Review

Genomic Selection—Considerations for Successful Implementation in Wheat Breeding Programs

Dylan Lee Larkin [1], Dennis Nicuh Lozada [2] and Richard Esten Mason [1,*]

[1]   Department of Crop, Soil, and Environmental Sciences, University of Arkansas, 115 Plant Sciences Building, Fayetteville, AR 72701, USA

[2]   Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164, USA

*   Correspondence: esten@uark.edu; Tel.: +1-479-575-5725

check for
updates

Abstract: In order to meet the goal of doubling wheat yield by 2050, breeders must work to improve breeding program efficiency while also implementing new and improved technologies in order to increase genetic gain. Genomic selection (GS) is an expansion of marker assisted selection which uses a statistical model to estimate all marker effects for an individual simultaneously to determine a genome estimated breeding value (GEBV). Breeders are thus able to select for performance based on GEBVs in the absence of phenotypic data. In wheat, genomic selection has been successfully implemented for a number of key traits including grain yield, grain quality and quantitative disease resistance, such as that for *Fusarium* head blight. For this review, we focused on the ways to modify genomic selection to maximize prediction accuracy, including prediction model selection, marker density, trait heritability, linkage disequilibrium, the relationship between training and validation sets, population structure, and training set optimization methods. Altogether, the effects of these different factors on the accuracy of predictions should be thoroughly considered for the successful implementation of GS strategies in wheat breeding programs.

1. Introduction

According to the United Nations Department of Economic and Social Affairs, the global population is currently expected to grow from 7.7 to approximately 9.7 billion people by 2050 and 10.9 billion by 2100 [1]. This rapid population growth is expected to exceed the rate of global food production, posing the risk of a potential food crisis in the next few decades. While wheat (*Triticum aestivum* L.) yield has increased rapidly since the 1950s, largely due to improved breeding and agronomic practices, such as the introduction of improved pesticides and synthetic fertilizers, the current rate is insufficient in meeting future caloric demands and in some cases remains stagnant [2]. It has been estimated that an approximate yield improvement of 2.4 percent per year, compared to the current 0.9 percent, is needed in order to double production of wheat by 2050 [3].

To meet future demands, breeders have looked toward new technologies, such as marker-assisted selection (MAS), phenomics, trans and cisgenic approaches, and the implementation of genomic selection (GS) [4–6]. Genomic selection is a modified form of MAS. The GS process begins with a panel of genotypes, referred to as a training population, that have been genotyped using genome wide markers and phenotyped for trait(s) of interest. The training population is then used to train a model to calculate genome estimated breeding values (GEBVs) for a panel of genotypes, the validation population or new breeding lines that have only been genotyped. The breeder can then use the calculated GEBVs to make selections from the validation population without the need for phenotyping [5,7]. GEBV is

based on the effects of all markers, and can therefore capture more genetic variation for traits under selection, making GS more advantageous than traditional MAS [8]. There are many ways to increase accuracy and maximize genetic gain by modifying the methods used for GS. For this review we will focus on some of these methods as they pertain to efforts in wheat and other crops, including prediction model selection, marker density, trait heritability, linkage disequilibrium, the relationship between training and validation sets, population structure, and training set optimization methods (see Table 1).

2. The Breeder's Equation and Genetic Gain

The breeder's equation ($R = i r \sigma_A$) has long been used to show how genetic response (R) changes in response to selection intensity (i), the square root of the additive genetic variance (σ_A), and selection accuracy, which is equivalent to narrow sense heritability (h^2) in phenotypic selection (r). If plant breeders want to increase the genetic gain of their breeding program, they must increase at least one of the three components of the breeder's equation within a given amount of time (t) in a breeding cycle or year ($R = i r \sigma_A / t$) [9,10].

It can be difficult to maximize genetic gains using only phenotypic selection. While phenotypic selection can be effective for traits with high heritability, selection for low heritability traits are often relegated to the latter stages of a breeding cycle, particularly with inbred crops [11]. Another problem with phenotypic selection is that environmental interactions can mask the expression of certain traits making selection difficult. The time between generations can also be long, as some traits require a full season in order to be expressed. The limitations of phenotypic selection can hamper the increase in genetic gain in a breeding program [10]. One of the ways to maximize genetic gain within a given breeding cycle is to optimize selection accuracy through the use of molecular genetics, through approaches such as MAS or GS [4,5].

3. Marker Assisted Selection

One of the technologies breeders have used to increase genetic gain is marker-assisted selection (MAS). In MAS, molecular markers closely linked or co-segregating with a desired trait can be used to differentiate breeding lines based on the allelic variation underlying a trait [11,12]. Practitioners of MAS have been able to report genetic gains twice that of phenotypic selection [13–15]. MAS has several advantages compared to phenotypic selection, some of which include (1) the ability to select for traits that are difficult, expensive, or time consuming to phenotype (2) the ability to select for traits that have low heritability and low expression and are therefore difficult to phenotype, (3) the ability to select for traits where phenotyping is dependent on specific environmental conditions or growth stages [16], and (4) the ability to select for monogenic or qualitative traits, making it an ideal selection strategy for introgression and backcrossing of single genes into germplasm along with pyramiding of disease resistance genes [4,16]. Strategies have also been developed so that multiple traits can be selected simultaneously via MAS [16,17].

While MAS has helped to revolutionize plant breeding and provide a source for increasing genetic gain, there are also downsides. One of the primary downsides to MAS is that it is less effective when screening for more complex or multi-genic quantitative traits [5,18,19]. In order to discover new quantitative trait loci (QTL) associated with a trait of interest, biparental mapping populations must be developed. The problem with biparental mapping populations is that they rarely account for the allelic diversity and genetic background of a full breeding program and their effects tend to vary across environments, particularly for quantitative traits. Therefore, multiple mapping populations must be developed for specific environments and breeding programs in order to validate the position and allelic effects of new QTL which is both expensive and time consuming [5,16]. Other mapping approaches such as genomewide association studies (GWAS), which rely on linkage disequilibrium to identify marker-trait associations in diverse germplasm, could serve as an alternative to biparental analyses; however, it is also limited by the fact that complex traits are controlled by multiple loci with small effects. If a breeder attempts to adopt small effect QTL into their breeding program through

MAS prior to validating said QTL using local germplasm, they could achieve genetic gains that are lower than conventional phenotypic selection [20]. Cost can especially limit the wide-spread use of MAS in a breeding program, especially when considering a genome as complex as wheat [21]. Altogether, the genetic architecture of complex traits, and the costs and resources needed to implement genetic mapping methods make other approaches such as GS an ideal complementary tool for breeding and selection.

Table 1. Modifications to genomic selection and impact on the breeder equation and accuracy.

Factor	Factor on Breeder's Equation Impacted	Effect on Selection Accuracy (*r*)	Consideration for Breeding
Marker density	Selection accuracy	Varies	When LD is high a small number of markers are needed
Choice of prediction model		Varies	Models vary in their ability to account for larger effect QTL and epistatic interactions
Training population size	Genetic variance	Increased r for increase in training population size	Greater diversity requires a larger training population size
Trait heritability	Selection accuracy	Varies	In the absence of large effect QTL, accuracy generally increases with trait heritability
Genetic relationship between training and validation	Selection accuracy	Increased r as more genetically related populations or families are used	Should use genetically related training and validation populations
Population structure	Selection accuracy	Varies	Sub-populations can lead to erroneous predictions
Training population composition	Selection accuracy	Varies	Should use a training population that captures genetic relationships with the validation population
Genotype by environment interaction	Additive genetic variance	Increased r when GxE effects fitted in the model	GxE together with pedigree and marker information could help improve accuracy
Multiple traits	Additive genetic variance	Increased r when multiple traits are fitted in the model	Accounts for additional genetic variance
Use of GWAS results	Additive genetic variance	Increased r for incorporation of GWAS results in the model (for most empirical studies)	Accounts for large effect QTL
Multiple generations per year	Breeding cycle time	Varies; r affected by one or a combination of factors above	Increases gain per time by reducing cycle time
Early generation parent selection	Breeding cycle time, selection intensity	Varies; r affected by one or a combination of factors above	Increases gain per time by expediting the recombination of favorable alleles
Selection intensity	Selection accuracy	Varies; r affected by one or a combination of factors above	Genomic selection allows more lines to be characterized and thus increased in selection intensity
Additive genetic variance	Selection accuracy	Varies; r affected by one or a combination of factors above	High additive genetic variance for increased gains

4. Impact of Genomic Selection on the Breeder's Equation

When considering the breeder's equation, GS has the potential to reduce the length of a breeding cycle by selecting on the basis of GEBVs as opposed to measured phenotypes, allowing for greater genetic gain in one cycle of GS compared to one cycle of phenotypic selection [5]. Two rounds of GS were performed in the time it took to perform one round of phenotypic selection when selecting for resistance to stem rust in wheat [22]. When GS was compared with MAS in maize (*Zea mays* L.), genetic response was between 18 and 43 percent greater in the GS population compared to the MAS population [19]. Differences in response to selection per cycle between GS, MAS, and phenotypic selection were minimal in oats (*Avena sativa* L.); however the ability of GS to make two generations of selection in the time it took to make one generation of selections for phenotypic and MAS made it favorable for maximizing genetic gain [23]. The best way to determine prediction accuracy in GS is by performing a Pearson correlation between the GEBVs and true breeding values or phenotypic values. In the context of the breeder's equation, the prediction accuracy is proportional to genetic gain

(R), where the prediction accuracy is equivalent to the selection accuracy (r) [9]. The accuracy of GS is affected by a combination of one or more factors as discussed below.

5. Different Factors Affecting the Accuracy of Genomic Selection

5.1. Marker Density

Since GS requires dense marker coverage across the entire genome, a proper genotyping platform must be considered [7]. Several genomewide platforms have been implemented in wheat. The first of these were restriction fragment length polymorphism (RFLP) markers [24], followed by amplified fragment length polymorphism (AFLP) markers [25], simple sequence repeat (SSR) markers [26], diversity arrays technology (DArT) [27], and single nucleotide polymorphisms (SNP), as part of the 9K and 90K Infinium arrays [28,29]. More recently, genotyping has been performed using SNP markers through genotyping by sequencing (GBS) platforms [30,31]. GBS has become the dominant genotyping platform for GS in wheat, particularly due to its low cost, high coverage, and reduced sampling bias compared to SNP arrays [32,33].

There is evidence that marker density only marginally impacts prediction accuracy and that the use of low to medium density marker datasets could be more cost-effective for GS for *Fusarium* head blight resistance in wheat [34]. Large numbers of markers can also result in model overfitting, resulting in lower prediction accuracies when predicting independent datasets [35]. In other plant species such as Asian rapeseed (*Brassica napus* L.), the use of lower density panels enabled high prediction accuracies for populations with strong LD comparable to those achieved using high-density marker sets [36].

Higher marker density can also be favorable for GS, as a larger amount of markers can increase the probability that markers will be in LD with QTL [5]. Higher marker density can also result in lower linkage disequilibrium (LD), significantly improving prediction accuracies [37]. In most cases lower LD can result in a larger recombination frequency and more accurate estimates of QTL effects [38]. Low LD combined with higher marker densities and larger training population sizes can strongly improve prediction accuracy as well [39].

In a study predicting the performance of agronomic traits in biparental maize and barley (*Hordeum vulgare* L.) populations, increasing the number of markers improved prediction accuracy. However, once the genome was sufficiently saturated with one marker for every 12.5 centimorgan (cM), the gain in prediction accuracy plateaued. This was also observed in mixed wheat and barley populations, where prediction accuracy plateaued after reaching a moderate marker density (2 cM in barley and 4.5 cM in wheat) [39]. While lower marker density might be adequate for GS, the cost of genotyping and phenotyping large training populations is not always sustainable for breeding programs. As a result, smaller training populations with higher marker densities are appropriate, especially in the case of biparental populations [40].

5.2. Prediction Models

Several statistical approaches have been proposed for predicting GEBVs and these methods differ based on their assumptions and treatment of marker effects. GS models include parametric, semi-parametric and non-parametric approaches [5,7,41]. The parametric methods include ridge regression-best linear unbiased prediction (RR-BLUP) and Bayesian methods which use a shrinkage procedure [5]. In RR-BLUP, the most commonly utilized model, all the marker effects are estimated simultaneously assuming that all marker effects are normally distributed (Gaussian distribution) with an equal variance [7]. The RR-BLUP approach treats marker effects as random effects and shrinks equally all QTL effects towards zero, thus it can capture minor effects using a large number of markers (minor QTLs). As RR-BLUP calculates GEBVs by contracting all marker effects equally, it can lead to an underestimation of large-effect QTL. To overcome this limitation, Bayesian methods have been proposed using marker-specific shrinkage effects [7]. A Bayesian approach assumes that the variance of the marker effects is different and is estimated by using a prior distribution for this variance [7].

Bayes A (Bayesian shrinkage regression) utilizes a scaled inverse chi-square distribution, which uses all markers, which are assumed to have unequal variances. Bayes B (Bayesian variable selection) uses an inverse chi-square distribution for marker effects and some markers are assumed to have a variance of zero. Bayes B presumes that a small number of large effect loci controls the trait. Another alternative Bayesian approach is Bayes C_π which combines all markers with non-zero effects and estimates a common variance for them [42]. Bayesian LASSO (least absolute shrinkage and selection operator) has an exponential prior on the marker variances, giving a double exponential (Laplace) distribution for the marker effects resulting in less and more shrinkage on large- and small-effect markers [43]. The semi parametric and non-parametric methods used in GS include kernel regression and reproducing kernel Hilbert spaces (RKHS), and random forest regression (RF) methods. Kernel regression is based on genetic distance, which regresses marker effects to a smoothing parameter to control the distribution of the QTL effects [44]. Random forest [45] is a machine learning method that captures nonlinear relations between phenotypes and marker genotypes by building a nonlinear prediction model. The advantage of semi- and non-parametric methods is that both methods capture non-additive effects.

When using smaller training sets, the RR-BLUP model outperformed the Bayes-$C\pi$ model [40]. This is likely due to the fact that a larger number of allelic observations are needed in order to predict small effect QTL, which is irrelevant for RR-BLUP since all allelic effects are assumed to be the same [46,47]. Epistatic models improved predictions in maize, thus RKHS and RF may hold advantage in this situation [48]. When compared with 10 other Bayesian and BLUP GS models using wheat, maize, and barley datasets, the RKHS often had much higher accuracy than other models. However, in the same analysis, RKHS over fitted more than the other models, in that too many irrelevant, low impact loci, were incorporated in the model leading to exaggerated genetic variances and trait heritability [49,50]. Another study in maize showed that RKHS outperformed Bayesian-LASSO for traits where epistasis was more relevant, whereas it was outperformed for more additive traits [51]. When RF was compared with RR-BLUP, Bayes A, Bayes B, and RKHS models, it consistently had the lowest prediction accuracies for seven drought tolerance traits in maize [52]. When the RF model was compared with 10 other models for wheat, maize, and barley predictions, it performed fairly well with good prediction accuracies and low computing time [49]. The RF model was also compared with Bayesian LASSO, Bayes B, Bayes $C\pi$, and RR-BLUP while predicting the performance of four agronomic traits in chickpea (*Cicer arietinum* L.). While the RF model was the best fitting model during three location-years, there was not a significant difference between any of the models [53].

5.3. Training Population Size

One of the important factors for implementing GS in a breeding program is determining the appropriate size of a training population. The goal for a breeding program would be to develop a training population that can maximize prediction accuracy while limiting population size as much as possible, largely due to expense of genotyping and phenotyping a large number of genotypes for the training population [35].

When training population size was tested for four different yield traits in wheat, prediction accuracy increased as the number of genotypes increased from 250 to 2000 genotypes; glaucousness increased from 0.78 to 0.92, grain yield increased from 0.72 to 0.85, thousand-kernel weight increased from 0.66 to 0.86, and relative maturity increased from 0.59 to 0.82 [49]. Even though prediction accuracy increased as the training population size increased, the gain in accuracy plateaued after 2000 genotypes [37]. When evaluating prediction accuracies for nine different end-use quality traits in biparental wheat populations, the increasing population size from 24 to 96 genotypes increased prediction accuracies as well [40]. It has also been observed that deviations between prediction accuracies for different traits are larger under smaller training population sizes, likely due to the effect of relatedness between the training population and validation set [33]. Smaller training populations also run the risk of overestimating the genotypic effect when predicting larger validation sets [35]. Training population sizes between 25 and 300 were evaluated for both wheat and rice (*Oryza sativa*

L.) for five and four agronomic traits, respectively. In wheat, prediction accuracies were highest at 300 genotypes for all five traits; however, there was evidence of a plateau after 300. In rice, prediction accuracies plateaued after 175 genotypes for florets per panicle and protein content, while there was actually a decrease in prediction accuracy for plant height and flowering time after 150 genotypes [54]. More diverse training populations usually need to be larger in order to account for the larger genetic diversity, particularly with low heritability traits [55].

5.4. Trait Heritability

Generally, as heritability increases for a trait, so does the predictive ability for the trait [35]. The impact of heritability on prediction accuracy for training and validation sets was evaluated for 30 agronomic, quality, and stress tolerance traits in rice and three quality traits in maize. Overall, prediction accuracy improved as heritability increased in both the training and validations sets, with a greater impact in training sets [56]. Heritability was also evaluated for mixed populations of barley and wheat, along with biparental populations of maize and barley. In all cases, prediction accuracy increased as heritability in the training populations increased [39]. A strong correlation between high heritability and stronger prediction accuracies was also observed when predicting quality traits in bi-parental wheat populations [40]. Even with higher heritability, certain traits may have lower prediction accuracies as the trait could be controlled by a larger number of small-effect genes, such as grain yield, as opposed to a small number of large-effect genes, such as grain dry weight, as evidenced when evaluating a maize test-cross population [57].

5.5. Genetic Relationship between Training and Validation Sets

The genetic similarity between training populations and their validation sets can significantly influence prediction accuracy [35,58]. More closely related individuals will often share a common ancestor in a relatively small number of generations prior, allowing for less recombination events, preserving QTL and marker linkage-phases. Closely related individuals are also more likely to share polymorphic loci that generate genetic variation, leaving fewer opportunities for genetic drift or mutations [59]. There are also interactions that can occur between QTL and genetic backgrounds, as closely related individuals are more likely to share a portion of their genetic backgrounds compared to distant individuals [59–61]. When evaluating training populations for *Fusarium* head blight (FHB) resistance in barley, it was observed that prediction accuracy is lower when a training population is predicting the performance of barley genotypes from outside breeding programs [62]. The same has also been observed when predicting across different full-sib families in bi-parental maize populations [63]. Training populations do not need to be as large when they are predicting the performance of closely related genotypes in the validation population. Along with that, genotypes in the training population do not need to have as large of a marker density when predicting the performance of closely related genotypes [64].

5.6. Population Structure

Population structure is the composition of a population that are divided by genetic background, geography, or natural selection. As a result, it plays an important role in the estimation of GEBVs. Changes in allele frequencies between sub-populations can result in erroneous associations between markers and traits, leading to biased prediction accuracies for GS. A common indicator of population structure interference is when small training populations have high prediction accuracies [65]. Several methods can be used to reduce the influence of population structure. These include the separation of breeding lines based on their origins or traits, inclusion of kinship or genotype matrices, inclusion of marker fixed effects, or including principle components in an analysis based on subpopulations [56,58,62]. Another way to control for population structure is with epistatic genomic prediction models that can account for both additive and non-additive effects, such as the RKHS, Gaussian kernel, or exponential kernel models [66]. Another method for population structure

partitioning is the K-means algorithm, where partitioning is based on genetic similarity [37]. In the case of large training populations, another common source of population structure is LD. Particularly when two loci with differences in allele frequencies are in LD across subpopulations. This can result in spurious associations with multiple QTL [67]. When the effects of population structure were evaluated in a historic winter wheat nursery, two major subpopulations emerged based on the presence or absence of the *t2BS:2GS·2GL:2BL* translocation derived from *Triticum timopheevii*. However, when this translocation was used to design an optimal training population, it did not have significantly higher predictive ability than a randomly selected training population [68].

5.7. Retraining and Training Population Composition

Retraining of the training population is important because as multiple generations of selections are made, the generational difference between the training population and the validation set grows larger. Since the training population and the validation set become more genetically distant, the prediction accuracy is reduced. Therefore, it is important to add new genotypes to the training population in order to maintain a genetic relationship to the validation set as new germplasm is introduced to the breeding program and more recombination events occur [59]. New germplasm must be added to breeding programs implementing GS because the increased selection intensity from GS reduces the effective population size in a breeding program and thus leads to a loss in genetic variability [5].

Several different methods are used to optimize training populations in order to maximize prediction accuracy while also minimizing training population size, allowing breeders to better allocate resources. Among these are the PEVmean, which identifies the optimal training set to minimize predictive error variance; CDmean which maximizes the coefficient of determination; Gmean which maximizes genetic relationship using only marker information; and Stratified Sampling (SS) which uses random sampling to maximize the variance explained; and Selection of Training Populations by Genetic Algorithm (STPGA) which utilizes both phenotypic and marker information [69].

When the CDmean, SS, Gmean, and STPGA algorithms were tested on a barley training population in order to compare prediction accuracies for grain yield and deoxynivalenol (DON) accumulation, all four significantly improved prediction accuracy for both traits compared to a randomly selected training set [69]. The Gmean algorithm outperformed the other algorithms while predicting grain yield, while the SS algorithm was the top performing algorithm for predicting DON accumulation [69]. PEVmean and CDmean were compared with each other in two empirical studies using two different maize diversity panels. CDmean was the more reliable algorithm as it took into account the reduction of variance due to relatedness between individuals [70]. The PEVmean algorithm, as described by Akdemir et al. was compared with random and clustered training population selection methods in winter wheat [71]. The PEVmean algorithm outperformed the random and clustered methods across all population sizes, ranging from 50 to 350 genotypes [68].

6. Genomic Selection in Multiple Environments

A vast majority of the applications for GS is in single environments and most genomic prediction models do not have the predictive power to make selections across multiple environments or take genotype by environment interactions into consideration [72]. Even so, genotype by environment interactions play a large role in plant breeding [73]. The first multiple environment models incorporated a factor analytic structure in order to model for genotype by environment interactions. Multiple environment models with pedigree and marker information performed better than models without pedigree or with only one of the components when making predictions for multiple environments in wheat [72]. The incorporation of weather data and weather covariates has also been beneficial for multiple environment GS as well, improving prediction accuracy by nearly 11 percent in winter wheat [74]. The problem with incorporating marker and environmental covariate interactions in a multiple environment model is that the number of interactions becomes so large that the ability to model such interactions is impossible, especially when using a large number of markers and environmental

covariates. The use of variance components can help to reduce the computational demand of the modeling marker and environmental component interactions [75]. When comparing marker by environment GBLUP models with models that ignored a genotype by environment interaction, the marker by environment GBLUP model significantly outperformed the naïve model over three wheat datasets in seven different irrigated and dryland environments [76]. In another study, a genotype by environment (GxE) GBLUP model and a standard GBLUP model were tested over 35 location years. The standard GBLUP model outperformed the GxE model when predicting the performance of new environments, however modeling for genotype by environment interactions improved the overall predictive ability [77].

7. Genomic Selection for Multiple Traits

Most GS models only predict for single traits, however predicting for multiple traits can be advantageous, especially when looking at multiple yield components, quality traits, disease resistance traits, and abiotic stress tolerance traits. In the past, multi-trait methods have been developed for QTL mapping [78–80]. Genomic prediction models have also been developed to predict multiple traits in dairy bulls. The three models are variations on the GBLUP, Bayes SVSS, and Bayes Cπ models, where the Bayes SVSS outperformed the Bayes Cπ and GBLUP models [81]. When comparing predictive ability of GBLUP, Bayes A, and Bayes Cπ multivariate models on a pine (*Pinus* L.) breeding dataset, the Bayesian models outperformed GBLUP under a major QTL structure and the multi-trait models strongly outperformed the single trait models. Under a polygenic genetic architecture, the three multi-trait models performed roughly the same and barely outperformed the single trait models. Multiple trait models tend to work better when predicting traits that are genetically correlated with each other [82]. When predicting cassava (*Manihot esculenta* Crantz) performance, multiple-trait models outperformed single trait models by nearly 40 percent in prediction accuracy [83]. In American cranberry (*Vaccinium macrocarpon* Ait), a multi-trait GBLUP model outperformed a standard, single trait GBLUP model in scenarios with medium to high genetic correlation between traits, however there was little difference between models when genetic correlation between traits was low [84]. A training population of 557 wheat genotypes were evaluated for grain yield and three proximal or remote sensing traits and genomic prediction models were tested using a univariate model for only grain yield and a multivariate model using all four traits [80]. Multivariate model prediction accuracies for grain yield outperformed univariate models by 70 percent, on average. This indicated that the use of proximal and remote sensing data as secondary traits in genomic prediction models could improve prediction accuracy for grain yield [85].

8. Use of GWAS Results in Genomic Selection

Traditionally, QTL mapping was performed using QTL studies, which consisted of biparental mapping populations consisting of recombinant inbred lines, F_2 lines, or backcrosses. All of the individuals within the mapping population were then genotyped and phenotyped for a trait of interest, where the resulting data would be analyzed using linkage mapping in order to identify QTL associated with the trait of interest. The problem with QTL mapping is that the resolution is often low, due to low recombination within the population and small population sizes. Then, when QTL have successfully been identified within a single biparental mapping population, markers within LD of the QTL could be used to select for the trait of interest. The problem is that it is difficult to validate QTL across mapping populations, often making it nearly impossible to implement MAS using markers developed for another breeding program [5,86].

An alternative to QTL mapping is the genomewide association study (GWAS), which can be used to identify significant marker trait associations (MTA) between genomewide SNP markers and individual traits. This method relies on LD and ancestral recombination events in natural populations [87,88]. GWAS can be performed on a wide array of diverse genotypes already existing in a population [86,89]. The problem with some GWAS models is that they can often fail to account for false positives due to

population structure and relationships between individuals in the association-mapping panel [90,91]. Using population structure and other relationship parameters as covariates can reduce these false positives; however, this can result in more false negatives and longer computational times [92]. In that regard, many GWAS models that can help control for false positives or negatives when identifying MTAs have been implemented. Some of these models include the generalized linear model (GLM), mixed linear model (MLM), efficient mixed-model association (EMMA), compressed mixed linear model (CMLM), multiple locus mixed linear model (MLMM), fixed and random model circulating probability unification (FarmCPU), and Bayesian-information and linkage-disequilibrium iteratively nested keyway (BLINK) models [92–94].

Most QTL discovered through QTL mapping and GWAS related to quantitative traits, such as grain yield, disease resistance, and quality, have not been successful for traditional MAS. This is largely due the fact that many small-effect genes are controlling variation in these traits [95]. While significant MTAs identified through GWAS for these traits may not be helpful regarding MAS, they can still be applied for GS [96–98]. Most simple genomic prediction models, such as RR-BLUP or GBLUP, assume that most markers have small genetic effects; as opposed to GWAS, which assumes there are multiple markers contributing to a larger amount of genetic variation. Both methods are incorrect in their assumptions [96,99]. More complex Bayesian and non-parametric models were developed in order to account for major-effect QTL using shrinkage algorithms, and variable selection using multiple distributions for marker effects and tuning parameters [7,42]. The problem with more complex models is that they are often more computationally intensive and do not work well with larger marker datasets. Parameter estimates can also be overly sensitive to priors and often do not translate to other studies [100].

An alternative is to use markers linked to major genes associated with traits of interest as fixed effects. Several GS experiments have shown that this strategy can improve prediction accuracies in wheat and maize [68,101–104]. Another proposed method is using only GWAS data from previously published GWAS studies as marker covariates, where the inclusion of GWAS data with a GBLUP model improved prediction accuracies for two out of three traits in cattle and nine out of 11 traits in rice over a Bayes B model [105]. Other studies have included GWAS data from the training populations themselves, otherwise referred to as GS + de novo GWAS (GS+GWAS) [97].

Spindel et al. previously performed GS+GWAS in a tropical rice breeding program using the RR-BLUP model with significant markers from a GWAS as fixed effects. The GS+GWAS model outperformed the RR-BLUP with historic GWAS data, RR-BLUP, Bayesian LASSO, RKHS, Random Forest, and multiple linear regression models for all three traits; however, the GS+GWAS model had a significantly higher prediction accuracy for plant height [97]. A GS+GWAS model was used on a maize nested association mapping (NAM) population, making predictions for three traits; one was a highly polygenic trait, plant height, whereas the other two were moderately polygenic disease resistance traits. The GS+GWAS model performed significantly better for the moderately polygenic disease resistance traits, whereas there was no significant difference in prediction accuracy between the GS+GWAS models and the standard models for plant height [96]. When predicting powdery mildew resistance in winter wheat, seven different training populations were developed ranging in size from 50 to 350 genotypes, A GWAS was performed on each training population and only the most significant MTA was selected as a fixed effect for the RR-BLUP model. The GS+GWAS model significantly outperformed the standard RR-BLUP model for prediction accuracy at each training population size [68]. The inclusion of significant markers in a model, however, does not always result to improved predictions. In a recent simulation study in maize and sorghum (*Sorghum bicolor* (L.) Moench), for instance, no significant increase or a decrease in prediction accuracy was observed for a majority of the traits evaluated when GWAS-derived SNPs were included as fixed effects in the GS model. Model performance should therefore be explored on a trait-by-trait basis before its implementation in the breeding program [106].

9. Putting the Pieces Together: Genomic Selection for Wheat Grain Yield

Improving grain yield is a major goal for any wheat breeder, as it is an economic trait directly affecting farmers while also contributing to global food security. As genetic gains in wheat grain yield have stagnated recently, wheat breeders have looked at GS in order to increase genetic gain [3]. Previously, GS was implemented on 2325 European winter wheat genotypes wherein four genomic prediction models, RR-BLUP, Bayes Cπ, RKHS, and EGBLUP were compared for prediction accuracy. The epistatic models (RKHS) outperformed the additive models, indicating that accounting for epistasis improved prediction accuracy, particularly in larger populations. Forward selection was also successfully implemented with limited decrease in prediction accuracy from one year to the next [107]. In another study, prediction accuracy was evaluated for grain yield using RR-BLUP over five independent breeding cycles in a winter wheat breeding program using 659 inbred lines. Over the five cycles, grain yield prediction accuracy was $r = 0.38$, however after outlier cycles were removed, accuracy improved to $r = 0.41$. The removal of outlier environments did not have a significant impact on prediction accuracy [108]. Genomic selection was also implemented into the preliminary yield trial (PYT) stage of a wheat breeding program and compared with traditional phenotypic selection and genomic assisted selection (GAS), where breeding values from the PYT stage were combined with GEBVs using a heritability index. The resulting prediction accuracies for all three methods for grain yield were 0.39, 0.33, and 0.48 respectively. This indicated that the use of heritability indices in order to supplement GS can increase prediction accuracy for grain yield in wheat [109]. Different training population sizes were evaluated for the prediction of grain yield in winter wheat. It was observed that there was a high predictive ability using RR-BLUP when a training population was selected using PEVmean for grain yield. There was no significant improvement in prediction accuracy as training population size increased above 50 genotypes, with the exception of a reduction in prediction accuracy at 100 genotypes. Prediction accuracies for grain yield averaged 0.64 with a training population size of 350 genotypes, selected using PEVmean [68].

10. Putting the Pieces Together: Genomic Selection for *Fusarium* Head Blight Resistance

Fusarium head blight (FHB) is a destructive fungal disease that negatively affects wheat production worldwide, resulting in heavy losses in grain yield and quality. In the United States, the most common causal pathogen is *Fusarium graminearum*, which is part of the phylum *Ascomycota* [110,111]. Symptoms on infected wheat plants include premature bleaching of individual spikelets shortly after flowering, eventually progressing through the entire spike, resulting in a fully bleached spike. Signs of *F. graminearum* are usually present during warm and moist conditions; they appear as salmon colored sporodochia on the rachis and glumes of the spikelets. Blue-black spherical perithecia, which serve as sexual structures for the pathogen, will often appear later in the growing season. As symptoms progress, the fungal mycelium will colonize the wheat kernels as they develop, resulting in shriveled grain with a pink to light-pink coloration. These damaged kernels are often referred to as 'tombstone-kernels' or '*Fusarium* damaged kernels' [112,113].

The *F. graminearum* pathogen produces the mycotoxin, deoxynivalenol (DON), in order to disable natural plant defenses. DON is considered a vomitoxin because it disrupts the digestive function of humans and animals that consume infected grain, resulting in nausea, headaches, vomiting, and in extreme cases, death. DON levels in human food should not exceed 1 ppm, however grain infected with FHB can exceed 20 ppm [114]. Any grain that does not meet the DON limits set by either the USDA or individual buyers can have the value of their grain docked or rejected entirely, causing significant economic damages for farmers [115].

Wheat breeders have worked to develop varieties that have genetic resistance to FHB. FHB resistance is a quantitative trait that is significantly influenced by genotype by environment interactions [116]. There are generally two major sources for FHB resistance in the United States, either through 'exotic' or 'native' germplasm. One of the most influential exotic genotypes, 'Sumai-3' was identified in China [117]. Several major effect QTL that have been validated for FHB resistance. The

first three identified include *Fhb1*, *Fhb2*, and *Qfhs.ifa-5A*, found in 'Sumai-3' [118–122]. The next two major QTL, *Fhb4* and *Fhb5* were found in the genotype 'Wangshiubai' [123,124]. *Qfhs.nau-2DL* was then found in the breeding line CJ9306 [125,126]. The most recently discovered QTL was *Fhb7* found in *Thinopyrum ponticum* [127]. The two primary forms of FHB resistance which exist in field conditions are Type I resistance or the resistance to initial infection by *F. graminearum*, and Type II resistance or the resistance to fungal spread within the infected spike [116]. The *Fhb1* and *Qfhs.nau-2DL* QTL primarily confer Type II resistance while also reducing DON accumulation, whereas *Qfhs.ifa-5A* primarily confers Type I resistance and reduces DON accumulation [128–131].

As FHB resistance in wheat is a complex, quantitative trait, it is an ideal candidate for GS. Previous studies explored the potential of GS to predict resistance to FHB in wheat under different subsets of markers, mapping populations, and prediction models and scenarios. A wheat panel of 322 genotypes was previously used to predict the performance of six traits associated with FHB resistance. Four genomic prediction models (RR-BLUP, Bayesian-LASSO, RKHS, and Random Forest) across two different marker sets (whole-genome versus a marker subset associated with FHB resistance) were compared. Random Forest and RKHS models had the highest prediction accuracies for most of the traits. In the case of DON, the Random Forest model plus the targeted and whole-genome markers had the highest prediction accuracy [132]. Hoffstetter et al. found that prediction accuracy for FHB resistance improved with the use of marker and training subsets [133]. Another study for predicting FHB resistance using six traits tested with three models: RR-BLUP, elastic net, and LASSO where different marker densities were tested between 500 and 4500 SNPs along with training population sizes between 96 and 2018 genotypes. Overall, there was a high prediction accuracy for FHB resistance. The top performing model was RR-BLUP, whereas the marker sets having the highest accuracies ranged between 1500 and 3000 SNPs. Depending on the trait, the optimum training population was less than 192 genotypes. [134]. Another study on spring wheat genotypes in the Pacific Northwest United States used a population of 170 genotypes. FHB incidence was observed to have the highest prediction accuracies and genotypes that were similarly related to the initial training population had a prediction accuracy of 0.60. The prediction accuracy was fairly high for FHB resistance overall, suggesting that GS could work well for FHB resistance [135].

11. Conclusions

In GS, all markers, including those with minor effects are included in the model simultaneously, allowing breeders to use GEBVs to make selections on performance when limited or no phenotypic data are available. This is especially important for quantitative traits conferred by a large number of genes each with a minor effect, which, in the workflow of a traditional breeding program, are not evaluated until the later generations. The availability of inexpensive genotyping technology makes GS a feasible and attractive approach. In terms of influencing the breeder's equation ($R = ir\sigma_A/t$), GS can increase genetic gain by decreasing the breeding cycle time and by increasing selection accuracy and intensity in early generations. In wheat, studies have shown the potential for positive genetic gains from GS for quantitative traits such as grain yield, quality and disease resistance. Future work should focus on empirical studies that validate prediction accuracy in new breeding material in order to determine the efficacy of GS as replacement for phenotypic selection, where applicable. As different factors affect the accuracy of GS, thoughtful considerations based on these parameters, the traits being evaluated, and the available resources should result in to a successful implementation of these approaches in wheat breeding programs.

References

1. United Nations; Department of Economic and Social Affairs. *Population Division World Population Prospects 2019: Highlights*; Department of Economic and Social Affairs: New York, NY, USA, 2019.

2. Graybosch, R.A.; Peterson, C.J. Genetic Improvement in Winter Wheat Yields in the Great Plains of North America, 1959–2008. *Crop Sci.* **2010**, *50*, 1882. [CrossRef]

3. Ray, D.K.; Mueller, N.D.; West, P.C.; Foley, J.A. Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS ONE* **2013**, *8*, e66428. [CrossRef] [PubMed]

4. Holland, J.B. Implementation of molecular markers for quantitative traits in breeding programs—Challenges and opportunities. In Proceedings of the 4th International Crop Science Congress, Brisbane, Australia, 26 September–1 October 2004.

5. Heffner, E.L.; Sorrells, M.E.; Jannink, J.-L. Genomic Selection for Crop Improvement. *Crop Sci.* **2009**, *49*, 1–12. [CrossRef]

6. Houle, D.; Govindaraju, D.R.; Omholt, S. Phenomics: The next challenge. *Nat. Rev. Genet.* **2010**, *11*, 855. [CrossRef] [PubMed]

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

8. Newell, M.A.; Jannink, J.-L. Genomic Selection in Plant Breeding. In *Crop Breeding*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 117–130.

9. Frankham, R. *Introduction to Quantitative Genetics*, 4th ed.; Falconer, D.S., Mackay, T.F.C., Eds.; Elsevier: Amsterdam, The Netherlands, 1996; ISBN 0582 24302 5.

10. Xu, Y.; Li, P.; Zou, C.; Lu, Y.; Xie, C.; Zhang, X.; Prasanna, B.M.; Olsen, M.S. Enhancing genetic gain in the era of molecular breeding. *J. Exp. Bot.* **2017**, *68*, 2641–2666. [CrossRef] [PubMed]

11. Collard, B.C.; Mackill, D.J. Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. *Philos. Trans. R. Soc. B Boil. Sci.* **2007**, *363*, 557–572. [CrossRef] [PubMed]

12. Mohan, M.; Nair, S.; Bhagwat, A.; Krishna, T.G.; Yano, M.; Bhatia, C.; Sasaki, T. Genome mapping, molecular markers and marker-assisted selection in crop plants. *Mol. Breed.* **1997**, *3*, 87–103. [CrossRef]

13. Crosbie, T.M.; Eathington, S.R.; Johnson Sr, G.R.; Edwards, M.; Reiter, R.; Stark, S.; Mohanty, R.G.; Oyervides, M.; Buehler, R.E.; Walker, A.K. *Plant Breeding: Past, Present, and Future*; Wiley Online Library: Hoboken, NJ, USA, 2006; pp. 3–50.

14. Ragot, M.; Lee, M.; Guimaraes, E. Marker-assisted selection in maize: Current status, potential, limitations and perspectives from the private and public sectors. In *Marker-Assisted Selection, Current Status and Future Perspectives in Crops, Livestock, Forestry and Fish*; Food and Agriculture Organization of the United Nations (FAO): Rome, Italy, 2007; pp. 117–150.

15. Knapp, S.J. Marker-Assisted Selection as a Strategy for Increasing the Probability of Selecting Superior Genotypes. *Crop Sci.* **1998**, *38*, 1164. [CrossRef]

16. Xu, Y.; Crouch, J.H. Marker-Assisted Selection in Plant Breeding: From Publications to Practice. *Crop Sci.* **2008**, *48*, 391. [CrossRef]

17. Ragot, M.; Gay, G.; Muller, J.-P.; Durovray, J. *Efficient Selection for the Adaptation to the Environment through QTL Mapping and Manipulation in Maize*; Food and Agriculture Organization of the United Nations (FAO): Rome, Italy, 2000; pp. 128–130.

18. Hospital, F. Challenges for effective marker-assisted selection in plants. *Genetica* **2009**, *136*, 303–310. [CrossRef] [PubMed]

19. Bernardo, R.; Yu, J. Prospects for Genomewide Selection for Quantitative Traits in Maize. *Crop Sci.* **2007**, *47*, 1082. [CrossRef]

20. Bernardo, R. What If We Knew All the Genes for a Quantitative Trait in Hybrid Crops? *Crop Sci.* **2001**, *41*, 1–4. [CrossRef]

21. William, H.M.; Trethowan, R.; Crosby-Galvan, E.M. Wheat breeding assisted by markers: CIMMYT's experience. *Euphytica* **2007**, *157*, 307–319. [CrossRef]

22. Rutkoski, J.; Singh, R.; Huerta-Espino, J.; Bhavani, S.; Poland, J.; Jannink, J.; Sorrells, M. Genetic Gain from Phenotypic and Genomic Selection for Quantitative Resistance to Stem Rust of Wheat. *Plant Genome* **2015**, *8*. [CrossRef]

23. Asoro, F.G.; Newell, M.A.; Beavis, W.D.; Scott, M.P.; Tinker, N.A.; Jannink, J.-L. Genomic, Marker-Assisted, and Pedigree-BLUP Selection Methods for β-Glucan Concentration in Elite Oat. *Crop Sci.* **2013**, *53*, 1894. [CrossRef]

24. Chao, S.; Sharp, P.J.; Worland, A.J.; Warham, E.J.; Koebner, R.M.D.; Gale, M.D. RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. *Theor. Appl. Genet.* **1989**, *78*, 495–504. [CrossRef] [PubMed]

25. Barrett, B.A.; Kidwell, K.K. AFLP-Based Genetic Diversity Assessment among Wheat Cultivars from the Pacific Northwest. *Crop Sci.* **1998**, *38*, 1261–1271. [CrossRef]

26. Röder, M.S.; Korzun, V.; Wendehake, K.; Plaschke, J.; Tixier, M.-H.; Leroy, P.; Ganal, M.W. A microsatellite map of wheat. *Genetics* **1998**, *149*, 2007–2023.

27. Akbari, M.; Wenzl, P.; Caig, V.; Carling, J.; Xia, L.; Yang, S.; Uszyński, G.; Mohler, V.; Lehmensiek, A.; Kuchel, H.; et al. Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theor. Appl. Genet.* **2006**, *113*, 1409–1420. [CrossRef] [PubMed]

28. Chao, S.; Dubcovsky, J.; Dvorak, J.; Luo, M.-C.; Baenziger, S.P.; Matnyazov, R.; Clark, D.R.; Talbert, L.E.; Anderson, J.A.; Dreisigacker, S. Population-and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (Triticum aestivum L.). *BMC Genom.* **2010**, *11*, 727. [CrossRef] [PubMed]

29. Wang, S.; Wong, D.; Forrest, K.; Allen, A.; Chao, S.; Huang, B.E.; Maccaferri, M.; Salvi, S.; Milner, S.G.; Cattivelli, L.; et al. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* **2014**, *12*, 787–796. [CrossRef] [PubMed]

30. Poland, J.A.; Brown, P.J.; Sorrells, M.E.; Jannink, J.-L. Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach. *PLoS ONE* **2012**, *7*, e32253. [CrossRef] [PubMed]

31. Saintenac, C.; Jiang, D.; Wang, S.; Akhunov, E. Sequence-Based Mapping of the Polyploid Wheat Genome. *G3 Genes Genomes Genet.* **2013**, *3*, 1105–1114. [CrossRef] [PubMed]

32. Heslot, N.; Rutkoski, J.; Poland, J.; Jannink, J.-L.; Sorrells, M.E. Impact of Marker Ascertainment Bias on Genomic Selection Accuracy and Estimates of Genetic Diversity. *PLoS ONE* **2013**, *8*, e74612. [CrossRef] [PubMed]

33. Poland, J.; Endelman, J.; Dawson, J.; Rutkoski, J.; Wu, S.; Manès, Y.; Dreisigacker, S.; Crossa, J.; Sanchez-Villeda, H.; Sorrells, M.; et al. Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. *Plant Genome* **2012**, *5*, 103. [CrossRef]

34. Jiang, Y.; Zhao, Y.; Rodemann, B.; Plieske, J.; Kollers, S.; Korzun, V.; Ebmeyer, E.; Argillier, O.; Hinze, M.; Ling, J. Potential and limits to unravel the genetic architecture and predict the variation of Fusariumhead blight resistance in European winter wheat (Triticum aestivum L.). *Heredity* **2015**, *114*, 318. [CrossRef] [PubMed]

35. Jannink, J.-L.; Lorenz, A.J.; Iwata, H. Genomic selection in plant breeding: From theory to practice. *Brief. Funct. Genom.* **2010**, *9*, 166–177. [CrossRef] [PubMed]

36. Werner, C.R.; Voss-Fels, K.P.; Miller, C.N.; Qian, W.; Hua, W.; Guan, C.-Y.; Snowdon, R.J.; Qian, L. Effective genomic selection in a narrow-genepool crop with low-density markers: Asian rapeseed as an example. *Plant Genome* **2018**, *11*. [CrossRef] [PubMed]

37. Norman, A.; Taylor, J.; Edwards, J.; Kuchel, H. Optimising Genomic Selection in Wheat: Effect of Marker Density, Population Size and Population Structure on Prediction Accuracy. *G3 Genes Genomes Genet.* **2018**, *8*, 2889–2899. [CrossRef]

38. Huang, B.E.; George, A.W.; Forrest, K.L.; Kilian, A.; Hayden, M.J.; Morell, M.K.; Cavanagh, C.R. A multiparent advanced generation inter-cross population for genetic analysis in wheat. *Plant Biotechnol. J.* **2012**, *10*, 826–839. [CrossRef] [PubMed]

39. Combs, E.; Bernardo, R. Accuracy of Genomewide Selection for Different Traits with Constant Population Size, Heritability, and Number of Markers. *Plant Genome* **2013**, *6*. [CrossRef]

40. Heffner, E.L.; Jannink, J.-L.; Iwata, H.; Souza, E.; Sorrells, M.E. Genomic Selection Accuracy for Grain Quality Traits in Biparental Wheat Populations. *Crop Sci.* **2011**, *51*, 2597. [CrossRef]

41. Lorenz, A.J.; Chao, S.M.; Asoro, F.G.; Heffner, E.L.; Hayashi, T.; Iwata, H.; Smith, K.P.; Sorrells, M.E.; Jannink, J.L. Genomic selection in plant breeding: knowledge and prospects. *Adv. Agron.* **2011**, *110*, 77–123.

42. Habier, D.; Fernando, R.L.; Kizilkaya, K.; Garrick, D.J. Extension of the bayesian alphabet for genomic selection. *BMC Bioinform.* **2011**, *12*, 186. [CrossRef] [PubMed]

43. Park, T.; Casella, G. The Bayesian Lasso. *J. Am. Stat. Assoc.* **2008**, *103*, 681–686. [CrossRef]

44. Gianola, D.; Fernando, R.L.; Stella, A. Genomic-Assisted Prediction of Genetic Value with Semiparametric Procedures. *Genetics* **2006**, *173*, 1761–1776. [CrossRef]

45. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

46. Endelman, J.B. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *Plant Genome* **2011**, *4*, 250. [CrossRef]

47. Gilmour, A. Mixed model regression mapping for QTL detection in experimental crosses. *Comput. Stat. Data Anal.* **2007**, *51*, 3749–3764. [CrossRef]

48. Dudley, J.; Johnson, G. Epistatic Models Improve Prediction of Performance in Corn. *Crop Sci.* **2009**, *49*, 1533. [CrossRef]

49. Sorrells, M.E.; Jannink, J.-L.; Heslot, N.; Yang, H.-P. Genomic Selection in Plant Breeding: A Comparison of Models. *Crop Sci.* **2012**, *52*, 146–160.

50. Jia, Z. Controlling the Overfitting of Heritability in Genomic Selection through Cross Validation. *Sci. Rep.* **2017**, *7*, 13678. [CrossRef]

51. Crossa, J.; de Los Campos, G.; Pérez, P.; Gianola, D.; Burgueño, J.; Araus, J.L.; Makumbi, D.; Singh, R.; Dreisigacker, S.; Yan, J. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* **2010**, *186*, 713–724. [CrossRef] [PubMed]

52. Shikha, M.; Kanika, A.; Rao, A.R.; Mallikarjuna, M.G.; Gupta, H.S.; Nepolean, T. Genomic Selection for Drought Tolerance Using Genome-Wide SNPs in Maize. *Front. Plant Sci.* **2017**, *8*, 63. [CrossRef] [PubMed]

53. Roorkiwal, M.; Rathore, A.; Das, R.R.; Singh, M.K.; Jain, A.; Srinivasan, S.; Gaur, P.M.; Chellapilla, B.; Tripathi, S.; Li, Y. Genome-enabled prediction models for yield related traits in chickpea. *Front. Plant Sci.* **2016**, *7*, 1666. [CrossRef]

54. Isidro, J.; Jannink, J.-L.; Akdemir, D.; Poland, J.; Heslot, N.; Sorrells, M.E. Training set optimization under population structure in genomic selection. *Theor. Appl. Genet.* **2015**, *128*, 145–158. [CrossRef] [PubMed]

55. Mujibi, F.; Nkrumah, J.; Durunna, O.; Stothard, P.; Mah, J.; Wang, Z.; Basarab, J.; Plastow, G.; Crews, D., Jr.; Moore, S. Accuracy of genomic breeding values for residual feed intake in crossbred beef cattle. *J. Anim. Sci.* **2011**, *89*, 3353–3361. [CrossRef] [PubMed]

56. Guo, Z.; Tucker, D.M.; Basten, C.J.; Gandhi, H.; Ersoz, E.; Guo, B.; Xu, Z.; Wang, D.; Gay, G. The impact of population structure on genomic prediction in stratified populations. *Theor. Appl. Genet.* **2014**, *127*, 749–762. [CrossRef]

57. Albrecht, T.; Wimmer, V.; Auinger, H.-J.; Erbe, M.; Knaak, C.; Ouzunova, M.; Simianer, H.; Schön, C.-C. Genome-based prediction of testcross values in maize. *Theor. Appl. Genet.* **2011**, *123*, 339–350. [CrossRef]

58. Clark, S.A.; Hickey, J.M.; Daetwyler, H.D.; Van Der Werf, J.H. The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. *Genet. Sel. Evol.* **2012**, *44*, 4. [CrossRef] [PubMed]

59. Lorenz, A.J.; Smith, K.P. Adding Genetically Distant Individuals to Training Populations Reduces Genomic Prediction Accuracy in Barley. *Crop Sci.* **2015**, *55*, 2657. [CrossRef]

60. Lorenz, K.; Cohen, B.A. Small and large effect QTL interactions underlie variation in yeast sporulation efficiency. *Genetics* **2012**, *192*, 1123–1132. [CrossRef] [PubMed]

61. Mohammadi, M.; Blake, T.K.; Budde, A.D.; Chao, S.; Hayes, P.M.; Horsley, R.D.; Obert, D.E.; Ullrich, S.E.; Smith, K.P. A genome-wide association study of malting quality across eight U.S. barley breeding programs. *Theor. Appl. Genet.* **2015**, *128*, 705–721. [CrossRef] [PubMed]

62. Lorenz, A.J.; Smith, K.; Jannink, J.-L. Potential and Optimization of Genomic Selection for Fusarium Head Blight Resistance in Six-Row Barley. *Crop Sci.* **2012**, *52*, 1609. [CrossRef]

63. Riedelsheimer, C.; Endelman, J.B.; Stange, M.; Sorrells, M.E.; Jannink, J.-L.; Melchinger, A.E. Genomic predictability of interconnected bi-parental maize populations. *Genetics* **2013**, *194*, 493–503. [CrossRef] [PubMed]

64. Meuwissen, T.H. Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. *Genet. Sel. Evol.* **2009**, *41*, 35. [CrossRef]

65. Windhausen, V.S.; Atlin, G.N.; Hickey, J.M.; Crossa, J.; Jannink, J.-L.; Sorrells, M.E.; Raman, B.; Cairns, J.E.; Tarekegne, A.; Semagn, K. Effectiveness of genomic prediction of maize hybrid performance in different breeding populations and environments. *G3 Genes Genomes Genet.* **2012**, *2*, 1427–1436. [CrossRef]

66. Sallam, A.H.; Endelman, J.B.; Jannink, J.-L.; Smith, K.P. Assessing Genomic Selection Prediction Accuracy in a Dynamic Barley Breeding Population. *Plant Genome* **2015**, *8*. [CrossRef]

67. Pritchard, J.K.; Stephens, M.; Rosenberg, N.A.; Donnelly, P. Association Mapping in Structured Populations. *Am. J. Hum. Genet.* **2000**, *67*, 170–181. [CrossRef]

68. Sarinelli, J.M.; Murphy, J.P.; Tyagi, P.; Holland, J.B.; Johnson, J.W.; Mergoum, M.; Mason, R.E.; Babar, A.; Harrison, S.; Sutton, R.; et al. Training population selection and use of fixed effects to optimize genomic predictions in a historical USA winter wheat panel. *Theor. Appl. Genet.* **2019**, *132*, 1247–1261. [CrossRef] [PubMed]

69. Tiede, T.; Smith, K.P. Evaluation and retrospective optimization of genomic selection for yield and disease resistance in spring barley. *Mol. Breed.* **2018**, *38*, 55. [CrossRef]

70. Rincent, R.; Laloë, D.; Nicolas, S.; Altmann, T.; Brunel, D.; Revilla, P.; Rodriguez, V.M.; Moreno-Gonzales, J.; Melchinger, A.E.; Bauer, E. Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: Comparison of methods in two diverse groups of maize inbreds (Zea mays L.). *Genetics* **2012**, *192*, 715–728. [CrossRef] [PubMed]

71. Akdemir, D.; Sanchez, J.I.; Jannink, J.-L. Optimization of genomic selection training populations with a genetic algorithm. *Genet. Sel. Evol.* **2015**, *47*, 342. [CrossRef] [PubMed]

72. Weigel, K.; Burgueño, J.; Campos, G.D.L.; Crossa, J. Genomic Prediction of Breeding Values when Modeling Genotype × Environment Interaction using Pedigree and Dense Molecular Markers. *Crop Sci.* **2012**, *52*, 707–719.

73. Burgueño, J.; Crossa, J.; Cotes, J.M.; Vicente, F.S.; Das, B. Prediction Assessment of Linear Mixed Models for Multienvironment Trials. *Crop Sci.* **2011**, *51*, 944. [CrossRef]

74. Heslot, N.; Akdemir, D.; Sorrells, M.E.; Jannink, J.-L. Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theor. Appl. Genet.* **2014**, *127*, 463–480. [CrossRef] [PubMed]

75. Jarquín, D.; Crossa, J.; Lacaze, X.; Du Cheyron, P.; Daucourt, J.; Lorgeou, J.; Piraux, F.; Guerreiro, L.; Pérez, P.; Calus, M. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor. Appl. Genet.* **2014**, *127*, 595–607. [CrossRef]

76. Lopez-Cruz, M.; Crossa, J.; Bonnett, D.; Dreisigacker, S.; Poland, J.; Jannink, J.-L.; Singh, R.P.; Autrique, E.; de los Campos, G. Increased prediction accuracy in wheat breeding trials using a marker× environment interaction genomic selection model. *G3 Genes Genomes Genet.* **2015**, *5*, 569–582. [CrossRef]

77. Lado, B.; Barrios, P.M.G.; Quincke, M.; Silva, P.; Gutiérrez, L. Modeling Genotype × Environment Interaction for Genomic Selection with Unbalanced Data from a Wheat Breeding Program. *Crop Sci.* **2016**, *56*, 2165. [CrossRef]

78. Jiang, C.; Zeng, Z.B. Multiple Trait Analysis of Genetic Mapping for Quantitative Trait Loci. *Genetics* **1995**, *140*, 1111–1127. [PubMed]

79. Banerjee, S.; Yandell, B.S.; Yi, N. Bayesian QTL mapping for multiple traits. *Genetics* **2008**, *179*, 2275–2289. [CrossRef] [PubMed]

80. Xu, C.; Wang, X.; Li, Z.; Xu, S. Mapping QTL for multiple traits using Bayesian statistics. *Genet. Res.* **2009**, *91*, 23–37. [CrossRef] [PubMed]

81. Calus, M.P.; Veerkamp, R.F. Accuracy of multi-trait genomic selection using different methods. *Genet. Sel. Evol.* **2011**, *43*, 26. [CrossRef] [PubMed]

82. Jia, Y.; Jannink, J.-L. Multiple-Trait Genomic Selection Methods Increase Genetic Value Prediction Accuracy. *Genetics* **2012**, *192*, 1513–1522. [CrossRef] [PubMed]

83. Okeke, U.G.; Akdemir, D.; Rabbi, I.; Kulakow, P.; Jannink, J.-L. Accuracies of univariate and multivariate genomic prediction models in African cassava. *Genet. Sel. Evol.* **2017**, *49*, 88. [CrossRef]

84. Covarrubias-Pazaran, G.; Schlautman, B.; Diaz-Garcia, L.; Grygleski, E.; Polashock, J.; Johnson-Cicalese, J.; Vorsa, N.; Iorizzo, M.; Zalapa, J. Multivariate GBLUP Improves Accuracy of Genomic Selection for Yield and Fruit Weight in Biparental Populations of Vaccinium macrocarpon Ait. *Front. Plant Sci.* **2018**, *9*, 1310. [CrossRef]

85. Rutkoski, J.; Poland, J.; Mondal, S.; Autrique, E.; Párez, L.G.; Crossa, J.; Reynolds, M.; Singh, R. Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. *G3 Genes Genomes Genet.* **2016**, *6*, 2799–2808. [CrossRef]

86. Thuillet, A.-C.; Yu, J.; Pressoir, G.; Romero, S.M.; Mitchell, S.E.; Doebley, J.; Kresovich, S.; Goodman, M.M.; Buckler, E.S.; Flint-Garcia, S.A.; et al. Maize association population: A high-resolution platform for quantitative trait locus dissection. *Plant J.* **2005**, *44*, 1054–1064.

87. Rafalski, A. Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Boil.* **2002**, *5*, 94–100. [CrossRef]

88. 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] [PubMed]

89. Breseghello, F.; Sorrells, M.E. Association Analysis as a Strategy for Improvement of Quantitative Traits in Plants. *Crop Sci.* **2006**, *46*, 1323. [CrossRef]

90. Finno, C.J.; Aleman, M.; Higgins, R.J.; Madigan, J.E.; Bannasch, D.L. Risk of false positive genetic associations in complex traits with underlying population structure: A case study. *Vet. J.* **2014**, *202*, 543–549. [CrossRef] [PubMed]

91. Shen, X.; Carlborg, O. Beware of risk for increased false positive rates in genome-wide association studies for phenotypic variability. *Front. Genet.* **2013**, *4*, 93. [CrossRef] [PubMed]

92. Huang, M.; Liu, X.; Zhou, Y.; Summers, R.M.; Zhang, Z. BLINK: A Package for Next Level of Genome Wide Association Studies with Both Individuals and Markers in Millions. *Gigascience* **2019**, *8*. [CrossRef] [PubMed]

93. Liu, X.; Huang, M.; Fan, B.; Buckler, E.S.; Zhang, Z. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLoS Genet.* **2016**, *12*, 1005767. [CrossRef]

94. Lipka, A.E.; Tian, F.; Wang, Q.; Peiffer, J.; Li, M.; Bradbury, P.J.; Gore, M.A.; Buckler, E.; Zhang, Z. GAPIT: Genome association and prediction integrated tool. *Bioinformatics* **2012**, *28*, 2397–2399. [CrossRef] [PubMed]

95. Bernardo, R. Molecular Markers and Selection for Complex Traits in Plants: Learning from the Last 20 Years. *Crop Sci.* **2008**, *48*, 1649. [CrossRef]

96. Bian, Y.; Holland, J.B. Enhancing genomic prediction with genome-wide association studies in multiparental maize populations. *Heredity* **2017**, *118*, 585–593. [CrossRef]

97. Spindel, J.E.; Begum, H.; Akdemir, D.; Collard, B.; Redoña, E.; Jannink, J.-L.; McCouch, S. Genome-wide prediction models that incorporate de novo GWAS are a powerful new tool for tropical rice improvement. *Heredity* **2016**, *116*, 395–408. [CrossRef]

98. Cericola, F.; Jahoor, A.; Orabi, J.; Andersen, J.R.; Janss, L.L.; Jensen, J. Optimizing training population size and genotyping strategy for genomic prediction using association study results and pedigree information. A case of study in advanced wheat breeding lines. *PLoS ONE* **2017**, *12*, e0169606. [CrossRef] [PubMed]

99. Gianola, D.; Van Kaam, J.B.C.H.M. Reproducing Kernel Hilbert Spaces Regression Methods for Genomic Assisted Prediction of Quantitative Traits. *Genetics* **2008**, *178*, 2289–2303. [CrossRef] [PubMed]

100. Gianola, D.; Campos, G.D.L.; Hill, W.G.; Manfredi, E.; Fernando, R. Additive Genetic Variability and the Bayesian Alphabet. *Genetics* **2009**, *183*, 347–363. [CrossRef] [PubMed]

101. Bernardo, R. Genomewide Selection when Major Genes Are Known. *Crop Sci.* **2014**, *54*, 68. [CrossRef]

102. Rutkoski, J.E.; Poland, J.A.; Singh, R.P.; Huerta-Espino, J.; Bhavani, S.; Barbier, H.; Rouse, M.N.; Jannink, J.-L.; Sorrells, M.E. Genomic selection for quantitative adult plant stem rust resistance in wheat. *Plant Genome* **2014**, *7*. [CrossRef]

103. Owens, B.F.; Lipka, A.E.; Magallanes-Lundback, M.; Tiede, T.; Diepenbrock, C.H.; Kandianis, C.B.; Kim, E.; Cepela, J.; Mateos-Hernandez, M.; Buell, C.R.; et al. A Foundation for Provitamin a Biofortification of Maize: Genome-Wide Association and Genomic Prediction Models of Carotenoid Levels. *Genetics* **2014**, *198*, 1699–1716. [CrossRef] [PubMed]

104. Mason, R.E.; Addison, C.K.; Babar, A.; Acuna, A.; Lozada, D.; Subramanian, N.; Arguello, M.N.; Miller, R.G.; Brown-Guedira, G.; Guedira, M.; et al. Diagnostic Markers for Vernalization and Photoperiod Loci Improve Genomic Selection for Grain Yield and Spectral Reflectance in Wheat. *Crop Sci.* **2018**, *58*, 242. [CrossRef]

105. Zhang, Z.; Ober, U.; Erbe, M.; Zhang, H.; Gao, N.; He, J.; Li, J.; Simianer, H. Improving the Accuracy of Whole Genome Prediction for Complex Traits Using the Results of Genome Wide Association Studies. *PLoS ONE* **2014**, *9*, e93017. [CrossRef] [PubMed]

106. Rice, B.; Lipka, A.E. Evaluation of RR-BLUP Genomic Selection Models that Incorporate Peak Genome-Wide Association Study Signals in Maize and Sorghum. *Plant Genome* **2019**, *12*. [CrossRef]

107. He, S.; Schulthess, A.W.; Mirdita, V.; Zhao, Y.; Korzun, V.; Bothe, R.; Ebmeyer, E.; Reif, J.C.; Jiang, Y. Genomic selection in a commercial winter wheat population. *Theor. Appl. Genet.* **2016**, *129*, 641–651. [CrossRef]

108. Michel, S.; Ametz, C.; Gungor, H.; Epure, D.; Grausgruber, H.; Löschenberger, F.; Buerstmayr, H. Genomic selection across multiple breeding cycles in applied bread wheat breeding. *Theor. Appl. Genet.* **2016**, *129*, 1179–1189. [CrossRef] [PubMed]

109. Michel, S.; Ametz, C.; Gungor, H.; Akgöl, B.; Epure, D.; Grausgruber, H.; Löschenberger, F.; Buerstmayr, H. Genomic assisted selection for enhancing line breeding: Merging genomic and phenotypic selection in winter wheat breeding programs with preliminary yield trials. *Theor. Appl. Genet.* **2017**, *130*, 363–376. [CrossRef] [PubMed]

110. Xu, X.; Nicholson, P. Community Ecology of Fungal Pathogens Causing Wheat Head Blight. *Annu. Rev. Phytopathol.* **2009**, *47*, 83–103. [CrossRef] [PubMed]

111. Milus, E.A. Evaluation of Foliar Fungicides for Controlling Fusarium Head Blight of Wheat. *Plant Dis.* **1994**, *78*, 697. [CrossRef]

112. Goswami, R.S.; Kistler, H.C. Heading for disaster: Fusarium graminearum on cereal crops. *Mol. Plant Pathol.* **2004**, *5*, 515–525. [CrossRef] [PubMed]

113. Wegulo, S.N.; Baenziger, P.S.; Nopsa, J.H.; Bockus, W.W.; Hallen-Adams, H. Management of Fusarium head blight of wheat and barley. *Crop Prot.* **2015**, *73*, 100–107. [CrossRef]

114. Sobrova, P.; Adam, V.; Vasatkova, A.; Beklova, M.; Zeman, L.; Kizek, R. Deoxynivalenol and its toxicity. *Interdiscip. Toxicol.* **2010**, *3*, 94–99. [CrossRef]

115. Dahl, B.; Wilson, W.W. Risk premiums due to Fusarium Head Blight (FHB) in wheat and barley. *Agric. Syst.* **2018**, *162*, 145–153. [CrossRef]

116. Steiner, B.; Buerstmayr, M.; Michel, S.; Schweiger, W.; Lemmens, M. Breeding strategies and advances in line selection for Fusarium head blight resistance in wheat. *Trop. Plant Pathol.* **2017**, *42*, 165–174. [CrossRef]

117. Buerstmayr, H.; Ban, T.; Anderson, J.A. QTL mapping and marker-assisted selection for Fusarium head blight resistance in wheat: A review. *Plant Breed.* **2009**, *128*, 1–26. [CrossRef]

118. Waldron, B.L.; Moreno-Sevilla, B.; Anderson, J.A.; Stack, R.W.; Frohberg, R.C. RFLP Mapping of QTL for Fusarium Head Blight Resistance in Wheat. *Crop Sci.* **1999**, *39*, 805. [CrossRef]

119. Anderson, J.A.; Stack, R.W.; Liu, S.; Waldron, B.L.; Fjeld, A.D.; Coyne, C.; Moreno-Sevilla, B.; Fetch, J.M.; Song, Q.J.; Cregan, P.B.; et al. DNA markers for Fusarium head blight resistance QTLs in two wheat populations. *Theor. Appl. Genet.* **2001**, *102*, 1164–1168. [CrossRef]

120. Buerstmayr, H.; Lemmens, M.; Hartl, L.; Doldi, L.; Steiner, B.; Stierschneider, M.; Ruckenbauer, P. Molecular mapping of QTLs for Fusarium head blight resistance in spring wheat. I. Resistance to fungal spread (Type II resistance). *Theor. Appl. Genet.* **2002**, *104*, 84–91. [CrossRef] [PubMed]

121. Buerstmayr, H.; Steiner, B.; Hartl, L.; Griesser, M.; Angerer, N.; Lengauer, D.; Miedaner, T.; Schneider, B.; Lemmens, M. Molecular mapping of QTLs for Fusarium head blight resistance in spring wheat. II. Resistance to fungal penetration and spread. *Theor. Appl. Genet.* **2003**, *107*, 503–508. [CrossRef] [PubMed]

122. Buerstmayr, H.; Stierschneider, M.; Steiner, B.; Lemmens, M.; Griesser, M.; Nevo, E.; Fahima, T. Variation for resistance to head blight caused by Fusarium graminearum in wild emmer (Triticum dicoccoides) originating from Israel. *Euphytica* **2003**, *130*, 17–23. [CrossRef]

123. Xue, S.; Li, G.; Jia, H.; Xu, F.; Lin, F.; Tang, M.; Wang, Y.; An, X.; Xu, H.; Zhang, L.; et al. Fine mapping Fhb4, a major QTL conditioning resistance to Fusarium infection in bread wheat (Triticum aestivum L.). *Theor. Appl. Genet.* **2010**, *121*, 147–156. [CrossRef] [PubMed]

124. Xue, S.; Xu, F.; Tang, M.; Zhou, Y.; Li, G.; An, X.; Lin, F.; Xu, H.; Jia, H.; Zhang, L.; et al. Precise mapping Fhb5, a major QTL conditioning resistance to Fusarium infection in bread wheat (Triticum aestivum L.). *Theor. Appl. Genet.* **2011**, *123*, 1055–1063. [CrossRef]

125. Jiang, G.-L.; Dong, Y.; Shi, J.; Ward, R.W. QTL analysis of resistance to Fusarium head blight in the novel wheat germplasm CJ 9306. II. Resistance to deoxynivalenol accumulation and grain yield loss. *Theor. Appl. Genet.* **2007**, *115*, 1043–1052. [CrossRef]

126. Jiang, G.-L.; Shi, J.; Ward, R.W. QTL analysis of resistance to Fusarium head blight in the novel wheat germplasm CJ 9306. I. Resistance to fungal spread. *Theor. Appl. Genet.* **2007**, *116*, 3–13. [CrossRef]

127. Guo, J.; Zhang, X.; Hou, Y.; Cai, J.; Shen, X.; Zhou, T.; Xu, H.; Ohm, H.W.; Wang, H.; Li, A.; et al. High-density mapping of the major FHB resistance gene Fhb7 derived from Thinopyrum ponticum and its pyramiding with Fhb1 by marker-assisted selection. *Theor. Appl. Genet.* **2015**, *128*, 2301–2316. [CrossRef]

128. Schroeder, H.W.; Christensen, J. Factors affecting resistance of wheat to scab caused by Gibberella zeae. *Phytopathology* **1963**, *53*, 831–838.

129. Mesterházy, Á. Types and components of resistance to Fusarium head blight of wheat. *Plant Breed.* **1995**, *114*, 377–386. [CrossRef]

130. Ban, T. Analysis of Quantitative Trait Loci Associated with Resistance to Fusarium Head Blight Caused by Fusarium graminearum Schwabe and of Resistance Mechanisms in Wheat (Triticum aestivum L.). *Breed. Sci.* **2000**, *50*, 131–137. [CrossRef]

131. Buerstmayr, H.; Lemmens, M. Breeding healthy cereals: Genetic improvement of Fusarium resistance and consequences for mycotoxins. *World Mycotoxin J.* **2015**, *8*, 591–602. [CrossRef]

132. Rutkoski, J.; Benson, J.; Jia, Y.; Brown-Guedira, G.; Jannink, J.-L.; Sorrells, M. Evaluation of Genomic Prediction Methods for Fusarium Head Blight Resistance in Wheat. *Plant Genome* **2012**, *5*, 51. [CrossRef]

133. Hoffstetter, A.; Cabrera, A.; Huang, M.; Sneller, C. Optimizing Training Population Data and Validation of Genomic Selection for Economic Traits in Soft Winter Wheat. *G3 Genes Genomes Genet.* **2016**, *6*, 2919–2928. [CrossRef] [PubMed]

134. Arruda, M.P.; Brown, P.J.; Lipka, A.E.; Krill, A.M.; Thurber, C.; Kolb, F.L. Genomic selection for predicting Fusarium head blight resistance in a wheat breeding program. *Plant Genome* **2015**, *8*, 1–12. [CrossRef]

135. Dong, H.; Wang, R.; Yuan, Y.; Anderson, J.; Pumphrey, M.; Zhang, Z.; Chen, J. Evaluation of the potential for genomic selection to improve spring wheat resistance to Fusarium head blight in the Pacific Northwest. *Front. Plant Sci.* **2018**, *9*, 911. [CrossRef]