

Review

Accelerating Tomato Breeding by Exploiting Genomic Selection Approaches

Elisa Cappetta, Giuseppe Andolfo, Antonio Di Matteo , Amalia Barone , Luigi Frusciante and Maria Raffaella Ercolano * 

Department of Agricultural Sciences, University of Naples Federico II, Via Università 100, 80055 Naples, Italy; elisa.cappetta@unina.it (E.C.); giuseppe.andolfo@unina.it (G.A.); antonio.dimatteo@unina.it (A.D.M.); ambarone@unina.it (A.B.); fruscian@unina.it (L.F.)

* Correspondence: ercolano@unina.it; Tel.: +39-081-253-9431

Received: 15 April 2020; Accepted: 15 September 2020; Published: 18 September 2020



Abstract: Genomic selection (GS) is a predictive approach that was built up to increase the rate of genetic gain *per* unit of time and reduce the generation interval by utilizing genome-wide markers in breeding programs. It has emerged as a valuable method for improving complex traits that are controlled by many genes with small effects. GS enables the prediction of the breeding value of candidate genotypes for selection. In this work, we address important issues related to GS and its implementation in the plant context with special emphasis on tomato breeding. Genomic constraints and critical parameters affecting the accuracy of prediction such as the number of markers, statistical model, phenotyping and complexity of trait, training population size and composition should be carefully evaluated. The comparison of GS approaches for facilitating the selection of tomato superior genotypes during breeding programs is also discussed. GS applied to tomato breeding has already been shown to be feasible. We illustrated how GS can improve the rate of gain in elite line selection, and descendent and backcross schemes. The GS schemes have begun to be delineated and computer science can provide support for future selection strategies. A new promising breeding framework is beginning to emerge for optimizing tomato improvement procedures.

Keywords: tomato; genetic breeding value; training population; genotyping; marker effect; phenotyping; selection schemes

1. The Tomato Genetic Background

Tomato (*Solanum lycopersicum*) is one of the most important vegetable crops worldwide. It possesses unique properties, offering a rich source of minerals (potassium, magnesium, phosphorus) and antioxidant compounds, which prevents cardiovascular and cancer diseases, and strengthens our immune system [1]. Tomato is an autogamous diploid species, with a modest genome size (~900 Mb) and a relatively short life cycle. As a model plant, numerous genetic and molecular tools have been developed for tomato species, including a high-quality draft genome sequence, high-density genetic maps, high-throughput molecular markers, introgression lines, and mutant collections (Tomato Genome Consortium—[2]). Moreover, hundreds of genomes from landraces, cultivars, and wild relatives have been re-sequenced, revealing a relatively low molecular diversity but a high rate of chromosome rearrangements due to traces of wild introgressions [3].

Tomato genetic basis became narrow along the process of domestication, preventing intra-population breeding strategies to provide satisfactory genetic gains [4]. Besides the low genetic variability that limits breeding gains of conventional and modern selection schemes, the tomato is tolerant to inbreeding and this allows the generation and maintenance of inbred lines. Therefore, the recombination of the genetic variability represents an excellent alternative for obtaining superior genotypes [4,5]. Moreover, the retaining of genome segments from wild relatives, used to introgress agronomically relevant traits such as resistance to diseases and quality traits, largely contributes to the genetic variability within the cultivated tomato gene pool [6–8].

In the early 1980s, the development of different molecular marker systems drastically changed the fate of plant breeding. A high saturated tomato reference linkage map based on *L. esculentum* LA925 (E6203) and *L. pennellii* LA716 interspecific population is available and the growing tomato sequencing projects are providing additional information for developing more resolving genetic markers [9]. However, the lack of enough DNA markers that detect polymorphism within the cultivated species remains a major issue in tomato crop [10]. Thus, most recently significant efforts were attempted to exploit intraspecific high-resolution genetic markers such as Single Nucleotide Polymorphisms (SNPs) and Insertion and Deletion (InDels) to detect polymorphism among closely related individuals [11,12].

Molecular markers were mainly integrated into traditional phenotypic selection (PS) by applying marker-assisted selection (MAS) to improve the plant selection process through the inclusion of chromosomal segments containing quantitative trait loci (QTLs) or single genes [13–17]. Several research articles concerning the identification of tomato QTLs and major genes conferring resistance to biotic and environmental stresses have been reviewed in [6,18]. Molecular markers have been also used in tomato to map genes or QTLs for environmental stresses and some flower and fruit-related traits, reviewed in [19]. However, MAS is more suitable for application concerning simple traits with a few major-effect genes [20,21]. Genomic selection (GS), which uses genome-wide markers to predict breeding values, may greatly improve the selection gain in breeding programs for complex traits controlled by several minor genes. In last year, pioneer studies concerning the application of GS to transfer yield-related traits in tomato varieties from wild related species were reported [22,23] as well as to assess the potential of GS to increase soluble solids content and fruit weight in F1 tomato varieties [24] and to develop bacterial spot resistant tomato lines [25]. GS models were widely exploited for predicting phenotypes of progeny and parents, although the efficiency varied depending on the parental cross combinations and the selected traits [26]. Optimized and validated GS protocols are still needed in the tomato. Several GS programs in this species are still in progress, thus the impact of factors affecting the implementation and the accuracy of the model has not yet been evaluated while their optimization for tomato breeding is still required. Among these factors, the genetic structure of species, phenotyping procedures, size of populations, genetic relationship between individuals of population assessed, genotyping platforms, marker quality metrics, and design of GS schemes should be further investigated.

Here, we illustrate the main progress achieved in plant GS and discuss the main challenges of its application in tomato breeding. Tomato GS schemes within and across breeding generations, as well as its potential to select parents based on their assessed GEBV, are also described.

2. Potential of GS in Plant

Genomic selection (GS) provides new opportunities for increasing the efficiency of plant breeding programs for traits with polygenic inheritance [21,27–30]. The potential breeding value of an individual is estimated using large scale genomic-based data such as SNPs. Recent high-throughput genotyping (HTG) systems help to generate several thousand of markers allowing entire genomes to be scanned. The allelic association of marker loci with the phenotypes can be used to predict the phenotypic value of a candidate for selection. A genomic estimated breeding value (GEBV), expressed as a linear function of marker effects, is calculated for each breeding candidate. GS combines genotypic and phenotypic data from a training population (TRN) in a training set (TRS) to obtain the GEBVs of a testing set (TST) which

has been genotyped but not phenotyped. The GS model will be then employed to predict breeding values of not phenotyped individuals in the next selection step (Figure 1). The first report concerning GS simulation in plants was provided by Meuwissen and colleagues [31]. The authors provided a comparison among linear regression, best linear unbiased prediction, and Bayesian prediction methods for estimating the relationship between true breeding values and estimated breeding values in order to develop suitable GS models in plant.

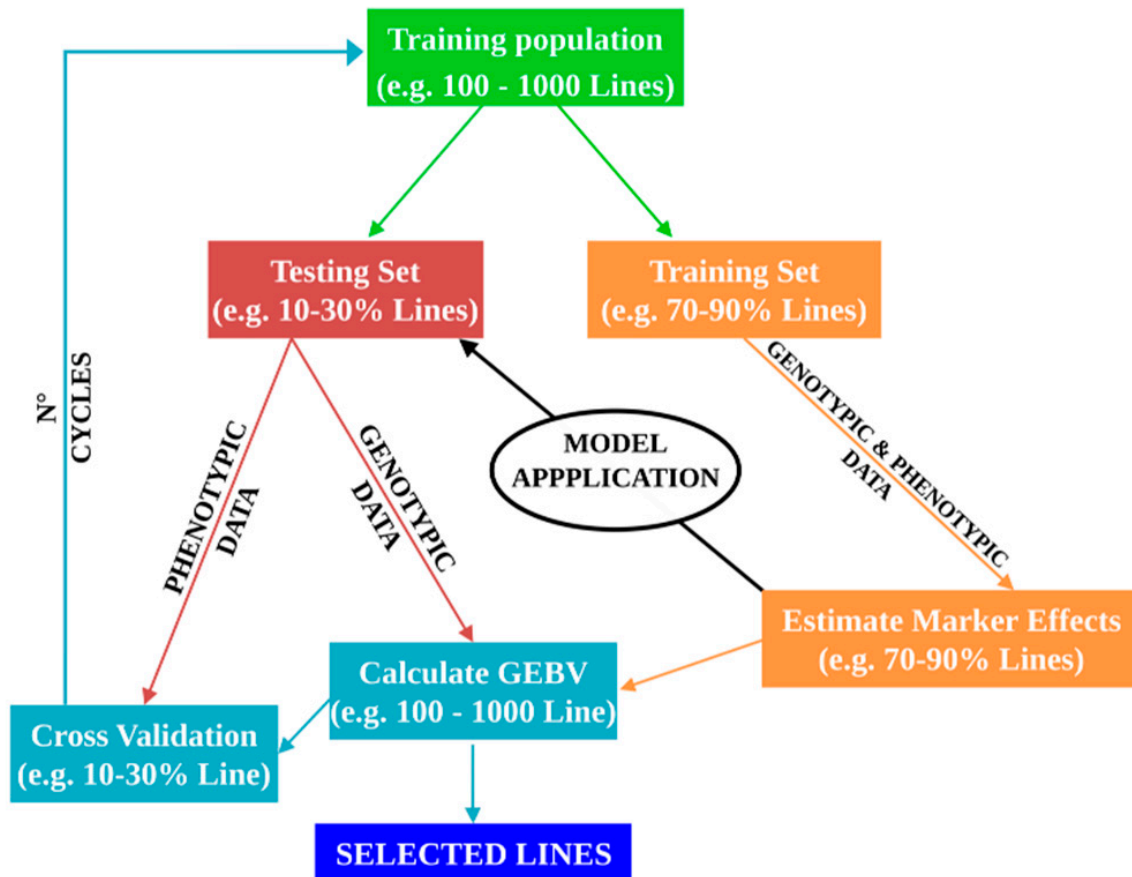


Figure 1. Flowchart of a genomic selection (GS) breeding program. GS overview with cross-validation using a training set (70–90% out of 100–1000 lines) to estimate marker effects in order to obtain a genomic estimated breeding value (GEBV) of lines in the testing set (10–30% out of 100–1000 lines). Finally, phenotypic and genotypic data of the training set are used to set up the prediction model.

Genomic screening of breeding populations can accelerate the genetic gain obtained at each cycle, reducing up two-third the time required for selection [21], especially when the selection is performed for traits with low heritability. Although the effect of each marker is very small, a large amount of genome-wide marker information has the potential to explain all the genetic variance [29]. However, the implementation of GS in real plant breeding schemes can be challenging for plant breeders. The number of candidates that can enter a plant breeding program is typically limited by costs and time of the breeding cycle that ultimately impact on the rate of genetic gain in crops. To date, several GS-based breeding strategies have been conceived in different crops including wheat, maize, rice, barley, soybean, tomato (Table 1) increasing rates of genetic gain [32–35]. The extensive use of GS in plants breeding requires to reduce GS costs, develop cost-effective genotyping, phenotyping platforms, create diverse and updatable TRN, develop highly efficient and multifunctional genomic prediction models, enhance agronomic procedures for shortening breeding cycle time, build up a strong decision support system, and establish open-source breeding programs. In addition, several primary factors that affect GS, including marker density, training-set size, species genetic relationship between TRN

and TST, population structure, phenotyping, and statistical model reliability should be carefully calibrated [36,37].

The prediction accuracy of a GS approach in plant breeding programs is strictly dependent on population linkage disequilibrium (LD), which increases with the number of recombination events [38]. GS is most accurate if the training and prediction populations are closely related and share long-range haplotypes [39,40]. Different studies underlined the impact of increasing the relatedness by including more related crosses in TRN rather than increasing the TRN size by adding unrelated or less-related crosses [26,41–43]. However, continuously using closely related populations to achieve better prediction would narrow down the genetic basis and ultimately reduce the genetic gain that would be achieved in long-term GS [38,44,45]. Therefore, the TRN–TST relationship should be balanced to preserve the genetic gain for both short-term and long-term selection [46]. It is noteworthy to underline that the quality of phenotyping strongly affect the accuracy of GS [47]. To date, phenotyping is labor and time-intensive and is also largely affected by human errors and biases. For these reasons, robotic devices and aerial vehicles are becoming a big opportunity to increase the accuracy of the phenotypic estimations, which in turn can be used in statistical models [48–52]. The development of statistical models capable of accurately predict marker effects has led to the breakthrough of GS increasing the rate of genetic gain per unit of time. However, the ability of different models to predict plant performances need to be evaluated. Heffner et al. [53] reported comparable values among Bayes-A, Bayes-B, and RR-BLUP models for 13 traits in 374 wheat lines. Besides, in 413 rice empirical data, small differences in accuracy were found with different statistical models (least absolute shrinkage and selection operator-LASSO, Bayes-B, and RR-BLUP) [54]. In addition, some studies report that the accuracy of prediction change when different statistical models were applied to different traits. For instance, Perez-Rodriguez et al. [55] applied seven different statistical models in wheat on two traits (heading date and grain yield), revealing that Reproducing Kernel Hilbert Spaces (RKHS) had the best accuracy for heading date, whereas Bayes-A and Bayes-B were the best in evaluating grain yield. Given these differences, it is advisable for researchers to carefully train models before the implementation of GS in crop breeding programs.

Table 1. Genomic selection studies in plant species.

Species	Traits	TRN Size and Type	No Markers	Statistical Model	Accuracy	References
Wheat	GY	374 inbred lines	1158 DArTs	RR-BLIP, Bayes-A, B, C	0.21	[53]
Wheat	GY, HD	306 lines CIMMYT	1717 DArTs	RR-BLUP, Bayes-A, B, LASSO, RKHS, RBFNN, BRNN	0.7 0.5–0.6	[55]
Wheat	GY	254 lines CIMMYT	2056 SNPs	LASSO, Bayes-b, RR-BLUP	0.43–0.51	[54]
Wheat	GY	94 lines CIMMYT	234 DArTs	Bayes-LASSO-RKHS	0.43–0.79	[56]
Wheat	GY	254 lines CIMMYT	34,749 SNPs	GBLUP	0.2–0.4	[57]
Rice	FT	413 varieties	36,901 SNPs	LASSO, Bayes-b, RR-BLUP	~0.5	[54]
Rice	YP, FT, WSY	210 Inbred lines	270,820 SNPs	LASSO	0.16–0.26–0.98	[58]
<i>Arabidopsis</i>	FT	199 inbred lines	215,908 SNPs	RR-BLUP	0.65–0.75	[54]
<i>Arabidopsis</i>	FT, DM	415 RILs	69 SSRs	BLUP	0.90–0.93	[59]
Soybean	YP, PO	540 (RILs)	2647 SNPs	RR-BLUP	0.81, 0.71, 0.26	[60]
Soybean	nematode resistance	363 Genotypes	84,416 SNPs	RR-BLUP	0.41–0.52	[61]
Maize	GY, ASI	255 inbred lines	37,403 SNPs	RR-BLUP	~0.5	[62]
Maize	GY, FF, ME, ASI	300 lines CIMMYT	1148 SNPs	M-BL	0.42–0.79	[27]
Barley	GY, AA	150 DHs	223 RFLPs	BLUP	0.64–0.83	[59]
Barley	PH, CC	140 DHs	107 RFLPs, AFLPs	BLUP	0.66–0.85	[59]
Tomato	SSC, FW	96 F1 varieties	337 SNPs	GBLUP, Bayesian Lasso, Wbsr, BayesC, RKHS, RF	0.56–0.68 0.22–0.27	[24]
Tomato	Metabolic and quality traits	163 Genotypes	5995 SNPs	RR-BLUP	0.05–0.81	[22]

YP = yield; PO = protein oil; FT = flowering time; WSY = white sugar yield; DM = dry matter; GY = grain yield; ASI = anthesis-silking interval; FF = female flowering; MF = male flowering; AA = amylase activity; PH = plant height; CC = chemical components; SSC = soluble solid content; FW = fruit weight; HD = heading date.

3. Lesson from Other Species

The effectiveness of GEBVs for prediction was mainly demonstrated in polyploid wheat [27,52–57] but studies are also available for diploids rice [54,58], barley [59,63], soybean [60,61], maize [27,62] and tomato [22–25]. Lorenzana and Bernardo [59] obtained GEBV accuracies between 0.64 and 0.83 using 150 DHs (doubled-haploid) barley lines and 223 Restriction Fragment Length Polymorphism (RFLP) markers to improve grain yield and amylase activity. Similarly, Xu [58], in rice, estimated the GEBV accuracy at 0.16 and 0.98 for yield and white sugar, respectively, using 210 inbred lines and 270,082 SNPs. Drought-tolerant maize varieties, referred to as the “*AQUAmax*” hybrids, significantly improving yield stability under water-limitation were generated by GS [64].

Wimmer et al. [54] obtained GEBV accuracies between 0.43 and 0.51 for grain yield using three different statistical models in 254 CIMMYT wheat lines, whereas Perez-Rodriguez and colleagues [55] recorded an accuracy of roughly 0.7 for the same trait but in 306 CIMMYT wheat lines.

These studies clearly indicated that the training population size is an important driver in determining the model accuracy. Indeed, Xu [58], using a TRN made of 210 rice varieties, obtained an accuracy of 0.26 for flowering time, whereas Wimmer et al. [54] performed an accuracy of 0.50 using 413 genotypes for the same trait. High accuracy levels were achieved in the par excellence model species *Arabidopsis thaliana*. In fact, an accuracy of ~0.9 was obtained for flowering time using 415 RIL population [59], whereas lower values (~0.7) were reported when a TRN made of 199 inbred lines was used [54]. Another key component for genomic selection success is the marker density. Generally, the higher the number of markers used, the better the accuracy obtained. A critical example was given in wheat by Heffner et al. [53] and Perez-Rodriguez et al. [55]. The former reported accuracy of 0.21 for grain yield using 374 individuals and 1158 Diversity arrays technology (DArT) markers. By contrast, the latter achieved better results for the same trait with a similar number of samples but higher marker density (accuracy of 0.7 obtained with 1717 DArTs). Generally, high marker density can augment prediction accuracy until the maximum level [65–68]. In addition, the marker density required for outcrossing species is higher than that for self-pollinated species [68,69] and the marker numbers required for natural populations with higher LD are normally higher than those for biparental populations [69,70]. However, accuracy and number of markers did not always have a positive correlation. Indeed, Poland and collaborators [57], in wheat, achieved a lower accuracy on grain yield with 34,749 SNPs than Wimmer et al. [54] that analyzed the same CIMMYT wheat lines with only 2056 SNPs (~0.3 vs. ~0.47). Additionally, Xu [58] achieved a lower accuracy for flowering time in rice with 270,820 SNPs than Wimmer et al. [54], which used 36,901 SNPs (0.26 vs. 0.5). Windhaussen et al. [62] in maize obtained only a slightly higher accuracy for grain yield (~0.5) with 37,403 SNPs than Crossa et al. [27] with 1148 SNPs (0.42), although both authors used approximately the same TRN size and composition.

4. Tomato GS Schema Implementation

The establishment of GS experiment optimal parameters in a crop species requires a careful evaluation of key factors [71]. Plant selection response depends on the precision of the phenotyping and genotyping methods used to obtain the GEBVs (including the size of TRN, marker density, marker technology), knowledge of the genome structure, and marker linkage disequilibrium [26].

The success of modern breeding programs based on genomic techniques strictly depends on the precision of measurements related to phenotyped traits [72]. Digital instruments with scalable technologies can improve the precision of phenotyping [73], reduce the requirement of human data annotation, and accelerate the selection. Recent technologies have been used to acquire specific data on tomato traits with the aim of boosting the precision and the throughput of measurements, the size of analyzed plant populations, and, thus, enhancing the accuracy of the predicted phenotypic value and the genetic gain [74,75].

The appropriate TRN size and composition are also critical for gaining high prediction accuracy. A positive correlation between prediction accuracy and TRN size was confirmed in several species [76,77]. However, the optimal TRN size seems to be highly influenced by the relatedness

of TRS and TST [25,78,79]. The highest prediction accuracies were found using TRS with a strong relationship to the TST [28,80,81]. Indeed, when the TRS and TST are unrelated, marker effects could be inconsistent due to the presence of different alleles, allele frequencies, and linkage phases. Developing ad hoc tomato TRN is crucial and update the TRN at each cycle could improve the prediction accuracy since the segregating population could accumulate genetic diversity and gene frequencies may change in each selection cycle [25].

To capture as many informative loci as possible an appropriate abundance of markers is required [82]. In this regard, genotyping-by-sequencing (GBS) can be used to efficiently generate high-density marker panels. Alternately, the cDNA-based GBS technique (RAR-seq restriction site-associated RNA sequencing) may detect conserved SNPs associated with a candidate mutation directly at the expression level [83]. Recently, a customizable method for tomato targeted genotyping, named single primer enrichment technology (SPET) was developed for improving the panel design and increasing the multiplexing levels of tomato genotyping [84]. Previous GS data can help to design an optimized suite of markers for the next steps. Liabeuf et al. [25] reduced the initial “SolCAP array” of 7700 SNPs [85] to screen populations with limited recombination. Moreover, the prediction accuracy may be also affected by the minor allele frequency threshold (MAF) [82]. Establishing methods for efficiently transferring validated genome signatures within tomato breeding selection procedures is also relevant. Human selection leads to changes in genomic regions that affect traits of agronomic interest, [86]. Detecting selection signatures is important for a better understanding of population history and genetic mechanisms affecting phenotypic cultivar differentiation [87]. Estimating allele or haplotype frequency differences between populations or generations within a population may improve traits selection. Linkage drag caused by recombination suppression can be reduced by estimating the effects of relevant markers improving prediction performance. Indeed, large gene introgression fragments in tomato cultivars from *Solanum* wild species caused drastic chromosome landscape changes. The *Solanum peruvianum* introgression carrying the tomato mosaic virus (ToMV) resistance gene *Tm2* can cover up to 79% of chromosome 9 in modern varieties [3]. To identify homologous recombination points, it was estimated that an average of 8× coverage should be used in tomato [88].

In the framework of GS, several statistical methods have been tested to estimate the marker effects in tomato [25]. The choice of the most appropriate method should be finalized to the specific context, considering the model complexity (genetic architecture, population size, and heritability) and the computation requirements [89,90]. Ridge regression best linear unbiased prediction (RR-BLUP) and genomic best linear unbiased prediction (G)BLUP [91], which assume a normal distribution of SNP effects, are suggested when assessing a trait that is affected by many small-effect genes using close TRN relatives. On the other hand, when traits are controlled by major-effect QTLs or when considering prediction of unrelated individuals, higher prediction accuracy can be obtained by Bayesian methods, considering a prior distribution of effects [92]. However empirical studies suggest that there are no major differences between regression-based and Bayesian methods in tomato [25].

5. Applying GS in Tomato Crop Improvement

Several constraints can affect the genetic gain of a GS program in the tomato. The implementation of GS requires the optimization of field trial management and agricultural practices, seed production, phenotyping, sample collection, and sequencing [93]. Moreover, as discussed above, parameters such as inbreeding level of populations, the number of individuals to be assessed, and marker metrics should be carefully evaluated to effectively run a GS-assisted breeding scheme. It can be estimated that, for tomato breeding programs, the genotyping work to complete GEBV predictions requires approximately three months. The selection decision will be achieved based on the higher GEBVs for each tested trait on the overall average of traits or as ‘indices’ of GEBV from several traits following selection priorities.

Once these issues have been addressed, the GEBVs can be calculated both to perform parental line selection and to evaluate the overall performance of the progenies in a descendent selection or backcross schemes.

The selection of elite parents to maximize the genetic variability exploitation is the first step in tomato F1 hybrid variety development. Elite germplasm represents a core collection of cross-compatible genotypes enriched for some favorable alleles [94]. Traditional breeding takes too much time in selecting elite lines. The main advantage of GS over traditional selection is that it can facilitate expeditious selection of superior variety/cultivar in less time by reducing breeding cycles [95].

In a GS-assisted breeding scheme for tomato F1 hybrid development, the decision to select parental lines is based on their breeding value (i.e., the mean performance of the progeny of a given parent) that consequently requires to be estimated accurately. Consistently, Yamamoto and collaborators [24] used a set of 96 big-fruited F1 tomato varieties to develop GS models, and the segregating populations obtained from crosses were used to validate the models. Consequently, the GS models were used to successfully predict parental combinations generating superior hybrids using progeny genotypic and phenotypic data for soluble solids content and total fruit weight. However, the efficiency of predictions varied depending on traits and parental combinations. While the need for fixing favorable alleles in the gene pool leads to increase inbreeding, the GS selection gain is dramatically reduced in small populations with narrow genetic variability. The managing of elite genetic diversity to increase the frequency of favorable alleles over time can highly benefit from GS approaches [94]. The prediction accuracy of parent cross ability could improve with the assessment of a higher number of selfing progenies. Thanks to the advances made in tomato genome knowledge and genotyping technologies, breeders can easily identify valuable alleles in elite germplasm [11,96] and create new lines combining these valuable alleles using a set of validated markers.

Generally, breeders take advantage of useful genetic variability by recycling the tomato best-performing varieties that have been successful for a given area by Single Seed Descendent (SSD) scheme where each generation derived from the former, taking only one seed from each parent plant. Nearly all steps can be conducted in the greenhouse, making this a method of choice for accelerating breeding in areas that do not benefit from a long enough growing season [97]. In the classical SSD scheme, the choice of tomato parental lines is very critical to ensure a higher additive breeding value since self-fertilization increases inbreeding level by 1/2 at each cycle. In the SSD scheme, no selection is conducted until the last generation (generally F6–F7), so the phenotyping of a larger number of lines could be challenging. The integration of the GS approach in the SSD could result in reducing the number of selfing generations thus shortening the overall selection process and decreasing the phenotyping effort (Figure 2). Because the prediction accuracy is generally higher when LD is high, an increase of the breeding gains is expected when applying GS in the earliest heterozygous segregating generations (i.e., F2–F4). Therefore, these generations could be successfully used for developing the GS model, and subsequently, GS prediction could assist selection in the following generations. Genomic data can accurately track the best performing plants along the generations, and the approach can successfully lead to the selection of individuals with the highest GEBV.

Backcrossing is another quite popular tomato breeding scheme employed to introgress a valuable trait from a donor parent into the genomic background of a recurrent parent. Backcrossing schemes with exotic or elite materials are widely used to introduce favorable traits. However, the constant introduction of novel alleles and the linkage drag, the crossing with old varieties or exotic material with low breeding value as well as the extended breeding cycles deriving from complex crossing scheme, can reduce the genetic gain per year. The response to genetic selection achieved through the selection of lines with high breeding value in a segregating population can be certainly improved by GS (Figure 2). A variant of the classical backcross scheme, where lines of each generation are selected based on recurrent parent breeding value, allowed researchers to obtain high rates of genetic gain [98,99]. By combining GS with single-marker assays, genes with major effects can be also selected within each offspring following the cross with the recurrent line. In this way, the GS approach is

expected to additively increase the genetic gain at each generation. Candidate genotypes for selection, carrying specific alleles (i.e., resistance traits), can be identified using genotyping platforms that include gene-specific diagnostic markers or integrate single-locus data obtained with different technologies. In addition, among markers used in the GS model implementation, a subset of them identifying undesirable segments of a wild donor can be selected. In fact, large wild genome segments (between the 30 and 70% of the whole chromosome) were found to be incorporated due to resistance gene introgressions on a specific chromosome in cultivated tomatoes [3]. As an extension of this approach, genome-wide selection with high-throughput markers in BC1 could be even more efficient and the recovering of the recurrent parent genome could be increased from generation BC1 to BC3 without affecting favorable trait introgression.

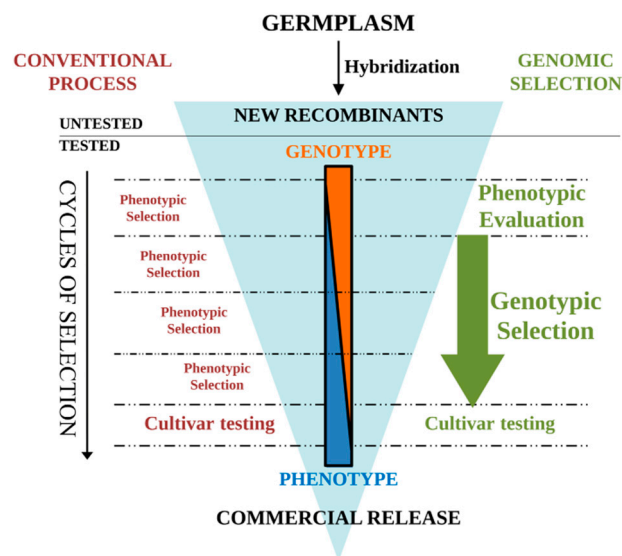


Figure 2. Comparison of genomic selection (GS) and conventional selection in tomato breeding programs. Screening of recombinant lines through GS approaches optimizes the genetic gain obtained in each selection cycle. Breeding cycles (horizontal dashed lines) are shortened by removing phenotypic evaluation of lines before training population (TRN) evaluation for the next cycle.

6. Conclusions

The evaluation of complex traits such as disease resistance genes, QTLs for quality traits and abiotic environmental stresses (such as salinity, drought, and heat) with high efficiency in a segregating population can be a difficult task for tomato breeders. Innovative breeding strategies such as marker-assisted selection (MAS), high-throughput phenotyping, high-throughput genotyping, reverse breeding, and genomic selection, are now increasingly being used to complement the conventional approaches for the effective improvement of the tomato. In particular, the implementation of GS in breeding programs can accelerate genetic achievable gain if selection schemes will be tailored to genomic-guided procedures. This technique offers the possibility to double improve genetic gain. The acquisition of theoretical knowledge about tomato genome structure, evolution, and recombination can help to improve the practical application. The connection between genomic and phenotypic variations gives us the unique opportunity to predict bases on the genome and early in the life of individuals. We may not understand the underlying mechanism, but we can predict the results. Major GS implementation challenges were highlighted here, including model development, genotyping quality, optimal GS incorporation stage and indications for overcoming these issues were also discussed. While the methodological procedures begin to be delineated, the optimal way to incorporate GS in a breeding scheme remains to be empirically defined. Important features for the success of GS under different breeding scenarios should be assessed. Advancements in genotyping efficiency and phenotyping technologies will facilitate the adoption of GS in tomato breeding. A future update of

existing selection schemes may be achieved using computer simulations for investigating different strategies to face the selection process gaps.

Author Contributions: E.C. was centrally involved in writing the manuscript and in drafting figures. G.A. revised the manuscript and produced the figures. A.D.M. revised the text. A.B. provided important suggestions for improving the manuscript. L.F. critically revised the manuscript. M.R.E. conceived the study, coordinated work and contributed to manuscript writing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was carried out within HARNESSTOM and TomGEM Projects that have received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 101000716 and No. 679796 respectively.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Frusciante, L.; Barone, A.; Carputo, D.; Ercolano, M.R.; Della Rocca, F.; Esposito, S. Evaluation and use of plant biodiversity for food and pharmaceuticals. *Fitoterapia* **2000**, *71*, 66–72. [[CrossRef](#)]
2. Sato, S.; Tabata, S.; Hirakawa, H.; Asamizu, E.; Shirasawa, K.; Isobe, S.; Kaneko, T.; Nakamura, Y.; Shibata, D.; Aoki, K. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **2012**, *485*, 635–641.
3. Schouten, H.J.; Tikunov, Y.; Verkerke, W.; Finkers, R.; Bovy, A.; Bai, Y.; Visser, R.G.F. Breeding has Increased the Diversity of Cultivated Tomato in The Netherlands. *Fron. Plant. Sci.* **2019**, *10*, 1606. [[CrossRef](#)] [[PubMed](#)]
4. Souza, L.M.; Paterniani, M.; Melo, P.C.T.; Melo, A.M.T. Diallel cross among fresh market tomato inbreeding lines. *Hortic. Bras.* **2012**, *30*, 246–251. [[CrossRef](#)]
5. Cappetta, E.; Andolfo, G.; Di Matteo, A.; Ercolano, M.R. Empowering crop resilience to environmental multiple stress through the modulation of key response components. *J. Plant. Physiol.* **2020**, *246*, 153134. [[CrossRef](#)] [[PubMed](#)]
6. Ercolano, M.R.; Sanseverino, W.; Carli, P.; Ferriello, F.; Frusciante, L. Genetic and genomic approaches for R-gene mediated disease resistance in tomato: Retrospects and prospects. *Plant. Cell Rep.* **2012**, *31*, 973–985. [[CrossRef](#)]
7. Sacco, A.; Di Matteo, A.; Lombardi, N.; Trotta, N.; Punzo, B.; Mari, A.; Barone, A. Quantitative trait loci pyramiding for fruit quality traits in tomato. *Mol. Breed.* **2013**, *31*, 217–222. [[CrossRef](#)]
8. D’Esposito, D.; Cappetta, E.; Andolfo, G.; Ferriello, F.; Borgonuovo, C.; Caruso, G.; De Natale, A.; Frusciante, L.; Ercolano, M.R. Deciphering the biological processes underlying tomato biomass production and composition. *Plant. Physiol. Bioch.* **2019**, *143*, 50–60. [[CrossRef](#)]
9. Available online: <https://solgenomics.net/cview/index.pl> (accessed on 15 April 2020).
10. Hamilton, J.P.; Sim, S.; Stoffel, K.; Van Deynze, A.; Buell, C.R.; Francis, D.M. Single nucleotide polymorphism discovery in cultivated tomato via sequencing by synthesis. *Plant. Genome* **2012**, *5*, 17–29. [[CrossRef](#)]
11. Sacco, A.; Ruggieri, V.; Parisi, M.; Festa, G.; Rigano, M.M.; Picarella, M.E.; Mazzucato, A.; Barone, A. Exploring a tomato landraces collection for fruit-related traits by the aid of a high-throughput genomic platform. *PLoS ONE* **2015**, *10*, e0137139. [[CrossRef](#)]
12. Esposito, S.; Cardi, T.; Campanelli, G.; Sestili, S.; Diez, M.J.; Soler, S.; Prohens, J.; Tripodi, P. ddRAD sequencing-based genotyping for population structure analysis in cultivated tomato provides new insights into the genomic diversity of Mediterranean ‘da serbo’ type long shelf-life germplasm. *Hortic. Res.* **2020**, *7*, 134. [[CrossRef](#)] [[PubMed](#)]
13. Collard, B.C.Y.; Mackill, D.J. Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2008**, *363*, 557–572. [[CrossRef](#)] [[PubMed](#)]
14. Andolfo, G.; Sanseverino, W.; Aversano, R.; Frusciante, L.; Ercolano, M.R. Genome-wide identification and analysis of candidate genes for disease resistance in tomato. *Mol. Breed.* **2014**, *33*, 227–233. [[CrossRef](#)]
15. Andolfo, G.; Jupe, F.; Witek, K.; Etherington, G.J.; Ercolano, M.R.; Jones, J.D.G. Defining the full tomato NB-LRR resistance gene repertoire using genomic and cDNA RenSeq. *BMC Plant. Biol.* **2014**, *14*, 120. [[CrossRef](#)] [[PubMed](#)]
16. Di Donato, A.; Andolfo, G.; Ferrarini, A.; Delledonne, M.; Ercolano, M.R. Investigation of orthologous pathogen recognition gene-rich regions in solanaceous species. *Genome* **2017**, *60*, 850–859. [[CrossRef](#)]

17. Capuozzo, C.; Formisano, G.; Iovieno, P.; Andolfo, G.; Tomassoli, L.; Barbella, M.M.; Pico, B.; Paris, H.S.; Ercolano, M.R. Inheritance analysis and identification of SNP markers associated with ZYMV resistance in *Cucurbita pepo*. *Mol. Breed.* **2017**, *37*, 1–12. [[CrossRef](#)]
18. Kissoudis, C.; Chowdhury, R.; van Heusden, S.; van de Wiel, C.; Finkers, R.; Visser, R.F.; Bai, Y.; van der Linden, G. Combined biotic and abiotic stress resistance in tomato. *Euphytica* **2015**, *202*, 317–332. [[CrossRef](#)]
19. Osei, M.K.; Prempeh, R.; Adjebeng, J.; Opoku, J.; Danquah, A.; Danquah, E.; Blay, E.; Adu-Dapaah, H. Marker-Assisted Selection (MAS): A Fast-Track Tool in Tomato Breeding. In *Recent Advances in Tomato Breeding and Production*; IntechOpen: London, UK, 2018; pp. 93–113. [[CrossRef](#)]
20. Dekkers, J.C.M.; Hospital, F. The use of molecular genetics in the improvement of agricultural populations. *Nat. Rev. Genet.* **2002**, *3*, 22–32. [[CrossRef](#)]
21. Heffner, E.L.; Sorrells, M.E.; Jannink, J. Genomic Selection for Crop Improvement. *Crop. Sci.* **2009**, *49*, 1–12. [[CrossRef](#)]
22. Duangjit, J.; Causse, M.; Sauvage, C. Efficiency of genomic selection for tomato fruit quality. *Mol. Breed.* **2016**, *36*, 29. [[CrossRef](#)]
23. Yamamoto, E.; Matsunaga, A.; Onogi, A.; Kajiya-Kanegae, H.; Minamikawa, M.; Suzuki, A.; Shirasawa, K.; Hirakawa, H.; Nunome, T.; Yamaguchi, H.; et al. A simulation-based breeding design that uses whole-genome prediction in tomato. *Sci. Rep.* **2016**, *6*, 19454. [[CrossRef](#)] [[PubMed](#)]
24. Yamamoto, E.; Matsunaga, H.; Onogi, A.; Ohyama, A.; Miyatake, K.; Yamaguchi, H.; Nunome, T.; Iwata, H.; Fukuoka, H. Efficiency of genomic selection for breeding population design and phenotype prediction in tomato. *Heredity* **2017**, *118*, 202–209. [[CrossRef](#)] [[PubMed](#)]
25. Liabeuf, D.; Sim, S.C.; Francis, D.M. Comparison of marker-based genomic estimated breeding values and phenotypic evaluation for selection of bacterial spot resistance in tomato. *Phytopathology* **2018**, *108*, 392–401. [[CrossRef](#)]
26. Robertsen, C.D.; Hjørtshøj, R.L.; Janss, L.L. Genomic Selection in Cereal Breeding. *Agronomy* **2019**, *9*, 95. [[CrossRef](#)]
27. Crossa, J.; De Los Campos, G.; Pérez, P.; Gianola, D.; Burguño, J.; Araus, J.L.; Makumbi, D.; Singh, R.P.; Dreisigacker, S.; Yan, J.; et al. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* **2010**, *186*, 713–724. [[CrossRef](#)] [[PubMed](#)]
28. Lorenz, A.J.; Chao, S.; Asoro, F.G.; Heffner, E.L.; Hayashi, T.; Iwata, H.; Hayashi, T.; Iwata, H.; Smith, K.P.; Sorrells, M.E.; et al. Genomic Selection in Plant Breeding. Knowledge and Prospects. *Adv. Agron.* **2011**, *110*, 77–123.
29. Voss-Fels, K.P.; Cooper, M.; Hayes, B.J. Accelerating crop genetic gains with genomic selection. *Theor. Appl. Genet.* **2018**, *132*, 669–686. [[CrossRef](#)]
30. Xu, Y.; Liu, X.; Fu, J.; Wang, H.; Wang, J.; Huang, C.; Prasanna, B.M.; Olsen, M.S.; Wang, G.; Zhang, A. Enhancing genetic gain through genomic selection: From livestock to plants. *Plant. Commun.* **2020**, *1*, 2641–2666. [[CrossRef](#)]
31. 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.
32. Gaynor, R.C.; Gorjanc, G.; Bentley, A.R.; Ober, E.S.; Howell, P.; Jackson, R.; Mackay, I.J.; Hickey, J.M. A two-part strategy for using genomic selection to develop inbred lines. *Crop. Sci.* **2017**, *57*, 2372. [[CrossRef](#)]
33. Gorjanc, G.; Gaynor, R.C.; Hickey, J.M. Optimal cross selection for long-term genetic gain in two-part programs with rapid recurrent genomic selection. *Theor. Appl. Genet.* **2018**, *131*, 1953–1966. [[CrossRef](#)]
34. Lin, Z.; Cogan, N.O.I.; Pembleton, L.W.; Spangenberg, G.C.; Forster, J.W.; Hayes, B.J.; Daetwyler, H.D. Genetic gain and inbreeding from genomic selection in a simulated commercial breeding program for perennial ryegrass. *Plant. Genome* **2016**, *9*. [[CrossRef](#)]
35. Voss-Fels, K.P.; Herzog, E.; Dreisigacker, S.; Sukurmaran, S.; Watson, A.; Frisch, M.; Hayes, B.J.; Hickey, L.T. Speed GS to accelerate genetic gain in spring wheat. In *Applications of Genetic and Genomic Research in Cereals*, 1st ed.; Miedaner, T., Korzun, V., Eds.; Woodhead Publishing: Cambridge, MA, USA, 2018.
36. Heslot, N.; Jannink, J.L.; Sorrells, M.E. Perspectives for genomic selection applications and research in plants. *Crop. Sci.* **2015**, *55*, 1–12. [[CrossRef](#)]
37. Heslot, N.; Yang, H.-P.; Sorrells, M.E.; Jannink, J.-L. Genomic selection in plant breeding. A comparison of models. *Crop. Sci.* **2012**, *52*, 146. [[CrossRef](#)]

38. 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](#)]
39. 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](#)]
40. Cooper, M.; Messina, C.D.; Podlich, D.; Totir, L.R.; Baumgarten, A.; Hausmann, N.J.; Wright, D.; Graham, G. Predicting the future of plant breeding. Complementing empirical evaluation with genetic prediction. *Crop. Pasture Sci.* **2014**, *65*, 311. [[CrossRef](#)]
41. Riedelsheimer, C.; Endelman, J.B.; Stange, M.; Sorrells, M.E.; Jannink, J.L.; Melchinger, A.E. Genomic predictability of interconnected biparental maize populations. *Genetics* **2013**, *194*, 493–503. [[CrossRef](#)] [[PubMed](#)]
42. Jacobson, A.; Lian, L.; Zhong, S.; Bernardo, R. General combining ability model for genomewide selection in a biparental cross. *Crop. Sci.* **2014**, *54*, 895–905. [[CrossRef](#)]
43. Brandariz, S.P.; Bernardo, R. Small ad hoc versus large general training populations for genomewide selection in maize biparental crosses. *Theor. Appl. Genet.* **2019**, *132*, 347–353. [[CrossRef](#)]
44. Hickey, L.T.; Hafeez, A.N.; Robinson, H.; Jackson, S.A.; Leal-Bertioli, S.C.M.; Tester, M.; Gao, C.; Godwin, I.D.; Hayes, B.J.; Wulff, B.B.H. Breeding crops to feed 10 billion. *Nat. Biotechnol.* **2019**, *37*, 744–754. [[CrossRef](#)] [[PubMed](#)]
45. Moeiniazade, S.; Hu, G.; Wang, L.; Schnable, P.S. Optimizing selection and mating in genomic selection with a look-ahead approach: An operations research framework. *G3 Genes Genom. Genet.* **2019**, *9*, 2123–2133. [[CrossRef](#)] [[PubMed](#)]
46. Li, Y.; Kadarmideen, H.N.; Dekkers, J.C. Selection on multiple QTL with control of gene diversity and inbreeding for long-term benefit. *J. Anim. Breed. Genet.* **2008**, *125*, 320–329. [[CrossRef](#)]
47. Goddard, M.E.; Hayes, B.J. Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat. Rev. Genet.* **2009**, *10*, 381–391. [[CrossRef](#)] [[PubMed](#)]
48. Guijarro-Real, C.; Navarro, A.; Esposito, S.; Festa, G.; Macellaro, R.; Di Cesare, C.; Fita, A.; Rodríguez-Burruezo, A.; Cardi, T.; Prohens, J.; et al. Large scale phenotyping and molecular analysis in a germplasm collection of rocket salad (*Eruca vesicaria*) reveal a differentiation of the gene pool by geographical origin. *Euphytica* **2020**, *216*, 53. [[CrossRef](#)]
49. Wang, X.; Xu, Y.; Hu, Z.; Xu, C. Genomic selection methods for crop improvement: Current status and prospects. *Crop. J.* **2018**, *6*, 330–340. [[CrossRef](#)]
50. Wendel, A.; Underwood, J.; Walsh, K. Maturity estimation of mangoes using hyperspectral imaging from a ground based mobile platform. *Comput. Electron. Agric.* **2018**, *155*, 298–313. [[CrossRef](#)]
51. Westling, F.; Underwood, J.; Orn, S. Light interception modelling using unstructured LiDAR data in avocado orchards. *Comput. Electron. Agric.* **2018**, *153*, 177–187. [[CrossRef](#)]
52. Gil-Docampo, M.L.; Arza-Garcia, M.; Ortiz-Sanz, J.; Martinez-Rodriguez, S.; Marcos-Robles, J.L.; Sanchez-Sastre, L.F. Above-ground biomass estimation of arable crops using UAV-based SfM photogrammetry. *Geocarto Int.* **2018**, *35*, 687–699. [[CrossRef](#)]
53. Heffner, E.L.; Jannink, J.L.; Sorrells, M.E. Genomic selection accuracy using multifamily prediction models in a wheat breeding program. *Plant. Genome* **2011**, *4*, 65–75. [[CrossRef](#)]
54. Wimmer, V.; Lehermeier, C.; Albrecht, T.; Auinger, H.J.; Wang, Y.; Schon, C.C. Genome-wide prediction of traits with different genetic architecture through efficient variable selection. *Genetics* **2013**, *195*, 573–587. [[CrossRef](#)] [[PubMed](#)]
55. Perez-Rodriguez, P.; Gianola, D.; González-Camacho, J.M.; Crossa, J.; Manes, Y.; Dreisigacker, S. Comparison between linear and non-parametric regression models for genome-enabled prediction in wheat. *G3 Genes Genom. Genet.* **2012**, *2*, 1595–1605. [[CrossRef](#)] [[PubMed](#)]
56. Crossa, J.; Perez, P.; Campos, G.D.L.; Mahuku, G.; Dreisigacker, S.; Magorokosho, C. Genomic selection and prediction in plant breeding. *J. Crop. Improv.* **2011**, *25*, 239–261. [[CrossRef](#)]
57. Poland, J.; Endelman, J.; Dawson, J.; Rutkoski, J.; Wu, S.; Manes, 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*, 103–113. [[CrossRef](#)]
58. Xu, S.Z. Genetic mapping and genomic selection using recombination breakpoint data. *Genetics* **2013**, *195*, 1103–1115. [[CrossRef](#)]
59. Lorenzana, R.E.; Bernardo, R. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theor. Appl. Genet.* **2009**, *120*, 151–161. [[CrossRef](#)]

60. Stewart-Brown, B.B.; Song, Q.; Vaughn, J.N.; Li, Z. Genomic selection for yield and seed composition traits within an applied soybean breeding program. *G3 Genes Genom. Genet.* **2019**, *9*, 2253–2265. [[CrossRef](#)]
61. Wen, L.; Chang, H.X.; Brown, P.J.; Domier, L.L.; Hartman, G.L. Genome-wide association and genomic prediction identifies soybean cyst nematode resistance in common bean including a syntenic region to soybean Rhg1 locus. *Hortic. Res.* **2019**, *6*, 9. [[CrossRef](#)]
62. 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.; et al. Effectiveness of genomic prediction of maize hybrid performance in different breeding populations and environments. *G3 Genes Genom. Genet.* **2012**, *2*, 1427–1436. [[CrossRef](#)]
63. Lorenz, A.J.; Smith, K.P.; Jannink, J.L. Potential and optimization of genomic selection for fusarium head blight resistance in six-row barley. *Crop. Sci.* **2012**, *52*, 1609–1621. [[CrossRef](#)]
64. Gaffney, J.; Schussler, J.; Löffler, C.; Cai, W.; Paszkiewicz, S.; Messina, C.; Groeteke, J.; Keaschall, J.; Cooper, M. Industry-scale evaluation of maize hybrids selected for increased yield in drought-stress conditions of the US corn belt. *Crop. Sci.* **2015**, *55*, 1608–1618. [[CrossRef](#)]
65. Cao, S.; Loladze, A.; Yuan, Y.; Wu, Y.; Zhang, A.; Chen, J.; Huestis, G.; Cao, J.; Chaikam, V.; Olsen, M.; et al. Genome-wide analysis of tar spot complex resistance in maize using genotyping-by-sequencing SNPs and whole-genome prediction. *Plant. Genome* **2017**, *10*. [[CrossRef](#)] [[PubMed](#)]
66. Lee, S.H.; Clark, S.; van der Werf, J.H.J. Estimation of genomic prediction accuracy from reference populations with varying degrees of relationship. *PLoS ONE* **2017**, *12*, e0189775. [[CrossRef](#)] [[PubMed](#)]
67. Xu, Y.; Wang, X.; Ding, X.; Zheng, X.; Yang, Z.; Xu, C.; Hu, Z. Genomic selection of agronomic traits in hybrid rice using an NCII population. *Rice* **2018**, *11*, 32. [[CrossRef](#)]
68. Juliana, P.; Poland, J.; Huerta-Espino, J.; Shrestha, S.; Crossa, J.; Crespo-Herrera, L.; Toledo, F.H.; Govindan, V.; Mondal, S.; Kumar, U.; et al. Improving grain yield, stress resilience and quality of bread wheat using large-scale genomics. *Nat. Genet.* **2019**, *51*, 1530–1539. [[CrossRef](#)]
69. Liu, X.; Wang, H.; Wang, H.; Guo, Z.; Xu, X.; Liu, J.; Wang, S.; Li, W.-X.; Zou, C.; Prasanna, B.M.; et al. Factors affecting genomic selection revealed by empirical evidence in maize. *Crop. J.* **2018**, *6*, 341–352. [[CrossRef](#)]
70. Hao, Y.; Wang, H.; Yang, X.; Zhang, H.; He, C.; Li, D.; Li, H.; Wang, G.; Wang, J.; Fu, J. Genomic prediction using existing historical data contributing to selection in biparental populations: A study of kernel oil in maize. *Plant. Genome* **2019**, *12*. [[CrossRef](#)]
71. Contaldi, F.; Cappetta, E.; Esposito, S. Practical workflow from High Throughput Genotyping to Genomic Estimated Breeding Values (GEBVs). In *Crop Breeding Methods and Protocols*; Springer: Berlin/Heidelberg, Germany, 2021; in press.
72. Esposito, S.; Carputo, D.; Cardi, T.; Tripodi, P. Applications and Trends of Machine Learning in Genomics and Phenomics for Next-Generation Breeding. *Plants* **2020**, *9*, 34. [[CrossRef](#)]
73. Araus, J.L.; Cairns, J.E. Field high-throughput phenotyping: The new crop breeding frontier. *Trends Plant. Sci.* **2014**, *19*, 52–61. [[CrossRef](#)]
74. Panthee, D.R.; Labate, J.A.; McGrath, M.T.; Breksa, A.P.; Robertson, L.D. Genotype and environmental interaction for fruit quality traits in vintage tomato varieties. *Euphytica* **2013**, *193*, 169–182. [[CrossRef](#)]
75. Daniel, I.O.; Atinsola, K.O.; Ajala, M.O.; Popoola, A.R. Phenotyping a Tomato Breeding Population by Manual Field Evaluation and Digital Imaging Analysis. *Int. J. Plant. Breed. Genet.* **2017**, *11*, 19–24. [[CrossRef](#)]
76. Meuwissen, T.; Hayes, B.; Goddard, M. Accelerating improvement of livestock with genomic selection. *Annu. Rev. Anim. Biosci.* **2013**, *1*, 21–237. [[CrossRef](#)] [[PubMed](#)]
77. 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*, 1–15. [[CrossRef](#)] [[PubMed](#)]
78. Schopp, P.; Müller, D.; Technow, F.; Melchinger, A.E. Accuracy of genomic prediction in synthetic populations depending on the number of parents, relatedness, and ancestral linkage disequilibrium. *Genetics* **2017**, *205*, 441–454. [[CrossRef](#)]
79. Edwards, S.M.; Buntjer, J.B.; Jackson, R.; Bentley, A.R.; Lage, J.; Byrne, E.; Burt, C.; Jack, P.; Berry, S.; Flatman, E.; et al. The effects of training population design on genomic prediction accuracy in wheat. *Theor. Appl. Genet.* **2019**, *132*, 1943–1952. [[CrossRef](#)]
80. Nielsen, N.H.; Jahoor, A.; Jensen, D.; Orabi, J.; Cericola, F.; Edriss, V.; Jensen, J. Genomic prediction of seed quality traits using advanced barley breeding lines. *PLoS ONE* **2016**, *11*, e0164494. [[CrossRef](#)]

81. 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](#)]
82. Zhang, H.; Yin, L.; Wang, M.; Yuan, X.; Liu, X. Factors affecting the accuracy of genomic selection for agricultural economic traits in maize, cattle, and pig populations. *Front. Genet.* **2019**, *10*, 189. [[CrossRef](#)]
83. Alabady, M.S.; Rogers, W.L.; Malmberg, R.L. Development of transcriptomic markers for population analysis using restriction site associated RNA sequencing (RARseq). *PLoS ONE* **2015**, *10*, e0134855. [[CrossRef](#)]
84. Barchi, L.; Acquadro, A.; Alonso, D.; Aprea, G.; Bassolino, L.; Demurtas, O.; Ferrante, P.; Gramazio, P.; Mini, P.; Portis, E.; et al. Single Primer Enrichment Technology (SPET) for High-Throughput Genotyping in Tomato and Eggplant Germplasm. *Front. Plant. Sci.* **2019**, *10*, 1005. [[CrossRef](#)]
85. Sim, S.C.; Van Deynze, A.; Stoffel, K.; Douches, D.S.; Zarka, D.; Ganal, M.W.; Chetelat, R.T.; Hutton, S.F.; Scott, J.W.; Gardner, R.G.; et al. High-Density SNP Genotyping of Tomato (*Solanum lycopersicum* L.) Reveals Patterns of Genetic Variation Due to Breeding. *PLoS ONE* **2012**, *7*, e45520. [[CrossRef](#)] [[PubMed](#)]
86. Smykal, P.; Nelson, M.N.; Berger, J.D.; von Wettberg, E.J.B. The impact of genetic changes during crop domestication. *Agron* **2018**, *8*, 119. [[CrossRef](#)]
87. Gao, L.; Gonda, I.; Sun, H.; Ma, Q.; Bao, K.; Tieman, D.M.; Burzynski-Chang, E.A.; Fish, T.L.; Stromberg, K.A.; Sacks, G.L.; et al. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat. Genet.* **2019**, *51*, 1044–1051. [[CrossRef](#)] [[PubMed](#)]
88. Gonda, I.; Ashrafi, H.; Lyon, D.A.; Strickler, S.R.; Hulse-Kemp, A.M.; Ma, Q.; Sun, H.; Stoffel, K.; Powell, A.F.; Futrell, S.; et al. Sequencing-based bin map construction of a tomato mapping population, facilitating high-resolution quantitative trait loci detection. *Plant. Genome* **2019**, *12*, 180010. [[CrossRef](#)] [[PubMed](#)]
89. Maltecca, C.; Parker, K.L.; Cassidy, J.P. Application of multiple shrinkage methods to genomic predictions. *J. Anim. Sci.* **2012**, *90*, 1777–1787. [[CrossRef](#)]
90. 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](#)]
91. Whittaker, J.C.; Thompson, R.; Denham, M.C. Marker-assisted selection using ridge regression. *Genet. Res.* **2000**, *75*, 249–252. [[CrossRef](#)]
92. De los Campos, G.; Hickey, J.M.; Pong-Wong, R.; Daetwyler, H.D.; Calus, M.P.L. Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics* **2013**, *193*, 327–345. [[CrossRef](#)]
93. Bassi, F.M.; Bentley, A.R.; Charmet, G.; Ortiz, R.; Crossa, J. Breeding Schemes for the Implementation of Genomic Selection in Wheat (*Triticum* Spp.). *Plant. Sci.* **2016**, *242*, 23–36. [[CrossRef](#)]
94. Falk, D.E. Generating and maintaining diversity at the elite level in crop breeding. *Genome* **2010**, *53*, 982–991. [[CrossRef](#)]
95. Rahim, M.S.; Bhandawat, A.; Rana, N.; Sharma, H.; Parveen, A.; Kumar, P.; Madhawan, A.; Bisht, A.; Sonah, H.; Sharma, T.R.; et al. Genomic Selection in Cereal Crops: Methods and Applications. In *Accelerated Plant Breeding: Cereal Crops*; Gosal, S.S., Wani, S.H., Eds.; Springer: Berlin/Heidelberg, Germany, 2020.
96. Ercolano, M.R.; Sacco, A.; Ferriello, F.; D'Alessandro, R.; Tononi, P.; Traini, A.; Barone, A.; Zago, E.; Chiusano, M.L.; Buson, G.; et al. Patchwork sequencing of tomato San Marzano and Vesuviano varieties highlights genome-wide variations. *BMC Genom.* **2014**, *15*, 138. [[CrossRef](#)] [[PubMed](#)]
97. Kanbar, A.; Kondo, K.; Shashidhar, H.E. Comparative efficiency of pedigree, modified bulk and single seed descent breeding methods of selection for developing high-yielding lines in rice (*Oryza sativa* L.) under aerobic condition. *Electron. J. Plant. Breed.* **2011**, *2*, 184–193.
98. Breseghello, F.; Morais, O.P.; Castro, E.M.; Prabhu, S.A.; Bassinello, P.Z.; Pereira, J.P.; Utumi, M.M.; Ferreira, M.E.; Soares, A.A. Recurrent selection resulted in rapid genetic gain for upland rice in Brazil. *Int. Rice Res. Notes* **2009**, *34*, 1–4. [[CrossRef](#)]
99. Shelton, A.; Tracy, W.; Shelton, A.C.; Tracy, W.F. Recurrent selection and participatory plant breeding for improvement of two organic open-pollinated sweet corn (*Zea mays* L.) populations. *Sustainability* **2015**, *7*, 5139–5152. [[CrossRef](#)]

