

Review

# NGS-Based Genotyping, High-Throughput Phenotyping and Genome-Wide Association Studies Laid the Foundations for Next-Generation Breeding in Horticultural Crops

Nunzio D'Agostino \*  and Pasquale Tripodi 

CREA Research Centre for Vegetable and Ornamental Crops, 84098 Pontecagnano Faiano, Italy; pasquale.tripodi@crea.gov.it

\* Correspondence: nunzio.dagostino@crea.gov.it; Tel.: +39-089-386-243

Received: 26 July 2017; Accepted: 13 September 2017; Published: 15 September 2017

**Abstract:** Demographic trends and changes to climate require a more efficient use of plant genetic resources in breeding programs. Indeed, the release of high-yielding varieties has resulted in crop genetic erosion and loss of diversity. This has produced an increased susceptibility to severe stresses and a reduction of several food quality parameters. Next generation sequencing (NGS) technologies are being increasingly used to explore “gene space” and to provide high-resolution profiling of nucleotide variation within germplasm collections. On the other hand, advances in high-throughput phenotyping are bridging the genotype-to-phenotype gap in crop selection. The combination of allelic and phenotypic data points via genome-wide association studies is facilitating the discovery of genetic *loci* that are associated with key agronomic traits. In this review, we provide a brief overview on the latest NGS-based and phenotyping technologies and on their role to unlocking the genetic potential of vegetable crops; then, we discuss the paradigm shift that is underway in horticultural crop breeding.

**Keywords:** plant genetic resources; vegetable crops; horticulture; single nucleotide polymorphisms; genetic diversity; phenotyping; genome-wide association studies; genomic selection

---

## 1. The Use of Plant Genetic Resources in Vegetable Crop Improvement

The main challenges for vegetable crop improvement are linked to the sustainable development of agriculture, food security, evolution of dietary styles, the growing consumers' demand for food quality, the spread of non-communicable diseases, and, finally, the under-nutrition due to deficiencies in vitamins and minerals (“hidden hunger”). The design and the development of breeding programs can provide effective responses to all of these challenges. One of the recognized hubs in breeding activities is the exploitation of crop genetic diversity aiming at the beneficial allele-hunting process.

The first part of this review will focus on the importance of biodiversity in horticultural crops, on the main initiatives for germplasm conservation and on the assessment of potentiality and constraints for the use of biodiversity in breeding programs.

### 1.1. Erosion of Genetic Diversity in Crops

Loss of genetic diversity in crop species, also referred to as genetic erosion, is a step-by-step process that has begun with human population growth and the expansion of human activities whose effects have had serious consequences in ancient, traditional, and modern agriculture.

Early farmers gradually abandoned their nomadic hunter-gatherer habits in favor of semi-sedentary/sedentary agriculture as the primary mode of supplying plant food resources [1].

This process, which is dated back to 12,000 years ago, is known as plant domestication and involved a broad spectrum of transitions, which has led to an increased adaptation of plants to cultivation and utilization by humans [2]. In vegetable crops, the “domestication syndrome” led to combinations of different traits, such as changes in plant architecture and reproductive strategy, the increase of fruit and seed size, and the loss of secondary metabolites [3].

Although the phases of domestication and diversification have not been homogeneous in all areas of the World, in the early stages they have caused the selection of naturally occurring variants in the wild ancestors of crops (i.e., crop wild relatives; CWR) to perpetuate only those species suitable to survive in various agro-ecological habitats (pre-domestication). Afterwards, continuous rounds of selection have completely re-shaped wild into domesticated species [1]. Such a process, known as selective breeding, lasted until early 1900’s generating a multitude of landraces (LR) with an intermediate level of variability and differentiation that have been mainly selected for specific adaptation to local environments and for desirable quality traits.

An exciting example on the history of domestication and genetic improvement concerns the cultivated tomato (*Solanum lycopersicum* L.). *S. pimpinellifolium*, the small-fruited species considered to be its wild ancestor, has been subjected to main domestication cycles, firstly in Peru and Ecuador, and later in Mexico [4]. The resulting domesticated cultivars have been brought to Europe and then spread all over the World, where an intensive selection has been carried out in the last two centuries leading to cultivars with an increased fruit size (~100 times larger than its ancestor) [5], but less flavorful than heirloom varieties [6].

The advent of modern agriculture and the impact of the “Green Revolution”, from the second half of the 20th century onwards, have introduced high-yielding and phenotypically uniform varieties that are better adapted to industrial agriculture (Figure 1). However, despite the gains breeders have seen during the past years, most of the natural variability has been lost as farmers abandoned traditional landraces in favor of hybrids that are generally more marketable and valuable. This has led to a sharp reduction of genetic diversity, which made crop substantially devoid of differences in quality traits [6]. Although cultivars carrying resistances are continuously released, a reduction of genetic diversity and the extensive use of few crop varieties can lead to critical consequences in the emergence of new pests and diseases [7].

It has been estimated that only 200 out of more than 275,000 species of flowering plants have been domesticated [8,9], and a relatively small number of these accounts for 95% of food supply [10]. Among horticultural crops, species belonging to the Solanaceae family along with those from Poaceae, Fabaceae, and Rosaceae explain more than 60% of the World calories [11]. Food production based on such restricted number of species represents a risk that should be not left out in the next decades with the human population is expected to reach 9.6 billion by 2050 [12], the occurrence of extreme weather events, the reduced availability of natural resources and the related increase in food demand [13]. If on one hand, the number of crop commodities contributing to food supply is expected to grow [14], on the other hand, minor crops cannot completely face the challenges of the new millennium. For the latter reason, CWR and LR, being tailored to extreme conditions and not having been particularly conditioned by the selection process, might be exploited for breeding purposes with innovative genomic and phenomic approaches since they are valuable repositories of traits. Indeed, the thorough exploration of plant genetic resources (PGR) and the preservation and use of CWR and LR are primary targets for the future progress of breeding programs.

## 1.2. Strategies for Collection and Conservation of Plant Genetic Resources

The first concerns on the effects of genetic erosion date back to the half of the 20th century, when Vavilov called attention in crop diversity conservation, motivating the institution of gene banks [15,16]. Since then, strategies of in-situ and ex-situ management have been proposed [17]. In-situ conservation is generally applied in the natural habitat where crop is cultivated and offers the possibility to maintain the ecosystem health as well as the dynamic evolution of CWR diversity in

relation to parallel environmental changes [18]. However, these advantages do not match the easiness to define a comprehensive network of protected areas for all species [17]. Ex situ strategies better fit with the possibility to conserve genetically representative populations in designed areas other than natural habitats.

Seeds are the most common form of conservation due to the relatively low cost of management. Other methods include parts of plants in tissue culture or cryo-preserved and mature individuals in field collections (i.e., arboretum). To date, more than 1750 gene banks conserve ~7 million of accessions worldwide in the form of seeds [19].

Although preserving PGR is essential to ensure the development of new crop varieties, the lack of adequate data on the accessions stored in gene banks is the main issue that hinders their effective use [20]. The large size of these collections complicates their characterization, making the identification of novel/beneficial alleles expensive. An approach that enhances the exploitation of the huge variability present in gene banks is the development of core collections (CC), which are represented by a reduced set of accessions capturing most of the variability with minimum repetitiveness [21]. While the definition of CC was mainly based on ecological and geographical information, the advent of novel genomics technologies allowed the efficiency of the selection and utilization of large germplasm collections stored in gene banks to be improved. Since then, several studies have been published, and different types of collections have been proposed [22].

By exploiting the information stored in different databases, a CC can be assembled fairly easily on an as-needed basis depending on target trait(s). We report as examples the tool developed by the Centre for Genetic Resource (CGN; The Netherlands), that allows a maximum diversity subset of accessions to be selected on-line based on user defined selection criteria (<https://goo.gl/XhEuwR>); the PowerCore program [23], and the Signal Processing Tool method recently developed by Borraro and Takeya [24].

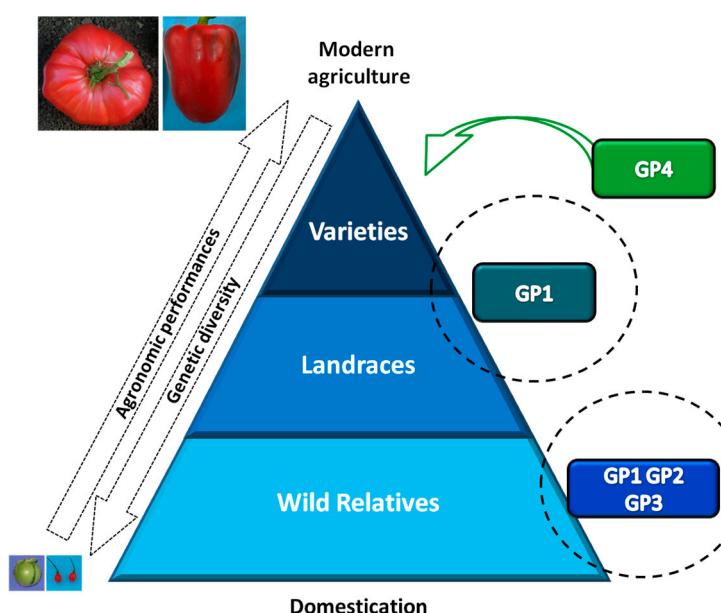
An additional approach, which relies on the combination of environmental data with specific plant characteristics, is based on the Focused Identification of Germplasm Strategy (FIGS). FIGS allows for the identification of a core of accessions with a higher probability of containing specific “target” traits subjected to the selection pressure of the environment from which they were originally sampled [25]. The method, which maximizes the possibility to capture specific adaptive traits by means of high-throughput geographic information system (GIS) technologies, has been mainly used to identify sources of resistance to abiotic and biotic stresses in cereals, but has not yet successfully applied in horticultural crops [26]. Despite the potentiality of CC, debate on their effectiveness and criteria of selection to adopt are still underway [27].

Collection, conservation, and utilization of PGR in agriculture are therefore major issues involving several international bodies and substantial investments. The International Treaty on Plant Genetic Resources for Food and Agriculture (FAO) and the Convention on Biological Diversity (CBD) have defined strategies to avoid loss of plant genetic diversity estimated to be 25–35% [28]. Established efforts have included actions to improve the effectiveness in-situ conservation, secure safety ex-situ storage, strengthening efforts of public and private breeders in PGR characterization and utilization [28]. More recently, global research alliances [29] and transnational projects [30,31] have been stipulated with the aim to unlock the potential of vegetable crop diversity by means of innovative genomics and phenomic approaches, and make it available for researchers and breeders.

### 1.3. Importance of Plant Genetic Resources and Biodiversity in Breeding Programs

In the last 35 years, several breeding programs involving the use of CWR have been developed for a wide range of crops. The pioneering studies conducted since the early nineties evidenced how wild relatives can contribute with their beneficial traits to crop improvement, highlighting the key role of genome mapping for the efficient use of genetic diversity [32]. Furthermore, advances in the phylogeny and taxonomy of plant species as well as the development of novel molecular technologies allowed the exploitation of genetic variability even in more distantly related *taxa*. The transfer of novel/beneficial

alleles from wild to cultivated species is not always easy to pursue, and it requires that reproductive barriers between different gene pools be overcome (Figure 1). Gene pools [33,34], referred to as the portion of genetic diversity available for breeding, are defined based on the cross-ability between the crop itself and the primarily non-domesticated species. Nevertheless, most of the potential resources to be used in breeding programs are in the primary gene pool, where, as general rule, gene transfer is easy and immediate. Conventional techniques that rely on the cross between crops and their close wild relatives are still in the norm [17], even if novel plant breeding techniques (NPBT) provide opportunities to: (i) transfer genes from genotypes of the same or sexually compatible species (cisgenesis and intragenesis); (ii) induce mutations in target genes; or, (iii) investigate species in distantly related *taxa* as a useful and extensive reservoir of alleles [35].



**Figure 1.** Global scenario from domestication to modern agriculture. Wild species of tomato and pepper (*Solanum habrochaites* and *Capsicum chacoense* at the base of the triangle) are characterized by wide genetic variability, which can be used to improve modern varieties (*Solanum lycopersicum* and *Capsicum annuum*, at the top of the triangle). An intermediate step of domestication is represented by landraces which have a broadening genetic variation linked to adaptation to local environments. Transfer of alleles can be possible within gene pools (GP) [33]. Four different GP levels include: (i) species with easy crossing ability resulting in fruitful hybrids and fertile off-springs (primary gene pool, GP1); (ii) less closely related species that generates weak or sterile hybrids and are characterized by difficulty in obtaining advanced generations (secondary gene pool, GP2); (iii) species requiring sophisticated techniques for gene transfer such as embryo rescue, somatic fusion, grafting, and bridge species (tertiary gene pool, GP3); and, (iv) distantly-related species belonging to different families or kingdoms for which gene transfer is not possible sexually but through direct gene transfer by means of genetic engineering (fourth gene pool, GP4).

Studies on the use of NPBT are being increasingly published. A non-exhaustive but extensive list of genome editing approaches applied to vegetable crops is reported in a recently published review we co-authored [36].

Despite their potentiality, CWR have not been well exploited due to their phenotype that is unsuitable for modern agriculture and to the poor value of economically important features, such as yield. Even more, the quantitative inheritance of relevant agronomic traits is often in linkage with undesirable characteristics (i.e., linkage drag) [15], which makes the use of wild species complicated since it would require efficient selection procedures [37].

The dissection of wild germplasm and the identification of genes underlying quantitative traits have been a central target over the past 35 years. Introgression of hundreds of genes from wild to cultivated species has been possible with the advent of DNA sequencing technologies, which have facilitated the establishment of numerous experimental mapping populations and related linkage maps in many horticultural species [38]. Bi-parental crosses offered the chance to fix alleles from exotic materials in advanced generations (inbred backcross lines, IBLs; recombinant inbred lines, RILs) leading to the identification of several quantitative trait *loci* (QTL) [39]. This approach, still largely used, has the main limitation that is due to the fact that only allelic diversity between the two parents is investigated. In addition, a lack of recombination due to continuous self-fertilization cycles reduces mapping resolution.

Advances in genomics have enabled the implementation of several genome-wide association studies (GWAS) [40] in order to investigate the larger variation present in CC. As it will be discussed in a greater extent at a later stage of this review, GWAS leads to better precision mapping through the identification of single nucleotide polymorphisms (SNPs) strictly related to traits of interest. Main constraints affect this method, such as the lacking of gene flow from wild relatives and several other drawbacks, as reviewed by Korte and Farlow [41].

To address these limitations, multi-parent advanced generation inter-cross (MAGIC) populations [42] are being developed to increase mapping resolution through multiple generations of recombination and provide a high statistical power afforded by a linkage-based design [43]. Furthermore, the combination of genomes from more founders allows for a larger allelic diversity to be explored, generating new phenotypes that constitute a highly valuable pre-breeding resource [44]. MAGIC, hence, represents an intermediate population, which overcomes the main constraints of bi-parental and GWAS. However, the statistical complexity of the analysis and the time required for its development make MAGIC utilization still challenging. As for horticultural species, MAGIC populations have been developed only in tomato to investigate the genetic basis of fruit weight [44].

Genome-wide introgression lines (ILs) represent a further option that increases both mapping resolution and statistical power for minor QTL detection. When compared with the populations mentioned above, each IL includes only small genomic regions from CWR [37] as the results of marker-assisted selection at early stages. Despite the large efforts required for their development, ILs represent a valid source for both genetic studies and breeding purposes, having the advantage of wiping out the linkage drag effect by transferring only the *loci* of interest [45].

A comprehensive use of exotic germplasm is still far away. High-throughput technologies in the field of genomics and phenomics will allow speediness and accuracy in germplasm characterization to be enhanced and will play a central role in the coming decades, leading to a unique opportunity for next generation precision breeding.

## 2. NGS-Based Genotyping for Genetic Diversity Evaluation

In the pre-genomic era, the assessment of genetic diversity has been traditionally carried out via morphological and cytogenetic characterization, or through the analysis of isozymes. Since '90s onwards, the techniques based on different types of DNA molecular markers have been preferentially used to measure the level of genetic variation [46]. Marker types and genotyping techniques have evolved over time; indeed time consuming, too costly, cumbersome and/or challenging techniques have been replaced by simpler, less expensive, and more efficient alternatives. Because of their extraordinary abundance in the genome and their usually bi-allelic nature, single nucleotide polymorphisms have quickly become the markers of choice to dissect the genetic variability of PGR [47].

The release into the public domain of complete, near-complete or partial genome sequence of the most important vegetable crops [36], the development of SNP detection assays [48,49] and high-density genotyping arrays [50–52] caused a shift from small-/medium-scale to large-scale SNP genotyping.

Both of the above methods allow for thousands of SNPs to be discovered and ample genetic variability across germplasm collections to be captured; however, their design relies on a priori

knowledge of the sequence space of the species under investigation and usually cannot be easily modified to fit in with custom experimental designs. This has encouraged the development of novel but still high-throughput and time- and cost-saving SNP discovery methods based on next-generation sequencing (NGS) technologies. Several strategies, methods and protocols for NGS-based genotyping have been developed so far [53,54]. We will go through some of these with a special focus on genotype-by-sequencing (GBS).

NGS, coupled with the availability of high-quality reference genomes of horticultural crops, have expedited the re-sequencing of many individuals to identify a large number of SNPs and investigate within- and between-species sequence variation [5,55–58]. Although the above cited re-sequencing projects were fruitful from a scientific standpoint, the re-sequencing was characterized by low depth of coverage for several individuals. Indeed, within these studies, the average sequence depth varied from 5-fold to 36-fold coverage. Obviously, the greater the depth of coverage is, than the greater the reliability of SNP calling. As general rule, a minimum of 10–20× coverage depth is indicated for reliable variant calling. Unfortunately, because of cost issues this standard is not always applicable.

An alternative strategy to whole genome re-sequencing is to generate a reduced representation of the genome by using on-array- or liquid-based hybridization methods to enrich and capture target genomic regions and possibly identify promising alleles having a potential application in crop breeding programs [59–61].

Unlike sequence capture and targeted re-sequencing, restriction enzyme-based enrichment techniques [62], while not allowing specific target sequences in the genome to be investigated, are the methods of choice for SNP discovery and genotyping.

These methods include three common steps: (i) DNA digestion with restriction enzymes (REs); (ii) ligation with sequencing platform-specific adapters; and, (iii) PCR amplification to increase the DNA yield into sequencing libraries. Conversely, they differ for the size selection step (that is necessary to filter out DNA fragments of desired size) that can be carried out at any point in the protocol workflow or it may be entirely dismissed. Obviously, different protocols result in different data outcomes. Even if it is desirable to dealing with dense distribution of SNP markers and uniform sequence coverage across samples, all of these methods have limitations due to inconsistency in the number of: (i) reads per sample; (ii) reads per polymorphic site; and, (iii) sequenced sites per sample. The combination of all these items can result in a huge number of low quality or missing data. Jiang, et al. [63] suggested a number of improvements to be made in individual steps of the wet-lab workflow to minimize biases in NGS library construction and to increase the degree of reliability and robustness of downstream sequence data.

The relevance of isolating, enriching, and sequencing of specific genomic *loci* for SNP discovery, was initially proved with restriction site-associated DNA sequencing (RAD-Seq) [64]. However, in a short while, RAD-Seq was paralleled and replaced by GBS [65]. At present, GBS is the most favorite technology for high-throughput SNP discovery and it is generating remarkable knowledge on the nature and extent of genetic diversity within germplasm collections [66,67]. GBS was originally applied in maize to identify SNP markers in a RIL population [65]. Since then, several studies have been reported in literature for large-scale SNP discovery and genotyping of RILs [68,69], bi-parental mapping population segregating for important agronomical traits [70,71]; ethylmethane sulfonate- (EMS) induced mutant populations [72,73]; and unrelated individuals in small or large size populations [74,75]. All of these efforts, most of which were carried out on vegetable crop species, aimed to have available a high number of SNP data points for concomitant or future genome-wide association studies.

Usually, ApeKI is the restriction endonuclease used to produce restriction fragments among individuals in a population. It is a methylation sensitive enzyme that shows preferential cleavage for lower or single copy regions of the genome that are generally the richest in genes. In 2012, Poland, et al. [76] modified the original protocol based on ApeKI, by using two restriction enzymes (PstI/MspI) to achieve a higher SNP density. Of course rare-cutter, frequent cutter, or methylation insensitive RE(s) can be used alone or in combination to obtain the desired number of SNP markers.

Intuitively, one major limitation of GBS is its random access to the genome. Such randomness can be adjusted by appropriately selecting the combination of REs to be used. Moreover, the choice of RE(s) is crucial since it affects DNA fragment size distribution as well as the number of fragments in the GBS library. These two parameters, in turn, influence sequencing depth and ultimately the number of SNPs identified [77]. GBS has been used in a large list of vegetable crops showing that this method is efficient for large-scale, low cost genotyping despite all of the limitations mentioned above.

As a rule, GBS tags are aligned to a reference genome to identify SNPs from aligned tags. Several reference-based SNP calling pipelines are available so far [78]. Although a reference genome can facilitate GBS data analysis, several reference-independent SNP calling pipeline (e.g., Stacks and TASSEL-UNEAK) have been developed [79,80] and successfully applied for PGR diversity studies [81,82].

The starting list of SNPs is then subjected to filtering (call rate; minimum depth of coverage; and minor allele frequency, etc.) and, in some cases, it can also be pruned by removing all of the SNPs that are in high linkage disequilibrium (LD). These operations generate high quality SNP datasets that are fed into the population structure analysis software. Such software are able to categorize individuals into ethnically similar cluster based on allele frequency estimates [83,84], or not [85]. Recently, we proved the strength of combining a parametric (STRUCTURE) with a non-parametric method (AWclust) in defining the genetic structure of a population of *Capsicum annuum* accessions [74].

More recently, next-generation marker genotyping platforms have evolved to address specific issues. Yang, et al. [86] developed a semi-automated primer design pipeline to convert GBS-derived SNPs into amplicon sequencing (AmpSeq) markers. This has become necessary because GBS alone is unworkable for highly heterozygous species, for which a large number of missing data and the under-calling of heterozygous sites are very common. The AmpSeq strategy starts from the design of primer pairs by using GBS tags as templates; then, the resulting amplicons are used for genotyping via NGS. In this way, it is possible to circumvent GBS-related issues and to develop reliable SNP-based markers for the high-throughput screening of heterozygous crops.

A further technique, named rAMPSeq, has been established by Buckler, et al. [87] to develop repetitive sequence-based markers for robust genotyping. Although the authors are well aware that the design of rAMPSeq sacrifice several strengths of the GBS method, they feel that the method they have proposed can be revolutionary for breeding and conservation biology.

All considered, the key role of NGS on SNP markers identification, genetic diversity assessment, and population structure analysis in horticultural crops is unquestionable, regardless of the genotyping strategies available. With its unprecedented throughput and scalability, NGS is enabling investigations on genetic diversity at a level never before possible.

### 3. Advanced Phenomics in Plant Breeding

Phenomics is becoming increasingly important in genetic studies and precision agriculture since it allows for the accurate characterization of multiple traits in crops and a better understanding of phenotypic changes due to underlying heritable genetic variation.

As reported in the previous section, the tremendous advancement of cutting-edge sequencing technologies allowed the identification of thousands of SNPs at affordable costs; at the same time, methods to assess plant traits progressed more slowly, generating what is known as the “phenotyping bottleneck” [88]. Biochemical and metabolomic phenotyping of quality traits and especially abiotic and biotic stress evaluation are the main cause of the “phenotyping bottleneck”. Further on in the text, we will provide several examples on how imaging methods are being used to collect phenotypic data points for complex traits in vegetable crops.

The need for increasing the ability to investigate a large amount of traits in a non-destructive manner and with an acceptable accuracy has become a major target in plant breeding. The labor-intensive and costly nature of phenotyping, motivated, in the past decade, the research community to develop automated technologies for high-throughput plant phenotyping (HTPP). This has guaranteed massive

progress in the dissection of a wide range of qualitative, agronomical, morphological, and physiological traits, as well in the investigation of traits related to biotic and abiotic stresses.

These technologies, which rely on automated non-invasive sensing methods, are able to capture plant features on the basis of image analysis. At present, several systems are available, including conventional RGB/CIR cameras, spectroscopy (multispectral and hyperspectral remote sensing), thermal infrared systems, fluorescence and tridimensional (3D) imaging, and magnetic resonance imagers (MRI) [89,90].

Features and potentialities of these tools are widely documented [91]; herein, we briefly list which implications (benefits and constraints) emerged from their use on horticultural crop phenotyping.

RGB/CIR cameras are extensively used to analyze traits such as plant canopy and biomass, which can be estimated by combining the imagery falling into red, blue, and green light (RGB), and color infrared (CIR). Derived indexes can be further calculated as the “leaf area index” (LAI) or “normalized difference vegetation index” (NDVI). Multispectral (MS) and hyperspectral (HS) analysis are more addressed to those traits linked to crop physiological status (e.g., nutrient and water content, photosynthetic efficiency, etc.) using a set of images which covers the entire range of radiation from visible (VIS; 400–765 nm) to near-infrared (NIR; 765 to 3200 nm). Thermal imaging is better suited to evaluate (i) the stage of infections by pathogens or if diseases are spreading through the crops and (ii) responses to abiotic stresses (e.g., drought and heat tolerance). It is based on measurements of the variation in foliage surface temperature, which is related to differences in stomatal conductance due to various stresses that alter water balance and transpiration [92]. Plant metabolic status can be assessed by means of fluorescence, capturing excitation, and emission spectra during the absorption of radiation in shorter wavelengths by chlorophyll [90]. Vegetation canopy structure and topographic maps can be estimated by 3D imaging, while MRI detects nuclear resonance signals from isotopes to generate images of the internal structures of the plant. These devices can be used alone or in combination to deliver different data outcomes.

These tools have been firstly applied in cereals, *Arabidopsis*, and industrial crops [90]. Studies in horticultural crops are currently being conducted. Chlorophyll fluorescence imaging has been used in tomato transgenic plants and young industrial chicory to investigate drought tolerance [93] and cold stress resistance [94], respectively. RGB cameras and 3D imaging have been applied in recombinant inbreds of pepper to inspect canopy structure and plant architecture [95]. MRI have been applied in bean to resolve root structure under resource competition [96].

While analyses in controlled-environments provided major breakthroughs for targeted applications, such as the possibility to investigate roots and seeds in a non-destructive manner by means of MRI-3D combination, analyses in open-field are much more laborious and troublesome due to the heterogeneous nature of soils and the inability to control external factors (e.g., climate).

When evaluating different classes of traits, their nature must be taken into account and how they interact with the external environment. As an example, fruit morphometric traits, can be easily and accurately assessed in a repetitive manner on plants cultivated in greenhouse as well as roots in hydroponics, aeroponics or pots. Contrariwise, yield potential as well as the detrimental effects of abiotic stresses, are better evaluated via field-based phenotyping (FBP) experiments [97]. FBP is generally applied to estimate the genotype  $\times$  environment (GxE) interaction in a large number of individuals in wide collections or extensive mapping populations [98]. Platforms carrying multiple sensors on wheeled vehicles can be used to measure plant traits on extensive surfaces and investigate GxE interaction. A successful example is provided by the Field ScanAnalyzer (LemnaTec, Aachen, Germany), which is a fully automated system designed to capture deep phenotyping data (growth and physiology traits throughout the crop cycle) from crops growing in field environments. Other options are provided by drones equipped with multispectral imaging systems that are able to collect thousands of images in a relatively small amount of time. These technologies are having a positive impact in the reliable and accurate estimate of phenotypic data points as well as in the development of high-yielding, stress tolerant and disease resistant plant varieties.

Although knowledge on semi- and fully-automatic phenotyping systems is rapidly spreading within scientific community, their accessibility still represents a major constraint due to high costs and difficulties in data management. Indeed, one of the main challenges of HTPP and FBP is the processing and analysis of millions of captured images. The development of software for managing and analyzing large and complex datasets is underway to better integrate the different sources of data underlying various plant developmental processes [98].

#### 4. Linking Genotype to Phenotype

In recent times, the selection process carried out by breeders has largely benefited from the improvement of genotyping and phenotyping techniques, but primarily it has taken advantage of the link between phenotypic and genotypic data sets. Uncovering genotype-phenotype relationships is a key requirement in establishing a basic understanding of complex traits and to elucidate how genetic differences can determine individual differences in agronomic performances.

Technical advancements in high-throughput genotyping, in the accuracy and mechanization of phenotyping, as well as the development of efficient statistical methods and bioinformatic tools laid the foundation for genome-wide association studies. This approach, introduced for the first time in human genetics a dozen of years ago [99], is becoming the most popular method to statistically associate genomic *loci* (also referred to as QTN; quantitative trait nucleotides) to simple or complex traits of interest [100]. Even more, the results from GWAS have also been employed to understanding the genetic bases of the domestication and artificial selection processes and to point out how breeding has influenced the genetic variability of modern crops [101]. Below are some examples from recent literature related to the application of GWAS in vegetable crops to reveal useful alleles in genes encoding for agronomic traits.

A genome-wide association study based on a SNP catalogue from whole-genome re-sequencing of 398 modern, heirloom, and wild tomato accessions, and a targeted metabolome quantification of sugars, acids, and volatiles permitted identification of candidate genetic *loci* capable of altering 21 of the chemicals contributing to consumer liking as well as to overall flavor intensity of tomatoes [6]. A similar study founded on polymorphic simple sequence repeat (SSR) markers and volatile quantification in a collection of 174 diverse tomato accessions was performed by Zhang, et al. [102]. The authors identified via GWAS already known as well as novel *loci* that could be important in controlling the volatile metabolism in tomato.

Genome-wide association studies based on SNPs generated by GBS have proved effective for the identification of new alleles affecting morpho-agronomic and fruit quality-related traits in horticultural crops. As an example, we report the work by Nimmakayala, et al. [103], which had the purpose of identifying SNP markers associated with capsaicinoid content and fruit weight traits in pepper (*Capsicum annuum* L.). A similar approach was used by Pavan, et al. [104], who performed preliminary GWAS and identified SNP(s)/trait associations for flowering time and seed-related traits in a collection of 72 accessions of *Cucumis melo* from Apulia. Additional GWAS for fruit firmness were performed in a larger collection of melon from Asia and Western hemisphere [105].

A further genome-wide association study by Cericola, et al. [106], focused on fruit-related traits but based on SNPs from RAD-tag sequencing, was performed in eggplant (*Solanum melongena* L.). A panel of 191 accessions, including breeding lines, old varieties, and landraces was genotyped and scored for anthocyanin pigmentation and fruit color. As result, different already known QTLs were validated and novel marker/trait associations were observed.

While the number of GWAS for quality traits is increasing rapidly, there are few papers describing the association analysis aimed at the identification of alleles for resistance or tolerance to abiotic or biotic stresses in horticultural crops.

To the best of our knowledge, no studies have been published so far aiming at the detection of significant genotype–phenotype associations for abiotic stress (e.g., drought; heat, flooding, salinity, cold).

As for biotic stresses, we did not find any work on fruit or leafy vegetables. We can report only a few examples in legumes. Hart and Griffiths [107] used GBS to genotype a set of RILs in common bean (*Phaseolus vulgaris* L.) and GWAS to identify SNPs associated with the resistance to *bean yellow mosaic virus*. The GBS approach was also used by Saxena, et al. [108] for SNP identification and genotyping of three different mapping populations (2 RILs and one F<sub>2</sub>) segregating for sterility mosaic disease (SMD) in pigeonpea (*Cajanus cajan* (L.) Millspaugh). The authors identified a total of 10 QTLs, including three major QTLs governing SMD resistance in pigeonpea.

GWAS are based on three cornerstones: (i) a high-density marker catalogue from genome-wide assessment of the diversity among individuals in a population; (ii) knowledge on population structure and allele frequency spectrum; and, (iii) robust phenotypic data for each individual within the population (Figure 2).

In addition to a large panel of markers (preferentially thousands of SNPs), GWAS require a large sample size (i.e., from hundreds to thousands of individuals) enclosing the maximum variability for the traits under study to achieve adequate statistical power. Population sampling is one of the key steps and should always include accessions collected from different locations and possibly hierarchically structured. Indeed, population structure represents a considerable source of confounding in GWAS [41]; for that reason, a suitable characterization of population structure is advisable to prevent false positive (type I error) or negative (type II error) SNP(s)-trait associations.

As for phenotypic records, it is more advisable to deal with quantitative data (integer or real-valued numbers) rather than binary data (presence/absence). As it is easily perceived, the former provides more statistical power than the latter.

Statistical analysis for capturing significant genotype–phenotype associations is performed by examining each SNP independently (multiple testing). In general, only those markers that meet the Bonferroni corrected *p*-value threshold ( $\alpha = 0.05$ ) can be considered significantly associated with phenotypic observations. Alternative approaches to Bonferroni correction have been proposed for establishing significance in GWAS: among these, the use of false discovery rate (FDR) or permutation testing [109].

Association analysis can be performed via general linear model (GLM) or mixed linear model (MLM). It has been demonstrated that the statistical power of MLM is higher than that of GLM and better suited to the quantitative nature of the traits under study, markedly reducing type I errors [110]. This is possible because MLM is able to handle the population structure and the covariance between individuals due to genetic relatedness (i.e., kinship matrix). In other words, MLM takes into account the confounding effect of the genetic background.

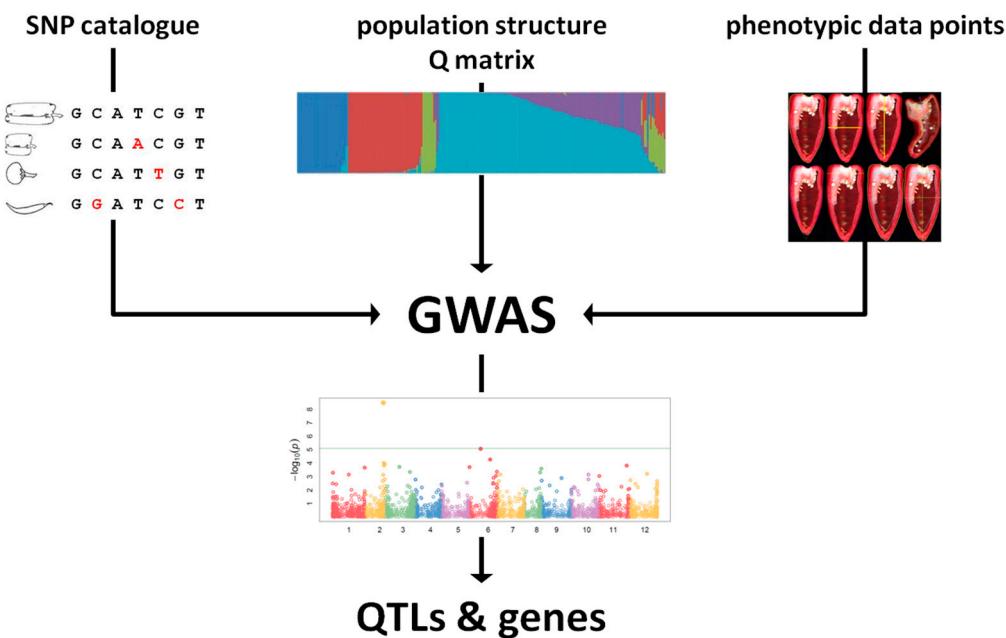
Even though GWAS is by now a recognized and powerful tool to investigate genotype–phenotype relationship, it still suffers of some limitations: (i) rare alleles cannot be detected; (ii) alleles having small effects on the phenotype are difficult to be identified; and (iii) the most significant SNPs are not always the true causative/predictive factors for a given trait.

It is very likely that genomic selection (GS) strategies will replace GWAS in the next future [111]. Albeit, comparable to GWAS, GS relies on high-throughput genotyping and phenotyping of unrelated individuals in a population (i.e., training population; TP) and allows markers in strong LD with causal variants to be identified [112], it has singular features. Results from TP are used to predict genomic estimated breeding values (GEBV) on a different set of individuals (i.e., breeding population; BP), which has been previously genotyped with dense marker coverage across genome, but whose phenotype is unknown. In this way, time and expense in phenotyping diminish considerably since phenotyping is needed only in the initial phase to enhance the accuracy of the prediction model. In crops GS has been mainly applied in cereals [113,114], while emerging applications in vegetables are reported only for tomato [115,116]. Both works on tomato assessed the potential of GS approach to improve fruit quality traits. Duangjit, et al. [115] used a panel of 163 tomato accessions that were genotyped using the SolCAP Infinium assay (Illumina) and phenotyped for a total of 35 metabolic

traits. Training and breeding populations were assembled including an increasing/decreasing number of accessions from the original population.

Yamamoto, et al. [116] used 96 big-fruited F<sub>1</sub> tomato varieties as a TP (that was phenotyped for soluble solids content and total fruit weight) and evaluated the predictability of the GS models in their progeny populations.

These pioneering studies demonstrated that GS models could predict phenotypes paving the way for a successful implementation of GS in vegetable crops.



**Figure 2.** Genome-wide association studies (GWAS) are based on three pillars: (i) a high-density single nucleotide polymorphisms (SNP) catalogue derived from diversity assessment of individuals in a germplasm collection; (ii) knowledge on population structure and allele frequency spectrum (Q matrix); and, (iii) phenotypic data points for each individual within the population. Generally results from GWAS are displayed as Manhattan plots showing genome-wide *p*-values of SNP(s)-trait associations. GWAS allow the causal gene(s) or QTL(s) associated with the trait(s) of interest to be identified.

## 5. Outlook

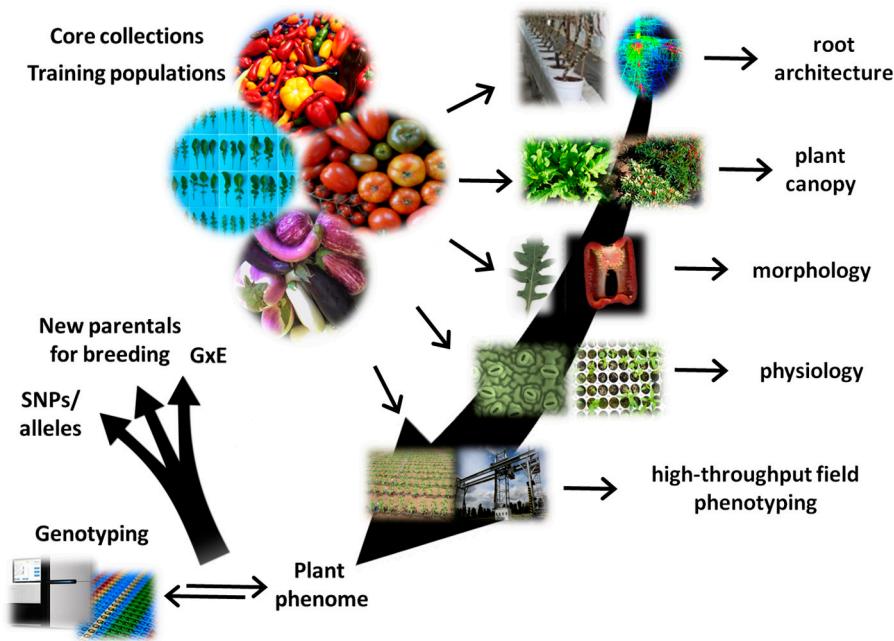
Future research on horticultural crops should aim at advances in shelf life extension and, at the same time, at improving the quality and enhance the health benefits derived from their consumption (e.g., increase the levels of bioactive compounds). Breeders should also consider taste and flavor, while breeding for products with improved nutritional attributes. Additional urgent needs concern the mitigation of abiotic and biotic stress factors to counter climate change related adverse effects on vegetable crop yield, as well as the development of "low input vegetable crops" because of increasing needs to achieve sustainable practices. We think the research community has developed over the last decade a complete toolbox, full of effective items, to face forthcoming challenges in horticulture.

This review provides useful and contemporary information at one place, and supports the notion that efficient management and exploitation of PGR is essential for recovering the repertoire of alleles that has been left behind by the artificial selection process.

With the widespread availability of NGS technologies SNP/allele discovery become feasible and affordable even for species with no genome sequence. Sequence-based information is increasingly used to assess genetic variability within and between closely-related species and restriction enzyme-based enrichment is making diversity studies on PGR more informative than before. Indeed, genetic diversity assessment of horticultural crop collections is a prerequisite for breeding purposes. It first allows

guiding decision making for their conservation. Secondly, it provides novel options (i.e., new/beneficial alleles) to breeders for the greater exploitation of horticultural germplasm aiming at the development of new hybrids and varieties.

On the other hand, advances in phenotyping technologies are essential for capitalizing on the developments in NGS-based genotyping; are accelerating the discovery of trait-allele associations and are improving our ability to uncover genotype-phenotype relationship (Figure 3).



**Figure 3.** Large-scale phenotyping and its impact on plant breeding. Core collection (CC) and/or training populations (TP) developed in various fruit and leafy vegetable crops (in the figure from the top clockwise: pepper, tomato, eggplant, and rocket salad) can be deeply assessed through innovative phenotyping tools for different categories of traits. The integration with genotyping data lead to the identification of: (i) SNPs and alleles associated with target traits; (ii) accessions which can be used as parentals for future breeding programs; and, (iii) the basis of the genotype per environment ( $G \times E$ ) interaction.

The combination of genotyping with accurate phenotyping is defining the current practices in the breeding of vegetable crops. Nonetheless, we have described how the association between genetic variations and a phenotypic change is nothing but simple. Since GWAS enter the scientific scenario, our ability to link genotypes to phenotypes is markedly improved. Thanks to their predictive power, GWAS have experienced rapid growth in the agri-genomics field and are allowing us to assemble some of the pieces of the puzzle that depict how genotypes articulate with phenotypes. Although the puzzle is partially assembled and with many pieces missing, it is already unraveling the genetic bases of complex agronomical traits and it is making available different alleles to be used to cope with forthcoming challenges.

From the review of recent literature it is clear that GWAS in horticultural crops are more common for traits related to plant morphology and quality, while SNP(s)/trait associations for abiotic and biotic stresses are quite neglected. This is explained by the fact that fruit and leafy vegetables are a major source of biologically active compounds with human health-promoting properties.

Likely, it is less time consuming to establish small or large size populations of unrelated individuals with contrasting phenotypes than develop mapping population segregating for traits related to abiotic/biotic stresses. Furthermore, phenotyping horticultural crops for abiotic/biotic stresses is still not within everyone's means and requires *ad-hoc* facilities and field trials across a broad

range of environments. Nevertheless, we expect a change of course in the near future and that GWAS under abiotic/biotic stresses will be undertaken in vegetable crops as in the case for staple crops.

We also believe that it is only a matter of time before GWAS will routinely applied in vegetable crops to identify the so-called non-phenotypic variations, such as eQTL (expression QTL) and/or methylation QTL (meQTL) [117,118].

While the identification of novel and valuable alleles via GWAS is a well-established and quite reliable process, their transfer to elite vegetable crops is still limited. As mentioned above, the increasing spread of NPBT will ease targeted genome editing in vegetable crops facilitating the combination of new sets of alleles.

Finally, we expect genomic selection to cause a paradigm shift in horticultural crop breeding by greatly expediting genomics-driven crop design [119].

The few published papers on GS application in horticultural crops, we briefly illustrated, show how GS is still in its infancy and that its great potential is far from being fully exploited. Indeed, GS could lead to a new scenario within commercial plant breeding programs, facilitating the selection of potential parents among those with the highest GEBV value.

Further efforts in the development of novel statistical models able to capture more faithfully genotype-phenotype relationships as well as software and databases implementation will, however, required for higher precision mapping.

GWAS, GS, and complex mapping populations (i.e., MAGIC) represent nowadays the main pillars for next generation precision breeding. Although less emphasis to research programs in plant breeding is given by universities and research institutes [120], the work of the scientific community in the next decades will be crucial in the development of appropriate populations with the aim to better utilize genetic diversity within vegetable crops.

**Acknowledgments:** This work has received funding from the EU Horizon 2020 research and innovation programme under grant agreement No. 677379 (G2PSOL, Linking genetic resources, genomes and phenotypes of Solanaceous crops).

**Author Contributions:** Nunzio D’Agostino provided the outlines of the review and the key ideas. Pasquale Tripodi contributed to the concept and layout of the manuscript. Nunzio D’Agostino and Pasquale Tripodi contributed to the writing and editing of the review.

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

## Abbreviations

AmpSeq	amplicon sequencing
CBD	convention on biological diversity
CC	core collections
CIR	color infrared
CWR	crop wild relatives
EMS	ethylmethane sulfonate
eQTL	expression QTL
FAO	food and agriculture organization of the United Nations
FBP	field-based phenotyping
FDR	false discovery rate
FIGS	focused identification of germplasm strategy
GBS	genotype-by-sequencing
GEBV	genomic estimated breeding values
GIS	geographic information system
GLM	general linear model
GP	gene pools
GS	genomic selection
GWAS	genome-wide association studies
HS	hyperspectral

HTPP	high-throughput plant phenotyping
IBLs	inbred backcross lines
ILs	introgression lines
LAI	leaf area index
LD	linkage disequilibrium
LR	landraces
MAGIC	multi-parent advanced generation inter-cross
meQTL	methylation QTL
MLM	mixed linear model
MRI	magnetic resonance imagers
MS	multispectral
NDVI	normalized difference vegetation index
NGS	next generation sequencing
NPBT	novel plant breeding techniques
PCR	polymerase chain reaction
PGR	plant genetic resources
QTN	quantitative trait nucleotides
QTL	quantitative trait loci
RAD-Seq	restriction site-associated DNA sequencing
REs	restriction enzymes
RILs	recombinant inbred lines
SMD	sterility mosaic disease
SNPs	single nucleotide polymorphisms
SSR	simple sequence repeat
TP	training population

## References

1. Meyer, R.S.; Purugganan, M.D. Evolution of crop species: Genetics of domestication and diversification. *Nat. Rev. Genet.* **2013**, *14*, 840–852. [[CrossRef](#)] [[PubMed](#)]
2. Gepts, P. Crop domestication as a long-term selection experiment. In *Plant Breeding Reviews*; John Wiley & Sons, Inc.: Oxford, UK, 2010; pp. 1–44.
3. Meyer, R.S.; DuVal, A.E.; Jensen, H.R. Patterns and processes in crop domestication: An historical review and quantitative analysis of 203 global food crops. *New Phytol.* **2012**, *196*, 29–48. [[CrossRef](#)] [[PubMed](#)]
4. Flint-Garcia, S.A. Genetics and consequences of crop domestication. *J. Agric. Food Chem.* **2013**, *61*, 8267–8276. [[CrossRef](#)] [[PubMed](#)]
5. Lin, T.; Zhu, G.; Zhang, J.; Xu, X.; Yu, Q.; Zheng, Z.; Zhang, Z.; Lun, Y.; Li, S.; Wang, X.; et al. Genomic analyses provide insights into the history of tomato breeding. *Nat. Genet.* **2014**, *46*, 1220–1226. [[CrossRef](#)] [[PubMed](#)]
6. Tieman, D.; Zhu, G.; Resende, M.F.; Lin, T.; Nguyen, C.; Bies, D.; Rambla, J.L.; Beltran, K.S.O.; Taylor, M.; Zhang, B.; et al. A chemical genetic roadmap to improved tomato flavor. *Science* **2017**, *355*, 391–394. [[CrossRef](#)] [[PubMed](#)]
7. Massawe, F.; Mayes, S.; Cheng, A. Crop diversity: An unexploited treasure trove for food security. *Trends Plant. Sci.* **2016**, *21*, 365–368. [[CrossRef](#)] [[PubMed](#)]
8. Diamond, J. Evolution, consequences and future of plant and animal domestication. *Nature* **2002**, *418*, 700–707. [[CrossRef](#)] [[PubMed](#)]
9. Gornall, R.J. Recombination systems and plant domestication. *Biol. J. Linn. Soc.* **1983**, *20*, 375–383. [[CrossRef](#)]
10. Brozynska, M.; Furtado, A.; Henry, R.J. Genomics of crop wild relatives: Expanding the gene pool for crop improvement. *Plant Biotechnol. J.* **2016**, *14*, 1070–1085. [[CrossRef](#)] [[PubMed](#)]
11. Roberts, M.J.; Schlenker, W. World supply and demand of food commodity calories. *Am. J. Agric. Econ.* **2009**, *91*, 1235–1242. [[CrossRef](#)]
12. World Population Ageing. Available online: <http://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2013> (accessed on 24 July 2017).

13. Godfray, H.C.J.; Beddington, J.R.; Crute, I.R.; Haddad, L.; Lawrence, D.; Muir, J.F.; Pretty, J.; Robinson, S.; Thomas, S.M.; Toulmin, C. Food security: The challenge of feeding 9 billion people. *Science* **2010**, *327*, 812–818. [[CrossRef](#)] [[PubMed](#)]
14. Khoury, C.K.; Bjorkman, A.D.; Dempewolf, H.; Ramirez-Villegas, J.; Guarino, L.; Jarvis, A.; Rieseberg, L.H.; Struik, P.C. Increasing homogeneity in global food supplies and the implications for food security. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 4001–4006. [[CrossRef](#)] [[PubMed](#)]
15. Tanksley, S.D.; McCouch, S.R. Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science* **1997**, *277*, 1063–1066. [[CrossRef](#)] [[PubMed](#)]
16. Dyer, G.A.; López-Feldman, A.; Yúnez-Naude, A.; Taylor, J.E. Genetic erosion in maize’s center of origin. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 14094–14099. [[CrossRef](#)] [[PubMed](#)]
17. Maxted, N.; Kell, S.P. *Establishment of a Global Network for the In Situ Conservation of Crop Wild Relatives: Status and Needs*; FAO Commission on Genetic Resources for Food and Agriculture: Rome, Italy, 2009; p. 266.
18. Maxted, N.; Ford-Lloyd, B.V.; Hawkes, J.G. Complementary conservation strategies. In *Plant Genetic Conservation, The In Situ Approach*; Maxted, N., Ford-Lloyd, B.V., Hawkes, J.G., Eds.; Springer: Dordrecht, The Netherlands, 1997; pp. 15–39.
19. Gruber, K. Agrobiodiversity: The living library. *Nature* **2017**, *544*, S8–S10. [[CrossRef](#)] [[PubMed](#)]
20. Koo, B.; Wright, B.D. The optimal timing of evaluation of genebank accessions and the effects of biotechnology. *Am. J. Agric. Econ.* **2000**, *82*, 797–811. [[CrossRef](#)]
21. Brown, A.H.D. Core collections: A practical approach to genetic resources management. *Genome* **1989**, *31*, 818–824. [[CrossRef](#)]
22. McKhann, H.I.; Camilleri, C.; Bérard, A.; Bataillon, T.; David, J.L.; Reboud, X.; Le Corre, V.; Caloustian, C.; Gut, I.G.; Brunel, D. Nested core collections maximizing genetic diversity in arabidopsis thaliana. *Plant. J.* **2004**, *38*, 193–202. [[CrossRef](#)] [[PubMed](#)]
23. Kim, K.-W.; Chung, H.-K.; Cho, G.-T.; Ma, K.-H.; Chandrabalan, D.; Gwag, J.-G.; Kim, T.-S.; Cho, E.-G.; Park, Y.-J. Powercore: A program applying the advanced m strategy with a heuristic search for establishing core sets. *Bioinformatics* **2007**, *23*, 2155–2162. [[CrossRef](#)] [[PubMed](#)]
24. Borrayo, E.; Takeya, M. Signal-Processing Tools for Core-Collection Selection from Genetic-Resource Collections. Available online: <https://f1000research.com/articles/4-97/v1> (accessed on 14 September 2017).
25. Khazaei, H.; Street, K.; Bari, A.; Mackay, M.; Stoddard, F.L. The figs (focused identification of germplasm strategy) approach identifies traits related to drought adaptation in vicia faba genetic resources. *PLoS ONE* **2013**, *8*, e63107. [[CrossRef](#)] [[PubMed](#)]
26. Street, K.; Bari, A.; Mackay, M.; Amri, A. How the focused identification of germplasm strategy (figs) is used to mine plant genetic resources collections for adaptive traits. In *Enhancing Crop Genepool Use: Capturing Wild Relative and Landrace Diversity for Crop Improvement*; Maxted, N., Dulloo, M.E., Ford-Lloyd, B.V., Eds.; CABI: Boston, MA, USA, 2016; p. 54.
27. Odong, T.L.; Jansen, J.; van Eeuwijk, F.A.; van Hintum, T.J.L. Quality of core collections for effective utilisation of genetic resources review, discussion and interpretation. *Theor. Appl. Genet.* **2013**, *126*, 289–305. [[CrossRef](#)] [[PubMed](#)]
28. Fao: Second Global Plan of Action for Plant Genetic Resources. Available online: <http://www.fao.org/agriculture/crops/thematic-sitemap/theme/seeds-pgr/gpa/en/> (accessed on 24 July 2017).
29. Divseek. Available online: <http://www.divseek.org> (accessed on 21 July 2017).
30. G2p-sol: Linking Genetic Resources, Genomes and Phenotypes of Solanaceous Crops. Available online: <http://www.g2p-sol.eu/> (accessed on 21 January 2017).
31. Traditom: Traditional Tomato Varieties and Cultural Practices. Available online: <http://traditom.eu/> (accessed on 21 July 2017).
32. Hajjar, R.; Hodgkin, T. The use of wild relatives in crop improvement: A survey of developments over the last 20 years. *Euphytica* **2007**, *156*, 1–13. [[CrossRef](#)]
33. Harlan, J.R.; de Wet, J.M.J. Toward a rational classification of cultivated plants. *Taxon* **1971**, *20*, 509–517. [[CrossRef](#)]
34. Gepts, P.; Papa, R. Possible effects of (trans)gene flow from crops on the genetic diversity from landraces and wild relatives. *Environ. Biosaf. Res.* **2003**, *2*, 89–103. [[CrossRef](#)]
35. Cardi, T. Cisgenesis and genome editing: Combining concepts and efforts for a smarter use of genetic resources in crop breeding. *Plant Breed.* **2016**, *135*, 139–147. [[CrossRef](#)]

36. Cardi, T.; D’Agostino, N.; Tripodi, P. Genetic transformation and genomic resources for next-generation precise genome engineering in vegetable crops. *Front. Plant. Sci.* **2017**, *8*, 241. [[CrossRef](#)] [[PubMed](#)]
37. Zamir, D. Improving plant breeding with exotic genetic libraries. *Nat. Rev. Genet.* **2001**, *2*, 983–989. [[CrossRef](#)] [[PubMed](#)]
38. Prohens, J.; Gramazio, P.; Plazas, M.; Dempewolf, H.; Kilian, B.; Díez, M.J.; Fita, A.; Herraiz, F.J.; Rodríguez-Burrueto, A.; Soler, S.; et al. Introgressomics: A new approach for using crop wild relatives in breeding for adaptation to climate change. *Euphytica* **2017**, *213*, 158. [[CrossRef](#)]
39. Holland, J.B. Genetic architecture of complex traits in plants. *Curr. Opin. Plant Biol.* **2007**, *10*, 156–161. [[CrossRef](#)] [[PubMed](#)]
40. Zhu, C.; Gore, M.; Buckler, E.S.; Yu, J. Status and prospects of association mapping in plants. *Plant Genome* **2008**, *1*, 5–20. [[CrossRef](#)]
41. Korte, A.; Farlow, A. The advantages and limitations of trait analysis with gwas: A review. *Plant Methods* **2013**, *9*, 29. [[CrossRef](#)] [[PubMed](#)]
42. Huang, B.E.; Verbyla, K.L.; Verbyla, A.P.; Raghavan, C.; Singh, V.K.; Gaur, P.; Leung, H.; Varshney, R.K.; Cavanagh, C.R. Magic populations in crops: Current status and future prospects. *Theor. Appl. Genet.* **2015**, *128*, 999–1017. [[CrossRef](#)] [[PubMed](#)]
43. King, E.G.; Merkes, C.M.; McNeil, C.L.; Hoofer, S.R.; Sen, S.; Broman, K.W.; Long, A.D.; Macdonald, S.J. Genetic dissection of a model complex trait using the drosophila synthetic population resource. *Genome Res.* **2012**, *22*, 1558–1566. [[CrossRef](#)] [[PubMed](#)]
44. Pascual, L.; Desplat, N.; Huang, B.E.; Desgroux, A.; Bruguier, L.; Bouchet, J.-P.; Le, Q.H.; Chauchard, B.; Verschave, P.; Causse, M. Potential of a tomato magic population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. *Plant Biotechnol. J.* **2015**, *13*, 565–577. [[CrossRef](#)] [[PubMed](#)]
45. Alseekh, S.; Ofner, I.; Pleban, T.; Tripodi, P.; Di Dato, F.; Cammareri, M.; Mohammad, A.; Grandillo, S.; Fernie, A.R.; Zamir, D. Resolution by recombination: Breaking up solanum pennellii introgressions. *Trends Plant Sci.* **2013**, *18*, 536–538. [[CrossRef](#)] [[PubMed](#)]
46. Henry, R.J. *Plant Genotyping: The DNA Fingerprinting of Plants*; CABI Publishing: Oxford, UK, 2001.
47. Mammadov, J.; Aggarwal, R.; Buyyapu, R.; Kumpatla, S. Snp markers and their impact on plant breeding. *Int. J. Plant Genom.* **2012**, *2012*, 728398. [[CrossRef](#)] [[PubMed](#)]
48. Smith, S.M.; Maughan, P.J. Snp genotyping using kaspar assays. In *Plant genotyping: Methods and Protocols*; Batley, J., Ed.; Springer: New York, NY, USA, 2015; pp. 243–256.
49. Shen, G.-Q.; Abdullah, K.G.; Wang, Q.K. The taqman method for.snp genotyping. In *Single Nucleotide Polymorphisms: Methods and Protocols*; Komar, A.A., Ed.; Humana Press: Totowa, NJ, USA, 2009; pp. 293–306.
50. LaFramboise, T. Single nucleotide polymorphism arrays: A decade of biological, computational and technological advances. *Nucleic Acids Res.* **2009**, *37*, 4181–4193. [[CrossRef](#)] [[PubMed](#)]
51. Sim, S.C.; Durstewitz, G.; Plieske, J.; Wieseke, R.; Ganal, M.W.; Van Deynze, A.; Hamilton, J.P.; Buell, C.R.; Causse, M.; Wijeratne, S.; et al. Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. *PLoS ONE* **2012**, *7*, e40563. [[CrossRef](#)] [[PubMed](#)]
52. Jaccoud, D.; Peng, K.; Feinstein, D.; Kilian, A. Diversity arrays: A solid state technology for sequence information independent genotyping. *Nucleic Acids Res.* **2001**, *29*, E25. [[CrossRef](#)] [[PubMed](#)]
53. Davey, J.W.; Hohenlohe, P.A.; Etter, P.D.; Boone, J.Q.; Catchen, J.M.; Blaxter, M.L. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* **2011**, *12*, 499–510. [[CrossRef](#)] [[PubMed](#)]
54. Ray, S.; Satya, P. Next generation sequencing technologies for next generation plant breeding. *Front Plant Sci.* **2014**, *5*, 367. [[CrossRef](#)] [[PubMed](#)]
55. Aflitos, S.; Schijlen, E.; Jong, H.; Ridder, D.; Smit, S.; Finkers, R.; Wang, J.; Zhang, G.; Li, N.; Mao, L.; et al. Exploring genetic variation in the tomato (solanum section lycopersicon) clade by whole-genome sequencing. *Plant. J.* **2014**, *80*, 136–148. [[PubMed](#)]
56. Schmutz, J.; McClean, P.E.; Mamidi, S.; Wu, G.A.; Cannon, S.B.; Grimwood, J.; Jenkins, J.; Shu, S.; Song, Q.; Chavarro, C.; et al. A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* **2014**, *46*, 707–713. [[CrossRef](#)] [[PubMed](#)]

57. Guo, S.; Zhang, J.; Sun, H.; Salse, J.; Lucas, W.J.; Zhang, H.; Zheng, Y.; Mao, L.; Ren, Y.; Wang, Z.; et al. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nat. Genet.* **2013**, *45*, 51–58. [CrossRef] [PubMed]
58. Causse, M.; Desplat, N.; Pascual, L.; Le Paslier, M.C.; Sauvage, C.; Bauchet, G.; Berard, A.; Bounon, R.; Tchoumakov, M.; Brunel, D.; et al. Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. *BMC Genom.* **2013**, *14*, 791. [CrossRef] [PubMed]
59. Terracciano, I.; Cantarella, C.; D’Agostino, N. Hybridization-based enrichment and next generation sequencing to explore genetic diversity in plants. In *Dynamics of Mathematical Models in Biology: Bringing Mathematics to Life*; Rogato, A., Zazzu, V., Guaracino, M., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 117–136.
60. Terracciano, I.; Cantarella, C.; Fasano, C.; Cardi, T.; Mennella, G.; D’Agostino, N. Liquid-phase sequence capture and targeted re-sequencing revealed novel polymorphisms in tomato genes belonging to the mep carotenoid pathway. *Sci. Rep.* **2017**, *7*, 5616. [CrossRef] [PubMed]
61. Ruggieri, V.; Anzar, I.; Paytuvi, A.; Calafiole, R.; Ciglano, R.A.; Sanseverino, W.; Barone, A. Exploiting the great potential of sequence capture data by a new tool, super-cap. *DNA Res.* **2017**, *24*, 81–91. [CrossRef] [PubMed]
62. Cronn, R.; Knaus, B.J.; Liston, A.; Maughan, P.J.; Parks, M.; Syring, J.V.; Udall, J. Targeted enrichment strategies for next-generation plant biology. *Am. J. Bot.* **2012**, *99*, 291–311. [CrossRef] [PubMed]
63. Jiang, Z.; Wang, H.; Michal, J.J.; Zhou, X.; Liu, B.; Woods, L.C.; Fuchs, R.A. Genome wide sampling sequencing for snp genotyping: Methods, challenges and future development. *Int. J. Biol. Sci.* **2016**, *12*, 100–108. [CrossRef] [PubMed]
64. Baird, N.A.; Etter, P.D.; Atwood, T.S.; Currey, M.C.; Shiver, A.L.; Lewis, Z.A.; Selker, E.U.; Cresko, W.A.; Johnson, E.A. Rapid snp discovery and genetic mapping using sequenced rad markers. *PLoS ONE* **2008**, *3*, e3376. [CrossRef] [PubMed]
65. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A robust, simple genotyping-by-sequencing (gbs) approach for high diversity species. *PLoS ONE* **2011**, *6*, e19379. [CrossRef] [PubMed]
66. Voss-Fels, K.; Snowdon, R.J. Understanding and utilizing crop genome diversity via high-resolution genotyping. *Plant. Biotechnol. J.* **2016**, *14*, 1086–1094. [CrossRef] [PubMed]
67. Taranto, F.; D’Agostino, N.; Tripodi, P. An overview of genotyping by sequencing in crop species and its application in pepper. In *Dynamics of Mathematical Models in Biology: Bringing Mathematics to Life*; Rogato, A., Zazzu, V., Guaracino, M., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 101–116.
68. Boutet, G.; Alves Carvalho, S.; Falque, M.; Peterlongo, P.; Lhuillier, E.; Bouchez, O.; Lavaud, C.; Pilet-Nayel, M.-L.; Rivière, N.; Baranger, A. Snp discovery and genetic mapping using genotyping by sequencing of whole genome genomic DNA from a pea ril population. *BMC Genom.* **2016**, *17*, 121. [CrossRef] [PubMed]
69. Verma, S.; Gupta, S.; Bandhiwal, N.; Kumar, T.; Bharadwaj, C.; Bhatia, S. High-density linkage map construction and mapping of seed trait qtls in chickpea (*Cicer arietinum* L.) using genotyping-by-sequencing (gbs). *Sci. Rep.* **2015**, *5*, 17512. [CrossRef] [PubMed]
70. Ren, Y.; McGregor, C.; Zhang, Y.; Gong, G.; Zhang, H.; Guo, S.; Sun, H.; Cai, W.; Zhang, J.; Xu, Y. An integrated genetic map based on four mapping populations and quantitative trait loci associated with economically important traits in watermelon (*Citrullus lanatus*). *BMC Plant. Biol.* **2014**, *14*, 33. [CrossRef] [PubMed]
71. Celik, I.; Gurbuz, N.; Uncu, A.T.; Frary, A.; Doganlar, S. Genome-wide snp discovery and qtl mapping for fruit quality traits in inbred backcross lines (ibls) of solanum pimpinellifolium using genotyping by sequencing. *BMC Genom.* **2017**, *18*, 1. [CrossRef] [PubMed]
72. Sidhu, G.; Mohan, A.; Zheng, P.; Dhaliwal, A.K.; Main, D.; Gill, K.S. Sequencing-based high throughput mutation detection in bread wheat. *BMC Genom.* **2015**, *16*, 962. [CrossRef] [PubMed]
73. Mishra, A.; Singh, A.; Sharma, M.; Kumar, P.; Roy, J. Development of ems-induced mutation population for amylose and resistant starch variation in bread wheat (*Triticum aestivum*) and identification of candidate genes responsible for amylose variation. *BMC Plant Biol.* **2016**, *16*, 217. [CrossRef] [PubMed]
74. Taranto, F.; D’Agostino, N.; Greco, B.; Cardi, T.; Tripodi, P. Genome-wide snp discovery and population structure analysis in pepper (*Capsicum annuum*) using genotyping by sequencing. *BMC Genom.* **2016**, *17*, 943. [CrossRef] [PubMed]

75. Pavan, S.; Lotti, C.; Marcotrigiano, A.R.; Mazzeo, R.; Bardaro, N.; Bracuto, V.; Ricciardi, F.; Taranto, F.; D’Agostino, N.; Schiavulli, A.; et al. A distinct genetic cluster in cultivated chickpea as revealed by genome-wide marker discovery and genotyping. *Plant. Genome* **2017**, *10*. [CrossRef] [PubMed]
76. 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]
77. Schröder, S. Optimization of genotyping by sequencing (gbs) data in common bean (*Phaseolus vulgaris* L.). *Mol. Breed.* **2016**, *36*, 6. [CrossRef]
78. Torkamaneh, D.; Laroche, J.; Belzile, F. Genome-wide SNP calling from genotyping by sequencing (gbs) data: A comparison of seven pipelines and two sequencing technologies. *PLoS ONE* **2016**, *11*, e0161333. [CrossRef] [PubMed]
79. Catchen, J.; Hohenlohe, P.A.; Bassham, S.; Amores, A.; Cresko, W.A. Stacks: An analysis tool set for population genomics. *Mol. Ecol.* **2013**, *22*, 3124–3140. [CrossRef] [PubMed]
80. Lu, F.; Lipka, A.E.; Glaubitz, J.; Elshire, R.; Cherney, J.H.; Casler, M.D.; Buckler, E.S.; Costich, D.E. Switchgrass genomic diversity, ploidy, and evolution: Novel insights from a network-based SNP discovery protocol. *PLoS Genet.* **2013**, *9*, e1003215. [CrossRef] [PubMed]
81. Russell, J.; Hackett, C.; Hedley, P.; Liu, H.; Milne, L.; Bayer, M.; Marshall, D.; Jorgensen, L.; Gordon, S.; Brennan, R. The use of genotyping by sequencing in blackcurrant (*Ribes nigrum*): Developing high-resolution linkage maps in species without reference genome sequences. *Mol. Breed.* **2014**, *33*, 835–849. [CrossRef]
82. Huang, Y.F.; Poland, J.A.; Wight, C.P.; Jackson, E.W.; Tinker, N.A. Using genotyping-by-sequencing (gbs) for genomic discovery in cultivated oat. *PLoS ONE* **2014**, *9*, e102448. [CrossRef] [PubMed]
83. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [PubMed]
84. Alexander, D.H.; Novembre, J.; Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **2009**, *19*, 1655–1664. [CrossRef] [PubMed]
85. Awclust. Available online: <https://sourceforge.net/projects/awclust/> (accessed on 20 July 2017).
86. Yang, S.; Fresnedo-Ramirez, J.; Wang, M.; Cote, L.; Schweitzer, P.; Barba, P.; Takacs, E.M.; Clark, M.; Luby, J.; Manns, D.C.; et al. A next-generation marker genotyping platform (ampseq) in heterozygous crops: A case study for marker-assisted selection in grapevine. *Hortic. Res.* **2016**, *3*, 16002. [CrossRef] [PubMed]
87. Buckler, E.S.; Ilut, D.C.; Wang, X.; Kretzschmar, T.; Gore, M.A.; Mitchell, S.E. Rampseq: Using repetitive sequences for robust genotyping. *bioRxiv* **2016**. [CrossRef]
88. Furbank, R.T.; Tester, M. Phenomics—Technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* **2011**, *16*, 635–644. [CrossRef] [PubMed]
89. Araus, J.L.; Cairns, J.E. Field high-throughput phenotyping: The new crop breeding frontier. *Trends Plant Sci.* **2014**, *19*, 52–61. [CrossRef] [PubMed]
90. Li, L.; Zhang, Q.; Huang, D. A review of imaging techniques for plant phenotyping. *Sensors* **2014**, *14*, 20078–20111. [CrossRef] [PubMed]
91. Fritsche-Neto, R.; Borém, A. *Phenomics: How Next-Generation Phenotyping Is Revolutionizing Plant Breeding*; Springer: Dordrecht, Switzerland, 2015; pp. 1–142.
92. Liew, O.; Chong, P.; Li, B.; Asundi, A. Signature optical cues: Emerging technologies for monitoring plant health. *Sensors* **2008**, *8*, 3205–3239. [CrossRef] [PubMed]
93. Mishra, K.B.; Iannacone, R.; Petrozza, A.; Mishra, A.; Armentano, N.; La Vecchia, G.; Trtílek, M.; Cellini, F.; Nedbal, L. Engineered drought tolerance in tomato plants is reflected in chlorophyll fluorescence emission. *Plant Sci.* **2012**, *182*, 79–86. [CrossRef] [PubMed]
94. Lootens, P.; Devacht, S.; Baert, J.; van Waes, J.; van Bockstaele, E.; Roldán-Ruiz, I. Evaluation of cold stress of young industrial chicory (*Cichorium intybus* L.) by chlorophyll a fluorescence imaging. II. Dark relaxation kinetics. *Photosynthetica* **2011**, *49*, 185–194. [CrossRef]
95. Van der Heijden, G.; Song, Y.; Horgan, G.; Polder, G.; Dieleman, A.; Bink, M.; Palloix, A.; van Eeuwijk, F.; Glasbey, C. Spicy: Towards automated phenotyping of large pepper plants in the greenhouse. *Funct. Plant Biol.* **2012**, *39*, 870–877. [CrossRef]
96. Rascher, U.; Blossfeld, S.; Fiorani, F.; Jahnke, S.; Jansen, M.; Kuhn, A.J.; Matsubara, S.; Märting, L.L.A.; Merchant, A.; Metzner, R.; et al. Non-invasive approaches for phenotyping of enhanced performance traits in bean. *Funct. Plant Biol.* **2011**, *38*, 968–983. [CrossRef]

97. Pauli, D.; Chapman, S.C.; Bart, R.; Topp, C.N.; Lawrence-Dill, C.J.; Poland, J.; Gore, M.A. The quest for understanding phenotypic variation via integrated approaches in the field environment. *Plant Physiol.* **2016**, *172*, 622–634. [[CrossRef](#)] [[PubMed](#)]
98. White, J.W.; Andrade-Sanchez, P.; Gore, M.A.; Bronson, K.F.; Coffelt, T.A.; Conley, M.M.; Feldmann, K.A.; French, A.N.; Heun, J.T.; Hunsaker, D.J.; et al. Field-based phenomics for plant genetics research. *Field Crop. Res.* **2012**, *133*, 101–112. [[CrossRef](#)]
99. Hirschhorn, J.N.; Daly, M.J. Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* **2005**, *6*, 95–108. [[CrossRef](#)] [[PubMed](#)]
100. Brachi, B.; Morris, G.P.; Borevitz, J.O. Genome-wide association studies in plants: The missing heritability is in the field. *Genome Biol.* **2011**, *12*, 232. [[CrossRef](#)] [[PubMed](#)]
101. Huang, X.; Han, B. Natural variations and genome-wide association studies in crop plants. *Annu. Rev. Plant Biol.* **2014**, *65*, 531–551. [[CrossRef](#)] [[PubMed](#)]
102. Zhang, J.; Zhao, J.; Xu, Y.; Liang, J.; Chang, P.; Yan, F.; Li, M.; Liang, Y.; Zou, Z. Genome-wide association mapping for tomato volatiles positively contributing to tomato flavor. *Front. Plant Sci.* **2015**, *6*, 1042. [[CrossRef](#)] [[PubMed](#)]
103. Nimmakayala, P.; Abburi, V.L.; Saminathan, T.; Alaparthi, S.B.; Almeida, A.; Davenport, B.; Nadimi, M.; Davidson, J.; Tonapi, K.; Yadav, L.; et al. Genome-wide diversity and association mapping for capsaicinoids and fruit weight in capsicum annuum L. *Sci. Rep.* **2016**, *6*, 38081. [[CrossRef](#)] [[PubMed](#)]
104. Pavan, S.; Marcotrigiano, A.R.; Ciani, E.; Mazzeo, R.; Zonno, V.; Ruggieri, V.; Lotti, C.; Ricciardi, L. Genotyping-by-sequencing of a melon (*Cucumis melo* L.) germplasm collection from a secondary center of diversity highlights patterns of genetic variation and genomic features of different gene pools. *BMC Genom.* **2017**, *18*, 59. [[CrossRef](#)] [[PubMed](#)]
105. Nimmakayala, P.; Tomason, Y.R.; Abburi, V.L.; Alvarado, A.; Saminathan, T.; Vajja, V.G.; Salazar, G.; Panicker, G.K.; Levi, A.; Wechter, W.P.; et al. Genome-wide differentiation of various melon horticultural groups for use in gwas for fruit firmness and construction of a high resolution genetic map. *Front. Plant Sci.* **2016**, *7*, 1437. [[CrossRef](#)] [[PubMed](#)]
106. Cericola, F.; Portis, E.; Lanteri, S.; Toppino, L.; Barchi, L.; Acciarri, N.; Pulcini, L.; Sala, T.; Rotino, G.L. Linkage disequilibrium and genome-wide association analysis for anthocyanin pigmentation and fruit color in eggplant. *BMC Genom.* **2014**, *15*, 896. [[CrossRef](#)] [[PubMed](#)]
107. Hart, J.P.; Griffiths, P.D. Genotyping-by-sequencing enabled mapping and marker development for the by-2 potyvirus resistance allele in common bean. *Plant Genome* **2015**, *8*. [[CrossRef](#)]
108. Saxena, R.K.; Kale, S.M.; Kumar, V.; Parupali, S.; Joshi, S.; Singh, V.; Garg, V.; Das, R.R.; Sharma, M.; Yamini, K.N.; et al. Genotyping-by-sequencing of three mapping populations for identification of candidate genomic regions for resistance to sterility mosaic disease in pigeonpea. *Sci. Rep.* **2017**, *7*, 1813. [[CrossRef](#)] [[PubMed](#)]
109. Bush, W.S.; Moore, J.H. Chapter 11: Genome-wide association studies. *PLoS Comput. Biol.* **2012**, *8*, e1002822. [[CrossRef](#)] [[PubMed](#)]
110. Zhang, Z.; Ersoz, E.; Lai, C.Q.; Todhunter, R.J.; Tiwari, H.K.; Gore, M.A.; Bradbury, P.J.; Yu, J.; Arnett, D.K.; Ordovas, J.M.; et al. Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* **2010**, *42*, 355–360. [[CrossRef](#)] [[PubMed](#)]
111. Bhat, J.A.; Ali, S.; Salgotra, R.K.; Mir, Z.A.; Dutta, S.; Jadon, V.; Tyagi, A.; Mushtaq, M.; Jain, N.; Singh, P.K.; et al. Genomic selection in the era of next generation sequencing for complex traits in plant breeding. *Front. Genet.* **2016**, *7*, 221. [[CrossRef](#)] [[PubMed](#)]
112. Hamblin, M.T.; Buckler, E.S.; Jannink, J.-L. Population genetics of genomics-based crop improvement methods. *Trends Genet.* **2011**, *27*, 98–106. [[CrossRef](#)] [[PubMed](#)]
113. Xu, Y.; Xie, C.; Wan, J.; He, Z.; Prasanna, B.M. Marker-assisted selection in cereals: Platforms, strategies and examples. In *Cereal Genomics II*; Gupta, P.K., Varshney, R.K., Eds.; Springer: Dordrecht, The Netherlands, 2013; pp. 375–411.
114. 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](#)] [[PubMed](#)]
115. Duangjitt, J.; Causse, M.; Sauvage, C. Efficiency of genomic selection for tomato fruit quality. *Mol. Breed.* **2016**, *36*, 29. [[CrossRef](#)]

116. 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](#)]
117. Druka, A.; Potokina, E.; Luo, Z.; Jiang, N.; Chen, X.; Kearsey, M.; Waugh, R. Expression quantitative trait loci analysis in plants. *Plant Biotechnol. J.* **2010**, *8*, 10–27. [[CrossRef](#)] [[PubMed](#)]
118. Vidalis, A.; Živković, D.; Wardenaar, R.; Roquis, D.; Tellier, A.; Johannes, F. Methylome evolution in plants. *Genome Biol.* **2016**, *17*, 264. [[CrossRef](#)] [[PubMed](#)]
119. Desta, Z.A.; Ortiz, R. Genomic selection: Genome-wide prediction in plant improvement. *Trends Plant Sci* **2014**, *19*, 592–601. [[CrossRef](#)] [[PubMed](#)]
120. Cheng, Z.-M.; Werner, D.J. Will the Traditional Horticultural Breeding and Genetics Research Be Fairly Valued in Academia. Available online: <https://www.nature.com/articles/hortres201553> (accessed on 14 September 2017).



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).