. [CrossRef] [PubMed]

33. Chen, Z.; Yao, Y.; Ma, P.; Wang, Q.; Pan, Y. Haplotype-based genome-wide association study identifies loci and candidate genes for milk yield in Holsteins. *PLoS ONE* **2018**, *13*, e0192695. [CrossRef]

34. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950838/ (accessed on 20 May 2019). [CrossRef]

35. Aulchenko, Y.S.; Ripke, S.; Isaacs, A.; van Duijn, C.M. GenABEL: An R library for genome-wide association analysis. *Bioinformatics* **2007**, *23*, 1294–1296. [CrossRef] [PubMed]

36. Helix, G. *SNP & Variation Suite*; Version 8.x; Golden Helix, Inc.: Bozeman, MT, USA, 2017. Available online: https://www.goldenhelix.com/products/SNP_Variation/index.html (accessed on 1 June 2019).

37. Curtis, R.E.; Kinnaird, P.; Xing, E.P. GenAMap: Visualization Strategies for Structured Association Mapping. In Proceedings of the 2011 IEEE Symposium on Biological Data Visualization (BioVis), Providence, RI, USA, 23–24 October 2011.

38. Zhou, X.; Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **2012**, *44*, 821–824. [CrossRef]

39. Misztal, I.; Lourenco, D.; Aguilar, I.; Legarra, A.; Vitezica, Z. *Manual for BLUPF90 Family of Programs*; pp. 1–142. Available online: http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all6.pdf (accessed on 20 May 2019).

40. Hu, Z.-L.; Park, C.A.; Reecy, J.M. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. *Nucleic Acids Res.* **2019**, *47*, D701–D710. [CrossRef]

41. Vignal, A.; Milan, D.; SanCristobal, M.; Eggen, A. A review on SNP and other types of molecular markers and their use in animal genetics. *Genet. Sel. Evol.* **2002**, *34*, 275. [CrossRef]

42. Brumfield, R.; Beerli, P.; Nickerson, D.; Edwards SBrumfield, R.T.; Beerli, P.; Nickerson, D.A.; Edwards, S.V. The utility of single nucleotide polymorphisms in inferences of population history. *Trends Ecol. Evol.* **2003**, *18*, 249–256. [CrossRef]

43. Fan, B.; Du, Z.-Q.; Gorbach, D.M.; Rothschild, M.F. Development and Application of High-density SNP Arrays in Genomic Studies of Domestic Animals. *Asian-Australas. J. Anim. Sci.* **2010**, *23*, 833–847. Available online: https://www.koreascience.or.kr/article/JAKO201018861576911.page (accessed on 20 May 2019). [CrossRef]

44. Scheben, A.; Yuan, Y.; Edwards, D. Advances in genomics for adapting crops to climate change. *Curr. Plant Biol.* **2016**, *6*, 2–10. Available online: https://www.sciencedirect.com/science/article/pii/S221466281630024X (accessed on 20 May 2019). [CrossRef]

45. Ganal, M.W.; Polley, A.; Graner, E.-M.; Plieske, J.; Wieseke, R.; Luerssen, H.; Durstewwitz, G. Large SNP arrays for genotyping in crop plants. *J. Biosci.* **2012**, *37*, 821–828. [CrossRef] [PubMed]

46. Lashmar, S.F.; Muchadeyi, F.C.; Visser, C. Genotype imputation as a cost-saving genomic strategy for South African Sanga cattle: A review. *S. Afr. J. Anim. Sci.* **2019**, *49*, 262–280. Available online: https://www.ajol.info/index.php/sajas/article/view/186645 (accessed on 20 May 2019). [CrossRef]

47. Nicolazzi, E.L.; Biffani, S.; Biscarini, F.; Wengel, P.O.T.; Caprera, A.; Nazzicari, N.; Stella, A. Software solutions for the livestock genomics SNP array revolution. *Anim. Genet.* **2015**, *46*, 343–353. [CrossRef]

48. Schmid, M.; Bennewitz, J. Invited review: Genome-wide association analysis for quantitative traits in livestock—A selective review of statistical models and experimental designs. *Arch. Anim. Breed.* **2017**, *60*, 335–346. Available online: https://aab.copernicus.org/articles/60/335/2017/ (accessed on 20 May 2019). [CrossRef]

49. Alqudah, A.M.; Haile, J.K.; Alomari, D.Z.; Pozniak, C.J.; Kobiljski, B.; Börner, A. Genome-wide and SNP network analyses reveal genetic control of spikelet sterility and yield-related traits in wheat. *Sci. Rep.* **2020**, *10*, 2098. Available online: https://www.nature.com/articles/s41598-020-59004-4 (accessed on 20 May 2019). [CrossRef] [PubMed]

50. Wang, Y.; Zhang, F.; Mukiibi, R.; Chen, L.; Vinsky, M.; Plastow, G.; Basarab, J.; Stothard, P.; Li, C. Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: II: Carcass merit traits. *BMC Genom.* **2020**, *21*, 38. [CrossRef] [PubMed]

51. Yang, N.; Lu, Y.; Yang, X.; Huang, J.; Zhou, Y.; Ali, F.; Wen, W.; Liu, J.; Li, J.; Yan, J. Genome Wide Association Studies Using a New Nonparametric Model Reveal the Genetic Architecture of 17 Agronomic Traits in an Enlarged Maize Association Panel. *PLoS Genet.* **2014**, *10*, e1004573. Available online: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004573 (accessed on 20 May 2019). [CrossRef]

52. Fernando, R.L.; Nettleton, D.; Southey, B.R.; Dekkers, J.C.M.; Rothschild, M.F.; Soller, M. Controlling the Proportion of False Positives in Multiple Dependent Tests. *Genetics* **2004**, *166*, 611–619. Available online: https://www.genetics.org/content/166/1/611 (accessed on 20 May 2019). [CrossRef]

53. Gao, X.; Starmer, J.; Martin, E.R. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet. Epidemiol.* **2008**, *32*, 361–369. [CrossRef]

54. Hochberg, Y.; Benjamini, Y. More powerful procedures for multiple significance testing. *Stat. Med.* **1990**, *9*, 811–818. [CrossRef] [PubMed]

55. De, R.; Bush, W.S.; Moore, J.H. Bioinformatics challenges in genome-wide association studies (GWAS). *Methods Mol. Biol.* **2014**, *1168*, 63–81. [PubMed]

56. Zeng, P.; Zhao, Y.; Qian, C.; Zhang, L.; Zhang, R.; Gou, J.; Liu, J.; Liu, L.; Chen, F. Statistical analysis for genome-wide association study. *J. Biomed. Res.* **2015**, *29*, 285–297. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4547377/ (accessed on 20 May 2019).

57. 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.; VHakravarti, A.; et al. Finding the missing heritability of complex diseases. *Nature* **2009**, *461*, 747–753. [CrossRef]

58. Moore, J.H.; Asselbergs, F.W.; Williams, S.M. Bioinformatics challenges for genome-wide association studies. *Bioinforma. Oxf. Engl.* **2010**, *26*, 445–455. [CrossRef]

59.	Sandoval-Motta, S.; Aldana, M.; Martínez-Romero, E.; Frank, A. The Human Microbiome and the Missing Heritability Problem. *Front. Genet.* **2017**, *8*, 80. Available online: https://www.frontiersin.org/article/10.3389/fgene.2017.00080 (accessed on 20 May 2019). [CrossRef] [PubMed]
60.	Li, C.; Fu, Y.; Sun, R.; Wang, Y.; Wang, Q. Single-Locus and Multi-Locus Genome-Wide Association Studies in the Genetic Dissection of Fiber Quality Traits in Upland Cotton (*Gossypium hirsutum* L.). *Front. Plant Sci.* **2018**, *9*, 1083. [CrossRef] [PubMed]
61.	Zhang, Y.-M.; Jia, Z.; Dunwell, J.M. Editorial: The Applications of New Multi-Locus GWAS Methodologies in the Genetic Dissection of Complex Traits. *Front. Plant Sci.* **2019**, *10*, 100. Available online: https://www.frontiersin.org/article/10.3389/fpls.2019.00100 (accessed on 20 May 2019). [CrossRef] [PubMed]
62.	Chu, B.B.; Keys, K.L.; German, C.A.; Zhou, H.; Zhou, J.J.; Sobel, E.; Sinsheimer, J.S.; Lange, K. Iterative hard thresholding in genome-wide association studies: Generalized linear models, prior weights, and double sparsity. *GigaScience* **2020**, *9*, 1–13. [CrossRef]
63.	Zhang, Z.; Ersoz, E.; Lai, C.-Q.; Todhunter, R.J.; Tiwari, H.K.; Gore, M.A.; Bradbury, P.J.; Yu, J.; Arnett, D.K.; Ordovas, J.M.; et al. Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* **2010**, *42*, 355–360. [CrossRef]
64.	Shenstone, E.; Cooper, J.; Rice, B.; Bohn, M.; Jamann, T.M.; Lipka, A.E. An assessment of the performance of the logistic mixed model for analyzing binary traits in maize and sorghum diversity panels. *PLoS ONE* **2018**, *13*, e0207752. Available online: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0207752 (accessed on 20 May 2019). [CrossRef] [PubMed]
65.	Wang, S.-B.; Feng, J.-Y.; Ren, W.-L.; Huang, B.; Zhou, L.; Wen, Y.-J.; Zhang, J.; Dunwell, J.M.; Xu, S.; Zhang, Y.-M. Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. *Sci. Rep.* **2016**, *6*, 19444. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4726296/ (accessed on 20 May 2019). [CrossRef] [PubMed]
66.	Zhang, Y.-W.; Tamba, C.L.; Wen, Y.-J.; Li, P.; Ren, W.-L.; Ni, Y.-L.; Gao, J.; Zhang, Y.-M. mrMLM v4.0.2: An R Platform for Multi-locus Genome-wide Association Studies. *Genom. Proteom. Bioinform.* **2020**, *18*, 481–487. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8242264/ (accessed on 20 May 2019). [CrossRef]
67.	Wen, Y.; Zhang, Y.; Zhang, J.; Feng, J.; Zhang, Y. The Improved FASTmrEMMA and GCIM Algorithms for Genome-Wide Association and Linkage Studies in Large Mapping Populations. *Crop J.* **2020**, *8*, 723–732. [CrossRef]
68.	Wen, Y.-J.; Zhang, Y.-W.; Zhang, J.; Feng, J.-Y.; Dunwell, J.M.; Zhang, Y.-M. An efficient multi-locus mixed model framework for the detection of small and linked QTLs in F2. *Briefings Bioinform.* **2018**, *20*, 1913–1924. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6917223/ (accessed on 20 May 2019). [CrossRef] [PubMed]
69.	Wang, S.; Dvorkin, D.; Da, Y. SNPEVG: A graphical tool for GWAS graphing with mouse clicks. *BMC Bioinform.* **2012**, *13*, 319. [CrossRef] [PubMed]
70.	Turner, S.D. qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Source Softw.* **2018**, *3*, 731. [CrossRef]
71.	Cook, D.E.; Ryckman, K.R.; Murray, J.C. Generating Manhattan Plots in Stata. *Stata J.* **2013**, *13*, 323–328. [CrossRef]
72.	Grace, C.; Farrall, M.; Watkins, H.; Goel, A. Manhattan++: Displaying genome-wide association summary statistics with multiple annotation layers. *BMC Bioinform.* **2019**, *20*, 610. [CrossRef]
73.	Haller, T.; Tasa, T.; Metspalu, A. Manhattan Harvester and Cropper: A system for GWAS peak detection. *BMC Bioinform.* **2019**, *20*, 22. [CrossRef]
74.	Pruim, R.J.; Welch, R.P.; Sanna, S.; Teslovich, T.M.; Chines, P.S.; Gliedt, T.P.; Boehnke, M.; Abecasisi, G.R.; Willer, C.J. LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* **2010**, *26*, 2336–2337. [CrossRef]
75.	Chittoor, G.; Haack, K.; Mehta, N.R.; Laston, S.; Cole, S.A.; Comuzzie, A.G.; Butte, N.F.; Voruganti, V.S. Genetic variation underlying renal uric acid excretion in Hispanic children: The Viva La Familia Study. *BMC Med. Genet.* **2017**, *18*, 6. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5240212/ (accessed on 20 May 2019). [CrossRef] [PubMed]
76.	Zimin, A.V.; Delcher, A.L.; Florea, L.; Kelley, D.R.; Schatz, M.C.; Puiu, D.; Hanrahana, F.; Pertea, G.; Tassell, C.V.; Sonstehard, T.S.; et al. A whole-genome assembly of the domestic cow, Bos taurus. *Genome Biol.* **2009**, *10*, R42. [CrossRef] [PubMed]
77.	Kanehisa, M.; Goto, S.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. Data, information, knowledge and principle: Back to metabolism in KEGG. *Nucleic Acids Res.* **2014**, *42*, D199–D205. [CrossRef] [PubMed]
78.	Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4*, 44–57. Available online: https://www.nature.com/articles/nprot.2008.211 (accessed on 20 May 2019). [CrossRef] [PubMed]
79.	Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [CrossRef]
80.	Pickrell, J.; Clerget-Darpoux, F.; Bourgain, C. Power of genome-wide association studies in the presence of interacting loci. *Genet. Epidemiology* **2007**, *31*, 748–762. Available online: https://onlinelibrary.wiley.com/doi/abs/10.1002/gepi.20238 (accessed on 20 May 2019). [CrossRef]
81.	Ledur, M.C.; Navarro, N.; Pérez-Enciso, M. Large-scale SNP genotyping in crosses between outbred lines: How useful is it? *Heredity* **2010**, *105*, 173–182. [CrossRef]
82.	Schmid, M.; Wellmann, R.; Bennewitz, J. Power and precision of QTL mapping in simulated multiple porcine F2 crosses using whole-genome sequence information. *BMC Genet.* **2018**, *19*, 22. [CrossRef]

83. Morin, P.A.; Luikart, G.; Wayne, R.K.; The SNP Workshop Group. SNPs in ecology, evolution and conservation. *Trends Ecol. Evol.* **2004**, *19*, 208–216. Available online: https://www.sciencedirect.com/science/article/pii/S0169534704000205 (accessed on 20 May 2019). [CrossRef]

84. Burrow, H.; Mans, B.; Cardoso, F.; Birkett, M.; Kotze, A.; Hayes, B.; Mapholi, N.; Dzama, K.; Marufu, M.C.; Githaka, N.; et al. Towards a new phenotype for tick resistance in beef and dairy cattle: A review. *Anim. Prod. Sci.* **2019**, *59*, 1401–1427. [CrossRef]

85. Mapholi, N.O.; Maiwashe, A.; Matika, O.; Riggio, V.; Banga, C.; MacNeil, M.D.; Muchenje, V.; Nephawe, K.; Dzama, K. Genetic parameters for tick counts across months for different tick species and anatomical locations in South African Nguni cattle. *Trop. Anim. Health Prod.* **2017**, *49*, 1201–1210. [CrossRef] [PubMed]

86. Prayaga, K.C.; Corbet, N.J.; Johnston, D.J.; Wolcott, M.L.; Fordyce, G.; Burrow, H.M. Genetics of adaptive traits in heifers and their relationship to growth, pubertal and carcass traits in two tropical beef cattle genotypes. *Anim. Prod. Sci.* **2009**, *49*, 413–425. Available online: https://www.publish.csiro.au/an/EA08247 (accessed on 20 May 2019). [CrossRef]

87. Burrow, H.M. Genetic Aspects of Cattle Adaptation in the Tropics. In *The Genetics of Cattle*; CAB International; Place: Oxfordshire, UK, 2014. Available online: https://rune.une.edu.au/web/handle/1959.11/17403 (accessed on 20 May 2019).

88. Hong, E.P.; Park, J.W. Sample size and statistical power calculation in genetic association studies. *Genom. Inform.* **2012**, *10*, 117–122. [CrossRef]

89. Kumar, J.; Pratap, A.; Solanki, R.K.; Gupta, D.S.; Goyal, A.; Chaturvedi, S.K.; Nadarajana, N.; Kumar, S. Genomic resources for improving food legume crops. *J. Agric. Sci.* **2011**, *150*, 289–318. Available online: https://www.cambridge.org/core/journals/journal-of-agricultural-science/article/abs/genomic-resources-for-improving-food-legume-crops/23984357B2E09BABF3F6A578631601A8 (accessed on 20 May 2019). [CrossRef]

90. Yano, K.; Yamamoto, E.; Aya, K.; Takeuchi, H.; Lo, P.-C.; Hu, L.; Yamasaki, M.; Yoshida, S.; Kitano, H.; Hirano, K.; et al. Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat. Genet.* **2016**, *48*, 927–934. [CrossRef] [PubMed]

91. Klein, R.J. Power analysis for genome-wide association studies. *BMC Genet.* **2007**, *8*, 58. [CrossRef]

92. Park, J.-H.; Wacholder, S.; Gail, M.H.; Peters, U.; Jacobs, K.; Chanock, S.J.; Chatterjee, N. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat. Genet.* **2010**, *42*, 570–575. [CrossRef]

93. Anderson, C.A.; Pettersson, F.H.; Clarke, G.M.; Cardon, L.R.; Morris, A.P.; Zondervan, K.T. Data quality control in genetic case-control association studies. *Nat. Protoc.* **2010**, *5*, 1564–1573. Available online: https://www.nature.com/articles/nprot.2010.116 (accessed on 20 May 2019). [CrossRef]

94. Teo, Y.Y. Common statistical issues in genome-wide association studies: A review on power, data quality control, genotype calling and population structure. *Curr. Opin. Lipidol.* **2008**, *19*, 133–143. [CrossRef] [PubMed]

95. Laird, N.M.; Horvath, S.; Xu, X. Implementing a unified approach to family-based tests of association. *Genet. Epidemiol.* **2000**, *19* (Suppl. S1), S36–S42. [CrossRef]

96. Wigginton, J.E.; Cutler, D.J.; Abecasis, G.R. A Note on Exact Tests of Hardy-Weinberg Equilibrium. *Am. J. Hum. Genet.* **2005**, *76*, 887–893. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1199378/ (accessed on 20 May 2019). [CrossRef] [PubMed]

97. Slatkin, M. Linkage disequilibrium—Understanding the evolutionary past and mapping the medical future. *Nat. Rev. Genet.* **2008**, *9*, 477–485. Available online: https://www.nature.com/articles/nrg2361 (accessed on 20 May 2019). [CrossRef] [PubMed]

98. Bush, W.S.; Moore, J.H. Chapter 11: Genome-wide association studies. *PLoS Comput. Biol.* **2012**, *8*, e1002822. [CrossRef]

99. Flint-Garcia, S.; Thornsberry, J.; Buckler, E. Structure of Linkage Disequilibrium in Plants. *Annu. Rev. Plant Biol.* **2003**, *54*, 357–374. [CrossRef]

100. Falconer, D.S.; Mackay, T.F.C. *Introduction to Quantitative Genetics*; Subsequent Edition; Benjamin-Cummings Pub. Co.: Harlow, UK, 1996; p. 464.

101. Kolmodin, R.; Strandberg, E.; Madsen, P.; Jensen, J.; Jorjani, H. Genotype by Environment Interaction in Nordic Dairy Cattle Studied Using Reaction Norms. *Acta Agric. Scand. Sect.-Anim. Sci.* **2002**, *52*, 11–24. [CrossRef]

102. Mota, R.R.; Tempelman, R.J.; Lopes, P.S.; Aguilar, I.; Silva, F.F.; Cardoso, F.F. Genotype by environment interaction for tick resistance of Hereford and Braford beef cattle using reaction norm models. *Genet. Sel. Evol.* **2016**, *48*, 3. [CrossRef]

103. Hayes, B.J.; Daetwyler, H.D.; Goddard, M.E. Models for Genome × Environment Interaction: Examples in Livestock. *Crop. Sci.* **2016**, *56*, 2251–2259. [CrossRef]

104. Biegelmeyer, P.; Nizoli, L.Q.; Cardoso, F.F.; Dionello, N.J. Aspectos da resistência de bovinos ao carrapato rhipicephalus (boophilus) microplus. *Arch. Zootec.* **2012**, *61*, 1–11. Available online: https://www.uco.es/ucopress/az/index.php/az/article/view/2954 (accessed on 20 May 2019). [CrossRef]

105. Mulder, H.A.; Bijma, P. Effects of genotype × environment interaction on genetic gain in breeding programs1. *J. Anim. Sci.* **2005**, *83*, 49–61. [CrossRef]

106. Sae-Lim, P.; Grøva, L.; Olesen, I.; Varona, L. A comparison of nonlinear mixed models and response to selection of tick-infestation on lambs. *PLoS ONE* **2017**, *12*, e0172711. [CrossRef] [PubMed]

107. Cooley, P.C.; Clark, R.F.; Folsom, R.E. *Assessing Gene-Environment Interactions in Genome-Wide Association Studies: Statistical Approaches*; Research Triangle Park (NC); RTI Press Research Report Series; 2014. Available online: http://www.ncbi.nlm.nih.gov/books/NBK532519/ (accessed on 20 May 2019).

108. Winham, S.J.; Biernacka, J.M. Gene-environment interactions in genome-wide association studies: Current approaches and new directions. *J. Child Psychol. Psychiatry* **2013**, *54*, 1120–1134. [CrossRef] [PubMed]

109. Marigorta, U.M.; Egibson, G. A simulation study of gene-by-environment interactions in GWAS implies ample hidden effects. *Front. Genet.* **2014**, *5*, 225. Available online: https://www.frontiersin.org/article/10.3389/fgene.2014.00225 (accessed on 20 May 2019). [CrossRef]

110. Arnau-Soler, A.; Scotland, G.; Macdonald-Dunlop, E.; Adams, M.J.; Clarke, T.-K.; MacIntyre, D.J.; Milburn, K.; Navrady, L.; Scotland, G.; Major Depressive Disorder Working group of the Psychiatric Genomic Consortium; et al. Genome-wide by environment interaction studies of depressive symptoms and psychosocial stress in UK Biobank and Generation Scotland. *Transl. Psychiatry* **2019**, *9*, 1–13. Available online: https://www.nature.com/articles/s41398-018-0360-y (accessed on 20 May 2019). [CrossRef] [PubMed]

111. Gali, K.K.; Sackville, A.; Tafesse, E.G.; Lachagari, V.R.; McPhee, K.; Hybl, M.; Mikic, A.; Smykal, P.; McGee, R.; Burstin, J.; et al. Genome-Wide Association Mapping for Agronomic and Seed Quality Traits of Field Pea (*Pisum sativum* L.). *Front. Plant Sci.* **2019**, *10*, 1538. Available online: https://www.frontiersin.org/article/10.3389/fpls.2019.01538 (accessed on 20 May 2019). [CrossRef] [PubMed]

112. Eltaher, S.; Baenziger, P.S.; Belamkar, V.; Emara, H.A.; Nower, A.A.; Salem, K.F.M.; Alqudah, A.M.; Sallam, A. GWAS revealed effect of genotype × environment interactions for grain yield of Nebraska winter wheat. *BMC Genom.* **2021**, *22*, 2. [CrossRef]

113. Turner, S.; Armstrong, L.L.; Bradford, Y.; Carlson, C.S.; Crawford, D.C.; Crenshaw, A.T.; de Andrade, M.; Dohey, K.F.; Haines, J.L.; Hayes, G.; et al. Quality Control Procedures for Genome-Wide Association Studies. *Curr. Protoc. Hum. Genet.* **2011**, *68*, 1–19. [CrossRef] [PubMed]

114. 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*, e1000529. Available online: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1000529 (accessed on 20 May 2019). [CrossRef] [PubMed]

115. Naj, A.C. Genotype Imputation in Genome-Wide Association Studies. *Curr. Protoc. Hum. Genet.* **2019**, *102*, e84. [CrossRef]

116. Spencer, C.C.A.; Su, Z.; Donnelly, P.; Marchini, J. Designing Genome-Wide Association Studies: Sample Size, Power, Imputation, and the Choice of Genotyping Chip. *PLoS Genet.* **2009**, *5*, e1000477. Available online: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1000477 (accessed on 20 May 2019). [CrossRef] [PubMed]

117. Garcia-Ruiz, A.; Ruiz-López, F.; Wiggans, G.; Van Tassell, C.; Montaldo, H. Effect of reference population size and available ancestor genotypes on imputation of Mexican Holstein genotypes. *J. Dairy Sci.* **2015**, *98*, 3478–3484. [CrossRef]

118. Holm, H.; Gudbjartsson, D.F.; Sulem, P.; Masson, G.; Helgadottir, H.T.; Zanon, C.; Magnusson, O.T.; Helgasin, A.; Saemundsdottir, J.; Gylfason, A.; et al. A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nat. Genet.* **2011**, *43*, 316–320. [CrossRef]

119. Pistis, G.; Porcu, E.; Vrieze, S.I.; Sidore, C.; Steri, A.M.; Danjou, F.; Busonero, F.; Mulas, A.; Zoledziuewska, M.; Maschio, A.; et al. Rare variant genotype imputation with thousands of study-specific whole-genome sequences: Implications for cost-effective study designs. *Eur. J. Hum. Genet.* **2014**, *23*, 975–983. [CrossRef] [PubMed]

120. Quick, C.; Anugu, P.; Musani, S.; Weiss, S.T.; Burchard, E.G.; White, M.J.; Keys, K.L.; Cucca, F.; Sidore, C.; Boehnke, M.; et al. Sequencing and imputation in GWAS: Cost-effective strategies to increase power and genomic coverage across diverse populations. *Genet. Epidemiol.* **2020**, *44*, 537–549. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7449570/ (accessed on 20 May 2019). [CrossRef] [PubMed]

121. Sargolzaei, M.; Chesnais, J.; Schenkel, F. FImpute-An efficient imputation algorithm for dairy cattle populations. *J. Dairy Sci.* **2011**, *94*, 421.

122. Tabor, A.E.; Ali, A.; Rehman, G.; Garcia, G.R.; Zangirolamo, A.F.; Malardo, T.; Jonsson, N.N. Cattle Tick Rhipicephalus microplus-Host Interface: A Review of Resistant and Susceptible Host Responses. *Front. Cell. Infect. Microbiol.* **2017**, *7*, 506. [CrossRef] [PubMed]

123. Turner, L.B.; Harrison, B.E.; Bunch, R.J.; Neto, L.R.P.; Li, Y.; Barendse, W.; Turner, L.B.; Harrison, B.E.; Bunch, R.J.; Neto, L.R.P.; et al. A Genome-Wide Association Study of Tick Burden and Milk Composition in Cattle. *Anim. Prod. Sci.* **2010**, *50*, 235–245. [CrossRef]

124. Regitano, L.; Prayaga, K. *Breeding for Disease Resistance in Farm Animals*, 3rd ed.; CABI Publishing: Oxon, UK, 2010; pp. 295–314.

125. Henshall, J.M. A genetic analysis of parasite resistance traits in a tropically adapted line of Bos taurus. *Aust. J. Agric. Res.* **2004**, *55*, 1109–1116. Available online: https://www.publish.csiro.au/ar/AR03085 (accessed on 20 May 2019). [CrossRef]

126. Budeli, M.A.; Nephawe, K.A.; Norris, D.; Selapa, N.W.; Bergh, L.; Maiwashe, A. Genetic parameter estimates for tick resistance in Bonsmara cattle. *S. Afr. J. Anim. Sci.* **2010**, *39*. Available online: https://www.ajol.info/index.php/sajas/article/view/51125 (accessed on 20 May 2019). [CrossRef]

127. Shyma, K.P.; Gupta, J.P.; Singh, V. Breeding strategies for tick resistance in tropical cattle: A sustainable approach for tick control. *J. Parasit. Dis.* **2013**, *39*, 1–6. [CrossRef] [PubMed]

128. Frazer, K.A.; Murray, S.S.; Schork, N.J.; Topol, E.J. Human genetic variation and its contribution to complex traits. *Nat. Rev. Genet.* **2009**, *10*, 241–251. Available online: https://www.nature.com/articles/nrg2554 (accessed on 20 May 2019). [CrossRef] [PubMed]

129. Aschard, H.; Lutz, S.; Maus, B.; Duell, E.J.; Fingerlin, T.; Chatterjee, N.; Kraft, P.; Steen, K.V. Challenges and opportunities in genome-wide environmental interaction (GWEI) studies. *Qual. Life Res.* **2012**, *131*, 1591–1613. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3677711/ (accessed on 20 May 2019). [CrossRef] [PubMed]
130. Tam, V.; Patel, N.; Turcotte, M.; Bossé, Y.; Paré, G.; Meyre, D. Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* **2019**, *20*, 467–484. [CrossRef]