# Research advances and prospects of spinach breeding, genetics, and genomics

Gehendra Bhattarai* and Ainong Shi*

*Department of Horticulture, University of Arkansas, Fayetteville, AR 72701, USA*
* Corresponding author, E-mail: gb005@uark.edu; ashi@uark.edu

## Abstract

Spinach (*Spinacia oleracea*) is a diploid (2n = 2x = 12), wind-pollinated and highly heterozygous crop. The plants are mostly dioecious, although some monoecious plants exist. Spinach is an economically important cool-season leafy vegetable crop. Demand for spinach is increasing worldwide, particularly due to its high nutritional content. Spinach is a versatile crop eaten raw or cooked and used as salads or mixed with other cuisines. This review article provides an overview of origin and domestication, genetic diversity and population structure, genetic and genomic resources, major diseases threatening spinach production, breeding progress, and synthesizing how these resources can help in spinach improvement. The rapid development of genomic and sequence resources of spinach has increased biological and genetics research and laid the foundation for adopting molecular breeding. Downy mildew is the most serious disease affecting spinach and breeding programs focus on developing cultivars resistant to continually emerging new races of downy mildew pathogens. The use of genomic and molecular resources and approaches offers promises in population improvement and hybrid development to address biotic and abiotic stresses production challenges and provide improved breeding materials and strategies against the rapidly changing pathogen races and climatic conditions.

## Introduction

Spinach (*Spinacia oleracea* L.) is a diploid, dioecious, wind-pollinated, and highly heterozygous crop. Spinach is one of the important leafy vegetable crops worldwide. China, the United States, Turkey, and Japan are the major countries producing spinach (Fig. 1a), with the annual worldwide production of spinach being in the region of 30.1 million tonnes, of which China produces 27.52 million tonnes comprising 91% of world production[1]. The United States produces around 0.44 million tonnes of spinach, of which almost 90% is produced in California and Arizona, with the majority of the remaining crop being produced in New Jersey and Texas[2]. Mild-cool temperatures favor spinach growth and the crop is grown year-round within the northern California and Arizona valleys in the United States. Over the past two decades (2000−2019), worldwide spinach production has been continuously increasing (Fig. 1b). The growing interest in low-calorie diets and increased understanding of nutritional and health benefits of spinach have increased its demand over the past few decades and is expected to continue to increase in the coming years.

Spinach is excellent in nutrition with a high amount of vitamins, proteins, minerals, flavonoids and is low in calories[3,4]. The leafy green spinach is rich in iron, lutein, folate, vitamins, minerals, antioxidants[5], and carotenoid and phenolic compounds[4,6,7]. Regular consumption of spinach benefits human health as the high antioxidant content in spinach leaves helps cells against oxidative stress with free-radical

species[8]. Some other possible benefits of spinach consumption include blood glucose control for diabetics, lower cancer risk, and supply of minerals and vitamins.

In recent years, advanced genomic resources have been developed, including the reference genome assembly, transcriptome sequences, and genome variant data for the spinach germplasm panel. However, the genetic advancement of spinach lags behind other important horticultural crops such as tomato and lettuce. An article in 2011 reviewed the genetic and genomic resources available in spinach, emphasizing two primary spinach diseases: downy mildew and white rust[9]. Recently, the genetic resources, domestication, and breeding history of spinach was reviewed[10]. This article summarizes available genetic resources and technological advances made in spinach, highlights the current status and progress on genomics and molecular resources, and outlines how these resources may help in improving spinach.

## Botany

Spinach (*S. oleracea*) belongs to the Amaranthaceae family in the Caryophyllales order. Some other economically and nutritionally important species in this family are quinoa (*Chenopodium quinoa* Willd.), amaranth (*Amaranthus* spp. L.), and sugar beet (*Beta vulgaris* L.)[11−13]. Spinach is a diploid crop (2n = 2x = 12) with an estimated genome size of around 989 Mb[14].

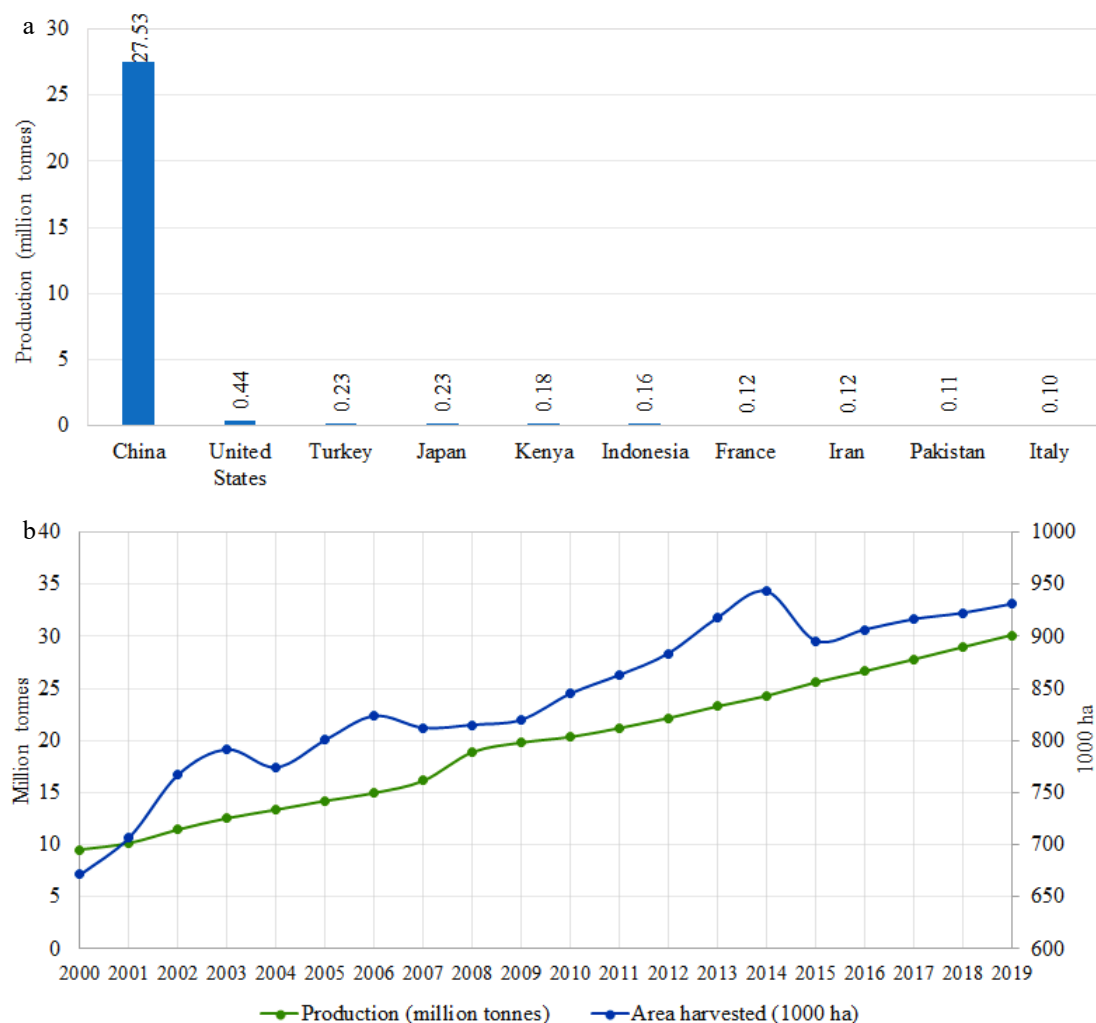Spinach is an annual plant with vegetative and reproductive phases. It is photoperiod sensitive and bolts with long

**Fig. 1** Worldwide spinach production. (a) The top ten spinach producing countries. (b) Global spinach production and harvest area in the past two decades.

daytime and high temperatures during summer[15]. For seed production, spinach is sown from beginning to mid-winter — plants transit from the vegetative to the reproductive stage with the arrival of long spring or summer days. Commercial spinach seeds are thus produced in regions with mild temperatures and short days during planting and long-warmer days to induce bolting during the reproductive phase. Early bolting is undesirable for many leafy vegetables, including spinach, as the leaves develop a bitter taste and are often not marketable[16]. New spinach cultivars are slow bolting and are available to adapt to wide photoperiods and climatic conditions.

Spinach is a predominantly dioecious crop and is wind-pollinated (Fig. 2). Dioecious spinach shows separate male and female plants with an equal sex ratio[17], although some genotypes produce monoecious plants[18]. Dioecious spinach has the X/Y sex determination system with heterogametic males (XY) and homogametic females (XX)[18,19]. The monoecious locus (M/m) was linked to the X/Y locus[20]. The sex-expressing genes have not been isolated in spinach and the mechanism of sex determination is unclear. Candidate genes linked to the male-specific region[21] and single nucleotide



**Fig. 2** Female (left) and male (right) spinach plants. Spinach plants are kept in isolated crossing blocks to generate seeds (center).

polymorphism (SNP) marker co-segregating with the sex-determining gene were reported[22]. Sex expression in spinach is flexible and shows sex reversion[15,23]. Environmental conditions and exogenous growth regulators modulate sex expression in dioecious spinach, and lack of pollen and environmental stress can induce sex reversion allowing

selfing in spinach. High air temperature and short days favor femaleness in spinach[24,25]. Also, exogenous application of gibberellic acid (GA3) increases male plants and abscisic acid (ABA) induces femaleness[26].

A morphological variation categorizes spinach as savoy, semi-savoy, and smooth or flat based on the leaf shape and structure. Savoy cultivars have crinkles or wrinkles on leaves, semi-savoy have reduced wrinkles, and flat or smooth cultivars lack wrinkles. Savoy leaf type spinach is dark green in color, taste better, and are preferred by home gardeners. Savoy-leafed spinach is popular in fresh markets, favored for long-range market transportation due to extended shelf life, and is also important for canned and processed products[27], although the savoy types are difficult to clean and are primarily used in cooking. On the other hand, the smooth spinach types are the most popular as they are easy to clean and package and thus are sold pre-washed and packaged for fresh-cut markets. Baby leaf spinach is primarily smooth leaf type and is harvested during early growth stages when leaves are small, tender, and sweet.

## Origin and domestication

Knowledge of the origin of crops and their wild progenitor species aids plant biologists and breeders in utilizing wild germplasm in crop improvement. Spinach is considered native to central and southwest Asia[27], but is now cultivated worldwide. The wild spinach *S. turkestanica* iljin and *S. tetrandra* Stev. are distributed in West Asia and are presumed to have been originated and domesticated in former Persia (present Iran) and later migrated towards Eastern Asia, North Africa, and Western Europe, and then to North America[28,29]. However, the exact origin of cultivated spinach is still unknown, and new pieces of information are generated on migration routes. Prior literature on the migration and spread of spinach to Southern and Eastern Asia, the Mediterranean, and Europe was presumed based on historical evidence and the predominance of diverse landraces and wild species in those regions. The two main wild *Spinacia* species are distributed in Western Asia over the two major geographical areas: the *S. turkestanica* in Central Asia in Turkmenistan, Uzbekistan, Kazakhstan, Tajikistan and in South Asia in Afghanistan and Pakistan and the *S. tetrandra* in the Middle East and the Caucasus regions in Armenia, Georgia, Iran, Iraq, and Turkey. Thus, the geographical distribution of wild spinach species and the diversity of cultivated spinach in Western Asia support their presumed origin in the region.

Domestication involves selecting preferred phenotypes over generations, leading to the loss of diversity found in the progenitors. Linguistic and historical records indicated spinach consumption during the 4th century AD in Mesopotamia, the 7th century in China, and later in the 10th century in the Mediterranean region, as discussed in Ribera et al. (2020)[10]. The Asian and Western cultivars selected in different geographical areas can be differentiated based on morphology and phylogenetic analysis[10,30,31]. On the other hand, both wild species *S. turkestanica* and *S. tetrandra,* are morphologically similar in appearance to the cultivated *S. oleracea* except for the inflorescences[10]. All three *Spinacia* species are diploid with six chromosomes, and the two wild species are interfertile with cultivated *S. oleracea*. Narrow geographical distribution of these wild species and their sexual compatibility with cultivated *S. oleracea* suggests spinach may have originated through the domestication of one or both wild species[28]. The two wild species are assigned to the primary gene pool of cultivated spinach, although a recent report found reduced fertility in the hybrid offsprings of *S. oleracea* and *S. tetrandra*[32], potentially preventing gene flow between species. The cause of such reduced fertility has not been fully investigated, but possibly was not with the differences in sex chromosomes and their rearrangements, while the apparent expectations are due to the genetic incompatibilities leading to pollen abortion[32].

Genetic analyses of transcriptome variants of 107 cultivated *S. oleracea* and 13 wild accessions (5 *S. tetrandra* and 8 *S. turkestanica)* found less nucleotide divergence in *S. oleracea* ($\pi$ = 0.67 × $10^{-3}$) and *S. turkestanica* ($\pi$ = 0.83 × $10^{-3}$) compared to *S. tetrandra* ($\pi$ = 6.40 × $10^{-3}$) with relatively high diversity[30]. A later study confirmed a reduction in nucleotide diversity in cultivated *S. oleracea* compared to the wild species[33]. Nucleotide diversity estimates of wild and cultivated spinach accessions assessed with whole genome sequence based variants[34] were double that reported using the transcriptome variants[30]. Based on phylogenetic and population structure analysis, *S. oleracea* was more closely related to *S. turkestanica* than the *S. tetrandra*[30] and seconded by later studies[35,36]. In addition, analysis of transcriptome variants among cultivated and wild spinach accessions provided molecular evidence of recent domestication in spinach, and the *S. turkestanica* is the direct progenitor of the cultivated spinach *S. oleracea*[30]. The minimal differences found between the genetic diversity of the cultivated species *S. oleracea* and the wild progenitor *S. turkestanica* indicates a weak bottleneck during domestication[30]. The study of Xu et al. (2017)[30] provided further insight into the genomic basis of potential domestication traits. The genome-wide scan for allele frequency differences between wild and cultivated *Spinacia* species identified domestication signatures at the 93 regions ranging in length from 10 Kb to 150 Kb (totaling 2.3 Mb) containing genes associated with spinach domestication. Such highly divergent selective sweep regions include the QTL associated with domestication traits in spinach, such as bolting, flowering, leaf number, stem length, and petiole color.

The available genetic resources, domestication, and breeding history of spinach have been recently reviewed and indicated the loss of dormancy and morphological changes in pistillate flowers as the main traits involved in domestication syndrome[10] and hypothesized the presumed migration routes of spinach during domestication from former Persia to China and then to Europe. Genetic studies of the two wild species (*S. turkestanica* and *S. tetrandra*) and the cultivated species *S. oleracea* were performed to generate additional genetic evidence on potential migration routes, crop ancestry, and domestication[36]. Their study used a small set of selected SNP (n = 60), identified in *S. oleracea* to amplify all three species panels and found a higher amplification rate in *S. turkestanica* than *S. tetrandra* accessions indicating a wider divergence of *S. tetrandra* from other species[36]. The *S. turkestanica* accessions were genetically closer to *S. oleracea*

landraces from Eastern and Southern Asia, mainly Afghanistan and Pakistan, which are also the native distribution region of *S. turkestanica* in the study[36]. They identified that spinach might have been domesticated around these regions than previous assumptions and reports in former Persia. More precise estimates on the origin and domestication of spinach and its distribution to other areas can be elucidated and explained using genome sequence data, particularly more valuable for quickly decaying crops like spinach that limits long-term preservation.

Genes related to domestication and diversification of crop species have been identified following quantitative trait loci (QTL) fine mapping, genome wide association study (GWAS), and whole genome resequencing (WGR) studies[37] and have contributed to our understanding of crop domestication. Upon continued selection during domestication, loss of genetic diversity occurs within the selected genes while the unselected region retains diversity. Such selective sweep regions with severe loss of diversity can be scanned and identified using genome and population-wide variant datasets as performed in spinach[30,33]. Genomic studies and the use of genetic and molecular data provide evidence on the origin and evolution of crops and the genetic separation between wild and cultivated species, which adds value in informative conservation and utilization of available genetic resources. For instance, the domestication history of cultivated spinach, distribution and migration routes to current cultivation, and the genetic basis of human selection may be better explained with population-wide resequencing and pan genomic studies in the future[38–40].

## Germplasm and genetic resources

Spinach landraces and wild accessions are rich sources of many important horticultural traits and have been exploited in breeding improved spinach cultivars since the early 1900s[9,10,15]. More than 2,000 spinach germplasm accessions have been collected and maintained in the Centre for Genetic Resources, the Netherlands (CGN), Wageningen University and Research (WUR) (https://ecpgr.cgn.wur.nl/LVintro/-spinach). Around 400 *Spinacia* accessions are available at the United States Department of Agriculture, Agricultural Research Service (USDA-ARS) North Central Regional Plant Introduction Station in Ames, Iowa. The USDA Germplasm Resources Information Network (USDA GRIN) mostly maintains *S. oleracea* cultivars, landraces, and breeding lines, and small numbers of *S. turkestanica* (n = 12) and *S. tetrandra* (n = 15). Germplasm collections from these genebanks have been extensively used as genetic resources for basic and applied research and crop improvement programs. These spinach germplasms have been investigated for many traits of economic interest, including disease and insect resistance, stress tolerance, and improved horticultural traits. Passport data for many important traits are available in the GRIN database providing ready-to-use phenotype sets to scan and identify the genetic basis of the traits. The available germplasm resources for *Spinacia* species will continue to support the characterization of additional natural genetic variations for the trait of interest and perform fundamental research on expanding the physiological basis of the trait

control. The rapid evolution of genomics resources of spinach and available genomic technologies shows promising opportunities to explore hidden genetic treasures of the cultivated and wild *Spinacia* species. However, a small number of germplasm accessions of *Spinacia* species are available in the public gene banks. Only 89 *S. turkestanica* and 59 *S. tetrandra* are available at the genetic resource collection centers and most were collected from recent collection expeditions[41]. Collections of additional wild accessions from previously unexplored regions in the Middle East and South-West Asia for *S. tetrandra* and South and South-West Asia for *S. turkestanica* may provide expanded genetic resources for future breeding efforts.

Cultivated spinach intraspecific genetic variation is generally low, and so spinach breeders have relied heavily on an introgression from other wild species for many important traits. Wild species and landraces have been widely used in introgressing novel quality and disease-pest resistance traits in crop breeding. The wild species *S. turkestanica* and *S. tetrandra* are particularly relevant to modern spinach breeding programs. These species are the primary source of resistance against the downy mildew pathogen and may be valuable to introgress many other unknown but useful genes to develop superior cultivars in the future. Wild spinach has not been widely characterized for major economic traits and its genetic structure is only recently being explored[33,36]. Further, the two wild species are geographically restricted, and small numbers of germplasm collections maintained in the genebanks may limit future breeding and crop improvement progress. Thus, additional collection and conservation of wild species will be of value for future crop improvements. Increased availability of SNP and simple sequence repeat (SSR) markers in spinach[31,33,35] will supplement fingerprinting, diversity studies, and duplicate identification, enabling diverse core set development.

The cultivated and wild spinach germplasm accessions are reservoirs of valuable traits, and wide variations in morphology, biochemical, and tolerance to biotic and abiotic stresses are known[9,10,15]. Previous germplasm screening studies infer that additional screening of wild relatives and landraces for novel traits may accelerate spinach breeding and improvement efforts, including adaptation to future climates. The utilization of molecular and genomic resources to assess the genetic diversity and characterize the genetic variability of phenotype in the available genetic resources for spinach breeding is outlined in Fig. 3. Characterizing the genetic architecture and mechanism of traits in wild germplasm is now feasible at a reasonable cost with the declining sequencing and genotyping cost. However, several plants per accession are bulked for genotyping and/or sequencing, although many spinach germplasm accessions are still segregating for easily observable traits, which often limits precise mapping and identification of major and minor QTLs associated with phenotypes of interests[42]. Germplasm seed multiplications are discussed in van Treuren et al. (2020)[41]. Generating genetically uniform germplasm accessions remains a challenge, for which the use of single plants in a cross and during seed increase may help. Further, multi-environment evaluation of germplasm accessions for quantitative traits will generate more reliable phenotype data
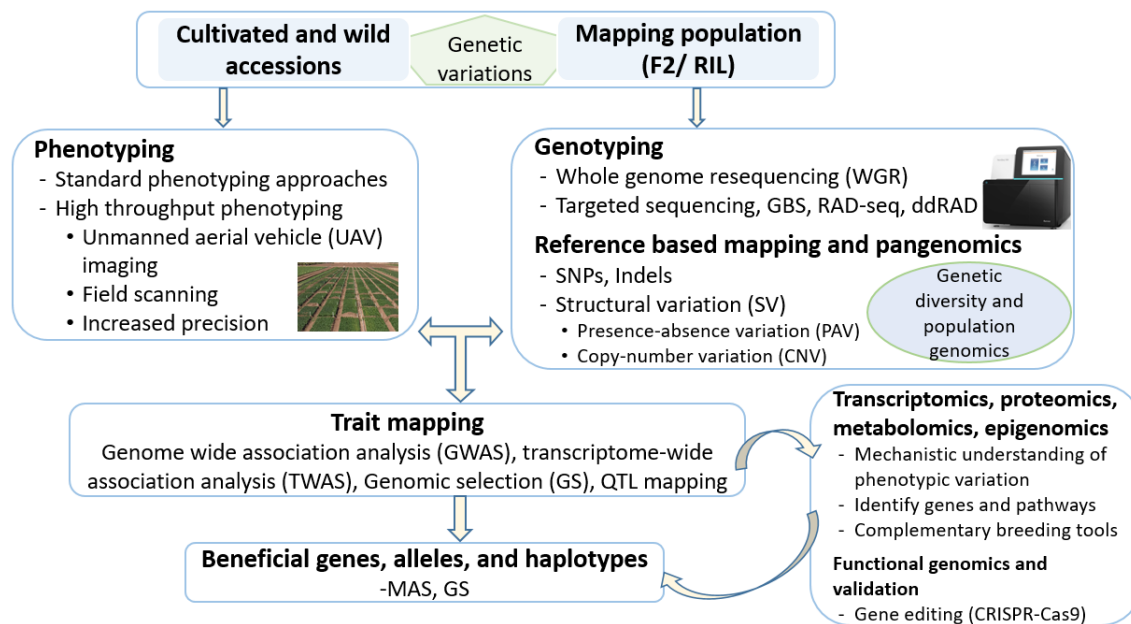
**Fig. 3** An overview of genome enabled breeding prospects of spinach.

and provide higher stability of the measured traits. The recent advancements in genomic resources in spinach and phenotyping methods make it possible to effectively profile and mine beneficial QTLs and alleles from the germplasm collections. The ongoing progress in developing genetic and genomic resources for wild and cultivated spinach germplasm and the parallel adoption of evolving phenomics and high-throughput phenotyping methods appears to advance knowledge on the biological mechanism of phenotype expression in spinach and facilitate breeding and spinach improvement efforts.

## Genetic diversity and population structure

Limited studies were reported on spinach genetic diversity and population structure in the past but have recently increased. Genetic diversity studies in spinach were initiated using 13 SSR markers in 33 *S. oleracea* hybrid cultivars from several breeding stations which revealed clustering of spinach hybrids based on their origin stations in the United States and Europe[43]. Another study analyzed 38 germplasm accession and ten commercial hybrid cultivars using the TRAP markers, although the accessions did not cluster based on the geographical origin[44]. Genetic differentiation and geographical variation assessment of diverse spinach accessions using SSR markers revealed the Western Asia accessions contain the highest gene diversity, and the East Asian and Japanese germplasm accessions showed clear differentiation from the European accessions[45].

Transcriptome and genome sequencing generated genome-wide marker datasets have been used to perform population genetic analyses of spinach in the recent past. Phylogenetic analyses performed using transcriptome sequencing derived 77,433 homozygous SNPs grouped three *S. oleracea*, three *S. turkestanica*, and three *S. tetrandra* into two clusters[46]. The *S. oleracea* and *S. turkestanica* accessions clustered together away from the *S. tetrandra* accessions

except for one *S. tetrandra* accession (PI 608712) grouped in the *S. oleracea* and *S. turkestanica* cluster. Previous reports have indicated the wrong classification of two accessions (Ames 23644 and PI 608712) as *S. tetrandra*[32], which were later suggested to belong to *S. turkestanica*[30]. In another study, phylogenetic analysis of 120 spinach accessions using 11,434 RNAseq derived SNPs formed three major groups comprising *S. turkestanica* and *S tetrandra* accessions, East Asian accessions, and accessions from Central/West Asia, Europe, and North America[30]. More recently, Gyawali et al. also found an admixture of PI 608712 and regarded it as a hybrid of *S. oleracea* and *S. turkestanica*[33].

Genetic diversity and the population structure analysis of the USDA germplasm accessions and commercial cultivars of spinach showed genetic grouping of these panels consistent with their geographical origins[31]. In addition, breeding lines from the University of Arkansas breeding program were uniquely differentiated from others[31]. Further, phylogenetic analysis of three *Spinacia* species showed clear separation of species except for some *S. turkestanica* population admixed with the *S. oleracea* group[36]. Genetic diversity and phylogenetic analysis were performed using SSR markers and grouped cultivated spinach germplasm according to their geographical origin[35]. Spinach accessions from East Asia were grouped separately from the worldwide collections[30,35,36], potentially indicating separate evolution events occurring in the East and Central/West Asian spinach accession.

The diversity and population structure of the cultivated and wild *Spinacia* species showed close clustering of the species indicating limited genetic divergence between the wild and cultivated *Spinacia* species. Overall these studies revealed that *S. turkestanica* was more closely related to cultivated species *S. oleracea*, and the cultivated species were genetically less diverse than the wild relatives. Less divergence of wild *Spinacia* species from the cultivated spinach supports recent domestication history from many of the

above-reported studies[30,33,35,36]. In reality, the wild species *S. turkestanica* and *S. tetrandra* have provided many resistance genes against the downy mildew pathogen and may have many other important genes conserved among the wild accessions.

Human selection and breeding efforts for growth and appearance, resistance against diseases and pests, and tolerance to abiotic stresses in different geographic regions and climatic conditions generate phenotypic differences among the cultivated and wild accessions that can be tapped for crop improvement[39,40]. It is thus essential to preserve, maintain, and characterize these important wild species as the availability of diverse genetic resources is key in crop improvement. Understanding the population structure and phylogenetic relationship of the germplasm accession maintained in the genebanks will facilitate efficient management and utilization. Further, the intraspecific and interspecific genetic diversity and relationship, identifying duplicates, and cataloging unique accessions will optimize the utilization of the genetic resources. Many previous genetic diversity studies of spinach involved low-density markers and/or a small panel of genotypes. The WGR effort and increased availability of molecular and genomic resources in spinach may offer a population genomics approach to extend investigation on domestication history, genetic diversity and structure, and genome-wide selection patterns. Moreover, the diversity lost during domestication should be strategically mined for genetic variants associated with beneficial effects on modern agriculture and utilized in breeding programs[37,40]. The immediate goal in spinach is to identify the most diverse and potentially most promising germplasm using genomics, transcriptomics, proteomics, metabolomics, and epigenomics platforms to investigate and gain new insights on the molecular basis of complex traits and unravel the biological networks.

## Molecular, genomic, and transcriptomic resources

During the early 2000s, genome sequencing and marker discovery efforts were expensive, and limited molecular-genomic resources were available for crop species. There appeared limited genomic resources in spinach until recently; however, from mid-2010, substantial genomic data and genomic resources have been generated and are publicly available, as summarized in Table 1. These new genomic resources are anticipated to increase genetic studies of many important traits and facilitate the molecular breeding effort. Several genome sequences and genomic resources are being generated, and with the reduced sequencing cost, additional genome resources are likely to be continually added. High-quality reference genome assembly is a prerequisite for in-depth genetic investigation and gene mapping experiments. The availability of the reference genome sequence and transcriptome sequences of cultivated and wild spinach accessions has expanded biological and genetic research towards crop improvement. The genotyping of natural population and segregating panels necessary for quantitative genetics, trait mapping, and population genetics studies is now feasible and affordable with the advancement of next-generation sequencing technologies. A rapid increase in molecular and genomic resources has been observed in spinach over the past five years. Essentially, the reduced cost of sequencing and genotyping will facilitate genetic mapping and generate molecular tools to accelerate the development of improved crops via genomics–assisted breeding. In addition, seeing the rapid progress on expanded genomic resources and multiple genome assemblies, researchers and the breeding community working on spinach would have ample genetic and genomic resources on hand to use in understanding and expanding the genetic mechanism of major traits and subsequently develop genetic tools and strategies to improve crop production and minimize crop loss. The proposed breeding strategy emphasizing the opportunity to mine the genetic variation of the phenotypes across the available genetic resources and utilizing the genome-enabled tools to improve spinach is outlined in Fig. 3.

### Molecular resources

Initially, the mitochondrial genome was sequenced and sequence homology between the nuclear and chloroplast genomes was studied[47–49]. Next, a 327 Kb physical map of the mitochondrial genome was constructed[50]. Afterward, the chloroplast genome from spinach was sequenced and 150 Kb nucleotide sequences were generated[51]. The spinach genome was characterized for the number of chromosomes, chromosomal length, and arm ratios using the fluorescent in situ hybridization (FISH) method, and 45S and 5S rDNA sites were physically localized on three chromosomes[52].

Molecular marker discovery in spinach began by mining and developing 12 SSR markers from spinach nuclear sequences deposited in the EMBL and GeneBank databases[53]. An initial genetic linkage map of spinach was constructed using 101 amplified fragment length polymorphism (AFLP) and 9 SSR markers from back cross-generated 161 progeny population segregating for male and female gender[17]. The markers in the map were grouped into seven linkage groups (LG), although spinach has six chromosomes. The total map length was 585 cM with an average marker interval of 5.18 cM and the sex determination locus in this study segregate in a 1:1 ratio of male and female, indicating a single locus controlling the sex determination in spinach[17]. Their study also emphasized that modifier genes at other loci may be involved in gender regulation[17]. The gender determining locus was mapped to the LG 3 at 101.5 cM flanked by five markers within 5.2 cM. In another study, target region amplification polymorphism (TRAP) markers were generated and used to evaluate the genetic diversity among spinach germplasm accessions and commercial cultivars[44].

Molecular and genomic resources have been rapidly generated over the past decade, particularly for minor crops like spinach, with the decreasing cost of high-throughput sequencing. Additionally, the bacterial artificial chromosome (BAC) library was constructed from a near-isogenic line (NIL1) to initiate and expand the spinach genomic resources and to increase the understanding of host resistance against downy mildew disease[54]. The 1920 random BACs were end sequenced to generate the new set of molecular markers, and 100 SSR loci were identified. The BAC-end sequence assembly provided newer insights into downy mildew resistance and

**Table 1.** Summary of major genetic and genomic resources available for spinach.

| Resources generated | Outcomes of the activities | References |
|---|---|---|
| DNA markers | 12 nuclear SSR markers were identified and amplified in spinach. | [53] |
| Genetic map | 110 markers (101 AFLP and 9 SSR markers) were assigned to linkage groups. The gender determination gene was mapped to the linkage group 3 at 101.5 cM nearby five other markers between 97.4–102.6 cM. | [17] |
| DNA markers | Target region amplification polymorphism (TRAP) markers were generated and were used to evaluate the genetic diversity among spinach germplasm accessions and commercial cultivars. | [44] |
| Bacterial artificial chromosome (BAC) library construction | BAC library was constructed from a near-isogenic line (NIL1). Fourteen sequences analogous to known plant disease resistance genes were identified. Of the 14, nine contained nucleotide binding sites leucine-rich-repeat (NBS-LRR) domains. One hundred SSR loci were identified. One primers pair designed from these putative resistance genes was closely linked to the *RPF*1 locus and cosegregated with the DM1 marker. | [54] |
| DNA markers | SSR loci were mined from Illumina sequences and a set of 85 polymorphic SSR markers were validated in 48 spinach accessions. | [55] |
| DNA markers | Thirtyfour new polymorphic SSR markers were identified and genetic diversity was assessed on Chinese spinach accessions. | [56] |
| DNA markers | Around 6,000 polymorphic SSRs reported in spinach following *in silico* genome sequence comparison. | [35] |
| Whole genome sequence | A draft genome sequence of spinach cultivar was assembled to 498 Mb. | [57] |
| Annotation and gene set development | Genome wide gene set was developed for spinach using the draft genome assembly[57] and mRNA sequencing from inbred spinach cultivar Viroflay. | [58] |
| Transcriptome sequences | Transcriptomes from nine *Spinacia* accessions (three each from cultivated *S. oleracea*, and two wild *S. turkestanica*, and *S. tetrandra*) were sequenced at a high depth, genes were functionally annotated, Gene Ontology (GO) terms were assigned, and metabolic pathways were predicted. The SNP variants detected from the transcriptome assembly among the sequenced accessions were used to infer phylogeny, evolution, and domestication history. | [46] |
| Whole genome assembly (Spov1) | The Spov1 genome sequence assembly developed for an inbred spinach line Sp75 provides a comprehensive genomic resource for spinach. The genome sequences were assembled to six linkage group that covers 463.4 Mb constituting 47% of the assembled genome. A total of 25,495 protein-coding genes were predicted, of which 139 NBS-LRR genes were identified that are known to provide disease resistance in plants. | [30] |
| Transcriptome assembly panel | The transcriptome sequencing of 120 cultivated and wild *Spinacia* accessions resulted in a large number of transcribed variants and gene expression profiles. Nucleotide diversity and selection sweeps analysis were performed. The *S. tetrandra* was highly diverse compared to the other two *Spinacia* species, and *S. turkestanica* was the progenitor of cultivated spinach *S. oleracea*. | [30] |
| Spinach genomic database | A publicly accessible database storing the reference spinach genome sequence, predicted functional annotations, gene expression profiles, and genetic variations based on transcriptome sequences of 120 cultivated and wild *Spinacia* accessions, and a platform to query and analyze genomic data were developed and hosted at SpinachBase (www.spinachbase.org). | [60] |
| Spinach genome assembly (Spov3) | Another comprehensive whole genome assembly Spov3 generated for monoecious cultivar Viroflay (https://phytozome-next.jgi.doe.gov/info/Soleracea_Spov3). The genome was assembled to 913.5 Mb and the six main pseudomolecules comprised 745 Mb (81.6% of the genome) and predicted 34,877 genes in spinach. | [61] |
| Spinach genome assembly (SOL_r1.1) | The SOL_r1.1 genome of breeding line 03-009 was assembled to 935.7 Mb of which 686.6 Mb (73.5%) was anchored to six chromosomes (http://spinach.kazusa.or.jp/index.html). | [62] |
| Spinach genome assembly (Monoe-Viroflay) | Chromosome-scale reference genome assembly comprising 894.3 Mb with 879.2 Mb (98.32%) anchored to six chromosomes, and 28, 964 protein-coding genes were predicted. Genome resequencing of 305 spinach accessions and genome wide variant data are available. | [34] |
| Whole genome resequenced (WGR) GWAS panel | A large set of 480 USDA accessions and commercial cultivars have been sequenced at a lower coverage to serve as the association panel for spinach. Several economically important traits have been phenotyped and genome wide association analysis (GWAS) are conducted to map the trait at a high-resolution mapping to identify causal variants. | [63] |

14 sequences analogous to known plant disease resistance genes, of which nine contained nucleotide binding sites leucine-rich-repeat (NBS-LRR) domains. One primer pair designed from these putative resistance genes was closely linked to the *RPF*1 locus and co-segregated with the DM1 marker[54]. Recently, 85 SSR loci mined from Illumina sequences were validated in 48 spinach accessions[55]. Another study reported a genetic diversity assessment of Chinese spinach germplasm collection using 41 SSR markers, including 34 new markers identified from Spov1 assembly and seven previously reported SSR markers[56]. Several thousand SSR markers were recently identified from the Spov1 assembly following computational screening for polymorphism by searching for repeat motif variation across

genome sequences of 21 spinach cultivars[35]. A subset of SSR markers was genotyped for molecular validation in 48 spinach accessions, and diversity analyses were performed[35].

## Genomic and transcriptomic resources

A draft genome sequence of spinach cultivar Viroflay was assembled to 498 Mb[57] that represents half of the estimated genome size (989 Mb) of spinach[14]. Comparative genome and evolutionary analysis of this draft spinach genome identified gene features that distinguish *Caryophyllales* species (sugar beet and spinach) from rosids and asterids[57]. The genome assembly [57] and mRNA-seq reads generated from inbred spinach cultivar Viroflay were annotated that provide a genome-wide gene set for spinach[58]. In addition,

the transcriptome sequence of nine spinach accessions that include three accessions each from the cultivated *S. oleracea*, wild *S. turkestanica,* and *S. tetrandra,* generated high quality assembled unigenes[46]. Furthermore, functional annotation of 42,019 spinach unigenes was performed by comparing against other protein databases, Gene Ontology (GO) terms were assigned for 34,522 unigenes, and 387 metabolic pathways represented by the 2,785 unigenes were predicted[46]. Phylogenetic analysis performed using the transcriptome-derived SNP variants indicated that the wild species *S. turkestanica* was more closely related to cultivated *S. oleracea* than the other wild species *S. tetrandra*[46].

Another draft genome assembly and annotations were generated from an inbred spinach line, Sp75, using the whole genome shotgun approach combined with BioNano Genomics optical maps, and the scaffolds were anchored using a high-density genetic map[30]. This Spov1 genome sequence represents a 168-fold coverage of the spinach genome assembled to 996 Mb with an N50 scaffold length of 919.2 Kb. The Spov1 genome sequences were assembled into six linkage groups covering 463.4 Mb, constituting 46.52% of the assembled genome and providing a comprehensive genomic resource for spinach[30]. From the Spov1 reference spinach genome, a total of 25,495 protein-coding genes were predicted, of which 139 NBS-LRR genes were identified that are known to provide disease resistance in plants[30]. The DM-1 marker tightly linked to the *RPF*1 locus[59] was mapped to the spinach genome and five NBS-LRR genes near the DM-1 markers were identified and were postulated as the potential downy mildew resistance candidate genes in spinach. In addition to whole genome sequencing, the transcriptome sequencing of 120 cultivated and wild *Spinacia* accessions resulted in a large number of transcribed variants and gene expression profiles[30]. A publicly accessible database storing the Spov1 reference spinach genome sequences, predicted functional annotations, gene expression profiles, and genetic variations based on transcriptome sequences of 120 culti-vated and wild *Spinacia* accessions, and a platform to query and analyze genomic data was developed[60]. Further analysis of the transcription variants among the spinach accessions has expanded our understanding of genetic diversity and domestication, explored loss and gain of genomic regions from domestication and human selection, and provided comprehensive genomic control of bolting and flowering traits. Based on nucleotide diversity among the transcriptome of three *Spinacia* species, the *S. tetrandra* was highly diverse compared to the other two species. And a small difference in genome-wide diversity between *S. oleracea* and *S. turkesta-nica* suggests a weak bottleneck during domestication and that *S. turkestanica* is the progenitor of cultivated spinach[30]. The selection sweeps analysis of the transcriptome data identified a highly divergent region that coincides with domestication traits such as bolting and flowering[30]. The highly divergent genomic region between *S. turkestanica* and *S. oleracea* includes QTLs associated with bolting, petiole color, and erectness traits.

Additionally, a second genome assembly has been generated for the monoecious spinach cultivar Viroflay[61]. Briefly, the Viroflay genome was sequenced using short–reads in Illumina and long-reads in Pacific Biosciences SMRT technology. The independent assembly and scaffolding of the Illumina sequences generated Spov2 assembly of 968.8 Mb and containing 26,862 genes. The Pacific Biosciences libraries and Illumina PE data were assembled to 913.5 Mb, repre-senting 70x genome coverage that yielded Spov3 assembly, available at https://phytozome-next.jgi.doe.gov/info/Sole-racea_Spov3. The six main pseudochromosome of Spov3 assembly comprised 745 Mb (81.56% of the genome) with 34,877 annotated genes, of which 1,004 disease resistance genes were identified. Thus, the Spov3 genome added 282 Mb chromosome anchored sequences and over 9,300 new genes in spinach, essentially improving genomic resources to support and advance genetic studies in spinach. In addition, the SOL_r1.1 genome assembly using PacBio long reads and Illumina short reads were released, constituting 935.7 Mb, of which 73.59% of the whole assembly were anchored to six pseudochromosomes[62]. A new chromosome-scale reference genome assembly comprising 894.3 Mb and an N50 contig size of 23.8 Mb was generated from an inbred line 'Monoe-Viroflay'[34], and around 98.32% of the assembly was anchored to six chromosomes. The chromosome anchored sequences in the Monoe-Viroflay assembly were twice the SpoV1 assembly and were significantly higher than the Spov3 and SOL_r1.1 assemblies. A total of 28,964 protein-coding genes were predicted in the Monoe-Viroflay assembly, with 115 NBS-LRR genes. The genome sequencing, with an average depth of 15.9x of 305 wild and cultivated spinach accessions (295 *S. oleracea*, 7 *S. turkestanica*, and 3 *S. tetrandra*), generated filtered sets of publicly available 5,511,663 SNPs and 55,330 structural variants (SVs) in the populations[34].

Essentially, the genome assemblies for multiple genotypes and resequencing of germplasm and commercial cultivars available in spinach are valuable resources in understanding genetic diversity, conducting comparative genomic studies, and studying molecular mechanisms underlying several important horticultural traits. The chromosome anchored reference genome assemblies have facilitated WGR of additional germplasm accessions, parents of mapping populations, and differential cultivars. In addition, a large set of USDA accessions, breeding lines, and commercial cultivars (n = 480) have been sequenced at an average depth of 10x, serving as the association panel in conducting GWAS to fine map the important phenotyped traits and identify functional genes and markers[63]. Furthermore, differential cultivars and parents of mapping populations (n = 30) were sequenced at 30x to generate high-resolution gene mapping and identify SVs associated with the phenotype[35]. Advanced genomic resources and the contiguous chromosome-scale genome assemblies for spinach and the WGR variants will likely facilitate genetic mapping and gene function studies. In the past decades, the molecular and genomic resources developed for *Spinacia* species enabled understanding of available genetic diversity, molecular dissection of prioritized traits, and the development of associated molecular markers, which have facilitated increasing selection efficiency and genetic gains. Strategic utilization of available genomic resources and new omics technologies in gene mapping to comprehensively understand the molecular mechanisms of horticulturally important traits and the functions of genes involved in trait expression will translate into genome-

enabled breeding and development of high yielding stress-tolerant, climate-resilient, and nutrient-dense spinach cultivars.

## Breeding strategy and focus

The main objective of the spinach breeding program is to increase yield and improve the important horticultural traits and crop performance. The major goal of spinach breeding programs is to develop varieties with traits including disease resistance and abiotic stress tolerance, slow bolting and improved yield and quality, and decreased levels of nitrate, oxalate, cadmium, and increased levels of folates[9,15,28]. Increased nitrogen-use and water-use efficiencies are emerging focus in spinach breeding programs. Traditional breeding approaches of mass selection and recurrent selection methods are commonly employed in spinach breeding programs. Other practices include a repeated selection from F1 progenies generated from a cross of two selected parents to generate a more homogeneous inbred population and the use of monoecious plants to develop inbreds via selfing. In the beginning, a genetic investigation on spinach was conducted for spiny versus smooth seed type, smooth versus savoy leaf texture, and short versus long petiole length[64]. All these studied traits were segregated in a Mendelian pattern and were controlled by a single gene. Until recently, breeding efforts in developing spinach cultivars relied on traditional selection approaches against the biotic and abiotic stresses, and many successful cultivars have been released.

A short overview of the effective means of breeding strategy to generate spinach hybrid cultivars is presented here. The dioecious and monecious flowering mechanism in spinach and lack of incompatibility permit hybrid seed production, and most of the cultivars in the United States, Europe, and Japan are F1 hybrids. Recently, breeding and releasing open-pollinated spinach cultivars have been prioritized at public institutions but have not yet been achieved. Initially, the F1 hybrids were produced by planting dioecious parental lines side-by-side and rouging the male plants from the female plant row before pollen sheds to ensure pollination from the intended male parent lines[65]. Hybrid seed generation in spinach uses monoecious inbred lines[25] and pure female lines to cross with other pollinator lines been outlined[20,65]. Currently, hybrid seed breeding employs pure female monoecious lines and the highly male monoecious inbred lines as pollen parents to produce hybrid F1 cultivars. However, female lines occasionally produce pollen under favorable environments leading to unintended inbred seeds in the hybrid seed lots that lack desired resistance genes from the male parent. In general, an isolation distance of three miles is practiced between two populations for commercial seed production. Long day length and mild cool temperatures favor commercial spinach seed production in Denmark, and so around 70% of worldwide spinach seeds are produced in Denmark[9,15]. The agro-climatic condition in the Pacific Northwest in Washington and Oregon in the United States fits spinach seed production and comprises around 20% of the world's spinach seeds[66].

The commercial production of spinach is affected by both biotic and abiotic stresses. Spinach cultivars are selected to meet the needs of commercial growers, production practices, and market demands. Spinach populations are selected for the seedling vigor beginning from the seedling stage and are continually selected over the production cycle. Many plant pathogens affect leafy green spinach, and insect damage is also of economic concern. The most important abiotic factors are drought or high water use efficiency and, more recently, cadmium. Similarly, major insect pests common in spinach production include aphids (*Myzus persicae*), leafminer (*Liriomyza langei*), and to some extent, thrips and spider mites. Many reports indicated that genetic variations exist in the available germplasm for many of the traits such as aphid resistance[67], leafminer resistance[68], cadmium tolerance[69], and drought tolerance[70]. This existence of quantified genetic variation among the genetic stocks allows breeding for tolerant cultivars.

For the morphological traits, spinach plants are selected for leaf type, shape, and texture. Plants are selected based on the savoy and smooth-leaf types and leaf shape that ranges from round to pointed. The smooth leaf with round leaf shapes has been more desirable in recent decades. In addition, spinach plants are selected for seed shape and type. Spinach seed types vary from round to spine, although round seed types are standard in all commercial cultivation practices[71]. Likewise, leaf color is another commercially important trait in spinach. A general variation in the green leaf color exists in spinach, with a dark green color being favored and selected distinctly for the baby leaf and processing cultivars. Recently a red leaf spinach cultivar was released containing significantly higher betacyanin concentrations than other red-veined and green leaf cultivars[72]. Spinach crops are also selected based on plant stature and leaf orientation, as upright foliages are attractive for mechanical harvesting. Likewise, late flowering and bolt resistance are other important traits in spinach breeding programs for which continuous selection against early flowering plants in the breeding population is practiced. Most modern spinach cultivars have been selected for resistance to bolting (slow-bolting) in the longer and warmer days, and the new cultivars can be grown all year round at varying latitudes.

The current emphasis in spinach breeding, besides the horticultural traits, is on improving disease resistance. Most of the effort in spinach is on breeding for downy mildew resistance as it is the economically most important disease[9]. In recent years, research interests have been increasing to evaluate and improve the nutritional content and biofortification, develop cadmium tolerant cultivars, and standardize vertical-indoor spinach cultivation. The establishment of commercial spinach production under vertical farming needs cultivars adapted to the indoor cultivation practices for which breeding and selection effort under such a new production system will be directed to fit the needs. Under those circumstances, many other abiotic and biotic stresses affecting commercial spinach production may appear for which new research and breeding needs may emerge. The cultivated and wild germplasm will be valuable resources to study and identify tolerance traits and genes that can be incorporated into cultivars.

## Major diseases affecting spinach

Numerous diseases caused by fungus, bacteria, and viruses affect spinach and reduces yield and quality[73]. Downy mildew, white rust, Fusarium wilt, leaf spot disease caused by *Stemphylium* species, and *Pythium* damping-off are the economically important diseases affecting commercial spinach production. Fresh market spinach production has dramatically increased in the past two decades in the United States[2]. Spinach is commercially produced in a high density of 5−10 million seeds/ ha in California, Arizona, Texas, and other areas (Fig. 4). Such high-density planting leads to dense canopy cover, extended leaf wetness, and poor air circulation that essentially increases humidity and creates ideal conditions for disease development. An equally important fact is that organic spinach production comprises around 50% of total spinach production in the United States[15]. Diseases represent a significant constraint in spinach production, particularly organic production, where disease management options are somewhat limited[73].

Downy mildew caused by an obligate oomycete *Peronospora effusa* (former *P. farinosa* f. sp. *spinaciae* [*Pfs*]) remains the most damaging disease affecting spinach. If not managed using fungicides, it causes complete crop loss but can only be used in conventional production (Fig. 5). Downy mildew in spinach was first reported in 1824, and a physiological race of downy mildew race 1 was observed in Texas and California during the 1950s[74]. Three additional races were reported until the late eighties[75−78]. Another physiological race of downy mildew, race 4, was identified in 1990[79]. Similarly,



**Fig. 4** High density spinach planting in Salinas, California and Yuma, Arizona, USA.



**Fig. 5** Signs and symptoms of downy mildew disease in spinach. Plants (middle two) inoculated with *P. effusa* in the greenhouse show sporulation and chlorosis. Baby leaf spinach plants growing in commercial field conditions, the resistant cultivars (left) are clean while the susceptible cultivars (right) are entirely infected.

race 5 of *P. effusa* was identified in 1996 in the United States and 1998 in Europe, race 6 in 1998 in the United States, and race 7 in 1999 in Europe[80]. An additional three races (races 8, 9, and 10) of spinach downy mildew pathogens were reported[81]. More races (11, 12, 13, 14, 15, 16, 17) were recently reported[82−85], making a total of 19 unique races. A standard differential set of spinach lines are available that can unequivocally differentiate all identified *P. effusa* races[81,83,84]. Downy mildew disease resistance reactions of the differential cultivars to all 19 known races of *P. effusa* are available (https://ucanr.edu/blogs/blogcore/postdetail.cfm?postnum=46392), and the virulence pattern of the races are updated by the International Seed Federation (ISF) at www.worldseed.org/our-work/plant-health/differential-hosts/.

Genetic resistance has been utilized to manage downy mildew disease and most released resistant spinach cultivars were bred using major genes effective against certain pathogen races[9]. Of the 19 reported *P. effusa* races, more than ten races were identified in the last three decades. Continuously emerging *P. effusa* races break down the resistance genes deployed in the newly released cultivars and remain the major setback to breeders and other stakeholders, including growers and the seed industry. A significant increase in the production area in the last three decades, planting in a higher density and year-round production, planting of resistant cultivars with a narrow genetic background that increases selection pressure on pathogens, continuous increase in the organic production area, are all conditions that are conducive for the emergence of a new race. New races are likely a result of asexual variation[86] and sexual recombination[86,87]within the pathogen populations. Recently, detached leaf assay was standardized to differentiate resistant and susceptible spinach genotypes and potentially characterize the *P. effusa* races[88]. Investigation on different moisture regimes on the viability of *P. effusa* found a significant reduction in viability under desiccation and lower moisture regimes[89]. Examination of oospores of *P. effusa* for sexual cross-compatibility on the detached leaf produced oospores, demonstrating the presence of opposite mating types in spinach production areas in the United States[89]. Additional studies in the future may expand knowledge on population genetics and the evolution of new pathogen races.

Most of the downy mildew resistant spinach cultivars were bred using single dominant resistance genes against races of *P. effusa* known as RPF (Resistance to *P. farinosa*). These hypothesized RPF loci were used to develop near isogenic lines (NILs) by backcrossing the resistant line with the susceptible cultivar Viroflay as the recurrent parent for four generations and continuous selection for resistance to achieve homozygosity at the resistance locus[9]. Regular breakdown of the resistance loci in spinach by the new races of the pathogen has increased research agendas to expand the current understanding of host-pathogen interaction in spinach downy mildew that mainly includes identifying and mapping genes from known available resistant materials, a functional test of the *RPF* genes, and characterizing functions of the effector genes[42]. Gene pyramiding and cultivating a mixture of lines are standard practices to slow down R-gene breakdown and are the strategies adopted to manage downy

mildew disease in spinach. Mechanistic understanding of spinach and *P. effusa* resistance and susceptible disease interactions, unraveling the involved genes and their functions, and characterization of virulence evolution of races of *P. effusa* will contribute to formulating a new breeding strategy to implement genetic resistances to manage downy mildew disease effectively.

White rust disease caused by obligate oomycete pathogen *Albugo occidentalis* G. W. Wils causes severe yield losses of spinach[9,90]. It is prevalent in southern Texas but does not occur in California or Arizona[91]. Symptoms begin with small chlorotic lesions[73], while leaf yellowing and stunted growth are major issues at later stages as the disease progress[92]. White rust spreads rapidly under a favorable environment and can substantially impact yield and quality. Quantitative resistance has been used to develop white rust resistant spinach cultivars[15]. White rust prevalence has recently been reported in many new spinach production areas in Greece[93], Mexico[94], and Turkey[95], threatening stakeholders around the world.

Fusarium wilt of spinach, caused by the soil-borne fungus *Fusarium oxysporium* f. sp. *spinaciae* (*Fos*), is a severe disease in spinach seed production in the Pacific Northwest USA[73]. Control of Fusarium wilt is a challenge as the fungus can survive for more than ten years as chlamydospores or colonize on the roots of non-host plants[96,97]. No major resistance genes have been identified, but QTLs have been reported[98]. Several bi-parental populations are being characterized to map the resistance locus from *S. oleracea* and *S. turkestanica* and harness the novel resistance alleles in cultivated spinach. Similarly, Verticillium wilt caused by *Verticillium dahliae* is a severe problem in spinach seed production but is also damaging to fresh or processing spinach production in the United States[99].

Leaf spot disease in spinach, caused by *Stemphylium vesicarium* and *S. beticola,* can reduce yield and quality and is of major concern in spinach production in the United States and other spinach producing countries[100,101]. The *Stemphylium* pathogen was identified in the Salinas Valley, California, starting in 1997. The pathogen can be seed-borne and seed-transmitted and has been present in the seed lots from the United States, Denmark, the Netherlands, and New Zealand[102]. A recent survey reported *Stemphylium* leaf spot predominates in Arizona, California, South Carolina, and Texas in the USA[100]. The USDA spinach collection and commercial cultivars were screened for an isolate of *S. botryosum*[103] (recently renamed as *S. beticola*), from which eight SNP markers associated with resistance were identified[104]. Around 300 USDA spinach accessions and 30 commercial cultivars were recently screened for resistance against *S. vesicarium* under greenhouse conditions and GWAS have identified markers associated with resistance[105,106].

*Pythium* species colonizes on spinach under wet soil conditions affecting seed germination, seedling emergence, and stand establishment. Infestation of *Pythium aphanidermatum* causes damping-off and root rot of spinach[107,108]. In recent years, *Pythium* infestations impact both field cultivation and soil-less spinach production. *Pythium* severity depends on cultivars, soil texture, irrigation, and pathogen populations. *Pythium* disease outbreak occurs at moderate to higher temperatures and humidity[109]. There is increasing interest in growing spinach and other leafy vegetable crops under controlled environment facilities and hydroponic production systems, for which *Pythium* infestation is one of the major setback. Thus, the spinach germplasm panel is being evaluated for resistance against *Pythium aphanidermatum* to identify molecular markers associated with the resistance.

## Molecular breeding progress

The lack of high-density genetic maps, molecular markers, and genomic resources slowed the integration of molecular breeding in spinach, and, until recently genetic research and breeding progress were mainly based on conventional approaches. The availability of a high-quality reference spinach genome has facilitated genetics and genomics research and molecular breeding in spinach. Genetic control of traits has been studied and linked, or associated markers have been developed in bi-parental or natural populations using a range of molecular markers. Several studies examined the genetic control of sex and have identified markers to determine sex in spinach. The first framework genetic map of spinach was constructed using the AFLP and SSR markers that mapped the sex locus to LG3[43]. Similarly, bulk segregation analysis identified markers closely linked to the *Y* locus[110], and a later study mapped the monoecious gene ($X^m$) close to the dioecious locus (*X/Y*)[20]. An SNP-based genetic linkage map was constructed in the spinach $F_{2:3}$ population segregating for nitrogen use efficiency (NUE)[111]. RNA sequences from two parents with contrasting NUE were used to identify SNPs and the 320 segregating F2 lines were genotyped by converting SNPs into Kompetetitive allele specific PCR (KASP) assay. The map contained 283 SNP markers grouped into six linkage groups with a total map distance of 433.6 cM, including 39 QTLs associated with the NUE in spinach. In 2017, a Loop-Mediated Isothermal Amplification (LAMP) marker linked to the male-determining locus and capable of distinguishing sex in spinach was reported[112]. Another study described a high-density genetic map of spinach using 4,080 specific-locus amplified fragment sequencing (SLAF-seq) markers spanning 1,125.97 cM in six LG with an average distance of 0.31 cM between markers[113]. The six LG corresponded to the total chromosome numbers in spinach. This study mapped the sex determination locus (*X/Y*) to two close regions of LG 4 (66.98–69.72 cM and 75.48–92.96 cM) containing 166 genes[113]. The male determining region was narrowed to 58.76–58.78 Mb region of Spov1 chromosome 4, and KASP markers to differentiate the XX, XY, and YY plants were reported[22].

The *RPF*1 locus governed by the single dominant allele was mapped to chromosome 3 based on linkage between progenies segregating from cultivar Lion (against race 6 of *P. effusa*) and AFLP markers[59]. Co-dominant sequence characterized amplified region (SCAR) marker Dm-1, which is 1.7 cM from the *RPF*1 locus, was developed and has been widely used to select for downy mildew resistance in spinach. Later, marker 5B14r was identified from the resistance gene analogs (RGA) co-segregating with the DM1 marker[54]. The other *RPF* loci (*RPF*1, *RPF*2, *RPF*3) were mapped to a 1.5 Mb region of chromosome 3 and diagnostic PCR markers were

reported to distinguish resistant and susceptible genotypes[114]. A total of 139 NBS-LRR genes involved in resistance against pathogens were identified in the spinach genome[30] and five resistance genes likely to provide resistance against the downy mildew disease were predicted near the *RPF* locus region. The *RPF*1 locus was reported within the 0.89 Mb region following protein homology comparison between resistant and susceptible segregating lines[115]. The downy mildew resistance region was finely mapped using the genotyping by sequencing (GBS) derived SNP markers in segregating populations generated from a cross of *P. effusa* race 13 resistant cultivars Swan, T-Bird, Squirrel, and Tonga with susceptible cultivars Whale and Polka[42]. In the same way, resistance against race 16 of *P. effusa* was mapped to a 0.57 Mb region of Spov1 from population segregating from cultivars Whale and Lazio, and resistance-associated SNP markers were identified[116]. The proximal end of Spov1 chromosome 3 (0.34−1.26 Mb region) contains 14 annotated plant disease resistance providing candidate genes. Previous studies mapped the *RPF* loci in this region at different resolutions and narrowed the resistance gene region to 1.5 Mb[114], 0.89 Mb[115], 0.84 Mb[42], and 0.57 Mb[116]. These mapping efforts pointed to different sets of genes involved in providing resistance with overlap between studies, as discussed in Bhattarai et al. (2021)[116]. Resistance genes are often found in clusters and are highly variable in copy numbers[117], and new resistance specificities are usually formed via unequal crossing-over and gene conversion[118,119]. The organization of such gene clusters complicates identifying genes responsible for resistance to specific races of the pathogen. Regardless of the wealth of genetic and genomic resources available and research conducted on spinach against *P. effusa* resistance, the functional R genes remain to be identified. Furthermore, additional downy mildew resistance mapping efforts have evaluated a set of USDA germplasm and commercial cultivars for tolerance to the natural population of *P. effusa* in the field condition[120,121] and race 5 of *P. effusa*[122] in the controlled environment facility to identify a set of minor and major QTLs.

The recent molecular breeding efforts in spinach used the genome-wide association analysis approach to examine the genetic polymorphism and their statistical association with the variation in the phenotype in diverse natural germplasm panels to take advantage of historical recombination events reflected in the linkage disequilibrium (LD). In addition, the rapid LD decay of between 4−12 Kb[30,33] in an open-pollinated *Spinacia* species facilitates mapping of the trait at a higher resolution though it depends on phenotyping accuracy and the number of genotyped markers. GBS discovered SNP markers on 268 USDA Grin spinach germplasm, 45 commercial hybrids, and 30 University of Arkansas cultivars/breeding lines[31] had been used to perform GWAS analysis to identify SNPs associated with several economically important traits in spinach. GWAS analysis reported SNP markers associated with surface texture, edge shape, and petiole color of spinach[123]; bolting, tallness, and leaf erectness traits[124]; leafminer (*Liriomyza* spp.) resistance[125]; oxalate content[126]; verticillium wilt caused by *Verticillium dahliae* Kleb.[127]; *Stemphylium* leaf spot[104]; mineral nutritional content in spinach leaves[128]; resistance against the anthracnose disease

caused by *Colletotrichum dematium*[129]; and white rust resistance evaluated in the disease nursery in Southern Texas[130]. High-throughput phenotyping platforms captured images from the unmanned aircraft have identified SNP associated with plant growth-related traits like canopy cover, canopy height, and canopy volume[131]. Additionally, several SNP markers have been identified that are associated with white rust resistance in USDA spinach germplasm, commercial hybrids, and Arkansas spinach lines evaluated for multiple years in the Southern Texas area (Shi et al., unpublished data). Several other preliminary GWAS studies have reported markers associated with spinach phenotypes, including vitamin C content[132], downy mildew resistance evaluated under natural inoculum pressure in the field[120,121], resistance to race 5 of downy mildew evaluated in the greenhouse[122], resistance to *Stemphylium vesicarium*[106], and Fusarium wilt resistance[98,133]. The identified SNP marker associated with each of the traits is important as spinach breeders can use these markers to select respective traits after validation in their genetic stocks.

Recently, RNA-seq analysis between resistant and susceptible cultivars inoculated with *P. effusa* at two different time-points provided comparative gene expression profiles during compatible and incompatible spinach-*P. effusa* interactions identified potential genes associated with resistance and provided insight into the molecular mechanism of spinach-*P. effusa* interactions[134]. Other new RNA-seq and proteomics experiments have been performed to understand the molecular mechanism of downy mildew resistance in NIL1[135] and NIL6[136] compared to susceptible cultivar Viroflay. These new transcriptomic and proteomic results are expected to provide infection stage-specific expression patterns and novel insights into resistance and susceptibility interactions. In addition, the seed spine locus *Fs* was mapped following BSA to a 0.27 Mb region of Spov1 chromosome 3 containing four candidate genes and reported a co-dominant marker co-segregating with the spiny seeds that distinguish the spineless seeds[71]. Many recent investigations and reports set an optimism towards incorporation and incremental use of molecular and genomic techniques to materialize genome enabled breeding effort in developing both abiotic and biotic stress tolerant spinach cultivars more efficiently in terms of time and cost.

Decreasing sequencing costs have made it feasible to perform WGR (coverage of 1−10x) to discover genome-wide variants in segregating or natural spinach populations. The spinach-working group at the University of Arkansas, USDA-Salinas, Texas A&M University, and Washington State University aim to generate a consensus map using multiple spinach biparental populations segregating for disease resistance (downy mildew, white rust, and Fusarium wilt) and map many other traits in the future. The WGR (10x coverage) generated variant data[34,63] will be used to conduct GWAS and genomic selection (GS) to identify causal variants for many important traits and provide insight into understanding the genetic mechanism of phenotypic variation. Markers identified from this coordinated effort may enable adopting marker assisted selection to breed improved spinach cultivars. Overall, the multistate and multistakeholder team aims to develop tools for molecular breeding that will

accelerate the introgression of economically important traits in developing improved spinach cultivars.

## Prospects of spinach breeding and genomics

Due to the high concentration of nutrients and health-promoting compounds in spinach, and an increasing interest in healthier diets, demand for spinach is growing in the United States and worldwide. Thus, developing improved cultivars is necessary to meet the increasing demand and research and breeding activities have also increased accordingly. The major constraints in increasing fresh market spinach production are diseases, particularly downy mildew, Fusarium wilt, white rust, and leaf spot diseases. Spinach, as well as containing beneficial nutrients, also contains a high amount of oxalic acid[137], affecting both taste and human health. This acid forms insoluble salt combined with calcium, magnesium, zinc, and iron[138]. Oxalic acid and calcium together form calcium oxalate crystals depositing to kidney stones. Additionally, oxalate gives a gritty taste to spinach. Spinach also contains a substantial amount of nitrate associated with lowering blood pressure and lowering the incidence of cardiovascular disease[139]. Similarly, spinach uptakes a higher amount of cadmium when grown in an area with high soil cadmium content, posing a risk to human health[69].

Spinach is a popular vegetable crop that can benefit from developing molecular tools to improve breeding efficiency. Attraction towards organic production practices that comprises half of the total production implies that the spinach breeding program should focus on breeding and selection for organic cultivation. Organic crop cultivation needs specific traits to adapt to the organic system as the varieties bred for conventional methods often do not perform well in the organic system[140]. Another important agenda in spinach is to develop cultivars with low oxalate content. Several past studies have shown a broad range of oxalate content in spinach accessions[137,141] and desirable alleles contributing to low oxalate content have been reported[34,126]. Thus, an opportunity to use the genetic resources of spinach to breed cultivars with reduced oxalate content exists. Similarly, efforts have been sought to identify low cadmium accumulating germplasm lines, alleles, and genes associated with low cadmium content and develop spinach cultivar with reduced cadmium content[69]. There is opportunity to evaluate genetic variation on health-promoting compounds among the available genetic resources soon to breed nutrient-dense spinach cultivars.

Management strategies to control downy mildew disease include treatment with chemical fungicide, cultural practices including growing several cultivars containing different resistance genes, minimizing leaf wetness and humidity, crop rotation for two to three years, and use of host genetic resistance. Genetic resistance seems to be the most effective disease management approach, which is likely the only option for organic production. Downy mildew-resistant cultivars are developed by regular introgression of resistant genes from wild plants. Breeders use the wild germplasm to search for new resistance genes and introgress them to keep with the novel virulent isolates. Commercial spinach hybrid cultivars have a narrow genetic base which make them vulnerable to pathogen attack and prone to abiotic stresses. The regular emergence of new pathogen strains, despite planting a mixture of cultivars containing different R genes, indicates a need to identify a sustainable strategy of using genetic resistances to manage the disease. Identifying additional resistance sources against the known *P. effusa* races is the best approach to control downy mildew by developing durably resistant cultivars via major R gene stacking and possibly utilizing the R gene on a quantitative resistance background. Field tolerance to downy mildew has been demonstrated[120,121], and combining both major genes and QTL in commercial cultivars and hybrids may potentially provide more durable disease resistance. Such cultivars may not be overcome by a new pathogen race and may help to minimize the pathogen population. However, the challenge is the gaps in our knowledge of genetics and molecular aspects of host-pathogen interactions for qualitative and quantitative resistance. Molecular and functional understanding of plant-pathogen interactions may provide a foundation to formulate new strategic interventions to minimize crop loss from pathogens. Further studies on available resistance sources will elucidate the inheritance of resistance locus, enable mapping of R genes, and identify tightly linked markers suitable for the marker-assisted selection approach. Molecular characterization of the interactions between the host and pathogen races and disease development has been prioritized in spinach. Many biparental and multi-parental populations and association mapping panels will be evaluated in the coming years, and a detailed list of genes involved in the resistance mechanism will be identified. A detailed fine mapping for all major genes providing resistance against different pathogen races and identifying and mapping the major and minor QTLs will provide options for breeding durable downy mildew resistant cultivars. Indeed, efforts should be sought to isolate genes, explore defense signaling networks, and elucidate the molecular mechanism of host-pathogen interaction, which may allow resistance breeding using multiple approaches. Alternatively, recent research focuses on identifying susceptibility genes[142,143] that could be an option to develop resistant cultivars following a loss-of-function of the susceptibility factors.

Climate change increases the surface air temperature and increased environmental stresses as droughts, floods, and high and low-temperature stress are expected to exacerbate in coming years/ decades[144–146]. Such climatic stresses will impact the sustainable production and food security of horticultural crops like spinach, as most elite cultivars are vulnerable to severe climate change effects. National and international consortiums should design a strategic plan to minimize the adverse impact and maximize sustainable production, foreseeing the possible climate change events. Further, identification of genetic loci providing tolerance to possible adverse climatic scenarios of high temperature, drought, and salinity should be prioritized as expanded understanding of the physiological mechanism and genetic control of abiotic and biotic stresses may help select and breed climate-resilient cultivars.

The breeding strategies and schemes to utilize new genomic tools in spinach improvement are shown (Fig. 3).

There are opportunities to explore the genetic diversity of wild and cultivated *Spinacia* species and phenotype variations, which remains the fundamental basis for improving spinach. A concerted effort has been sought to sequence a large panel of germplasm accessions and commercial cultivars to generate whole-genome sequences and identify genome-wide genetic variants to dissect the many important traits at a sequence level resolution and linking the variants to gene function[63]. The molecular understanding may be further supported using proteomics and metabolomics approaches[147,148] to elucidate the association between genetic and phenotypic variations to improve crop nutritional quality and understand biotic and abiotic stresses. In recent years, pangenome studies have been gaining importance on crop communities to identify core genes (present in all members of the panel) and variable genes (present or absent in some accessions)[149–151]. Such variability represents species diversity and can present genes enriched among the species members with beneficial traits. The presence and absence of genes among the accessions may provide an insight into gene functions to support spinach improvement programs. The GBS and low-coverage whole genome sequencing approaches have been used for genotyping populations segregating for major traits (downy mildew, white rust, Fusarium wilt) and precisely map the important phenotypes. Further, the ongoing mapping efforts are aimed to generate a consensus genetic map of spinach. However, the parents used in the crosses are not inbreds, and segregating F2 populations are generated by sister-crossing of the resultant F1s that often show a lack of fit to the 1:2:1 expected segregation ratios. Such unfit segregations do not permit QTL analyses; alternatively, association mapping can be used to map the traits from such populations as performed in previous studies [42,116].

In the last decade, the adoption of next-generation sequencing has accelerated genetic studies and generated substantial molecular biological information and generated genomic resources. Several publications have reported SNP markers associated with major horticultural traits in spinach, and many phenotypic evaluations and GWAS/QTL studies are ongoing. The current progress on identifying accessions with favorable traits and the practical ability to mine corresponding alleles associated with the traits, has laid the foundation of molecular and genomics-assisted breeding in spinach. Overall, cataloging the beneficial alleles, genes, and haplotypes should be prioritized to assemble the optimal allelic combination of desirable phenotypes at high precision[152]. The automated high-throughput phenotyping approaches[153,154] can be highly effective and precise in quantifying disease resistances and other traits and may further contribute to efficient and effective gene mapping and allele discovery. Initiations are made to establish the use of high-throughput phenotyping tools in spinach[131]. It is thus imperative to develop standardized tools to assist with rapid phenotyping of the spinach population to quantify morphological and physiological response parameters and screen for yield, disease resistance, and tolerances to environmental stresses.

New cultivar development may take advantage of genome-enabled tools (marker-assisted selection, genomic selection,

haplotype-based selection) to expedite the selection process. Availability of a wealth of sequence resources generated in spinach and the decreasing cost of genotyping and sequencing is accelerating the adoption of genomic assisted selection and will possibly shorten the time to develop elite spinach cultivars. A more detailed investigation on identifying candidate genes and developing a breeder-friendly diagnostic marker assay on target gene sequence is necessary to increase the efficiency and precision of developing new cultivars. There is also a need to identify genes and pathways to improve our understanding of their role and regulation in providing resistance and control of other important traits to enable precise and efficient improvement of commercially important traits in spinach empowered by the genome-enabled breeding platforms. With the availability of new genomic resources and high-resolution mapping results, the challenge remains to translate this knowledge and resources into breeding and crop improvement.

## Conclusions

Spinach is a popular low calorie vegetable crop with an excellent source of vitamins, proteins, and flavonoids. The availability of genomic resources, including the reference genome sequences and ongoing sequencing effort for many of the worldwide spinach germplasm collections, have promoted fundamental research, including the QTL mapping and identifying molecular markers, association mapping, molecular breeding, comparative genomics, and comparative transcriptomic and proteome characterization between healthy and stress conditions that will provide new biological pieces of information and adds genomic resources for spinach. In addition, the development of molecular tools benefits the spinach community by improving breeding efficiency.

## Conflict of interest

The authors declare that they have no conflict of interest.

## REFERENCES

1.  FAOSTAT. 2019. Production domain. In Crops. FAO, Rome https://www.fao.org/faostat/en/#data/QCL
2.  USDA-NASS. 2019. Vegetables 2018 summary. United States Department of Agriculture National Agricultural Statistics Service, Wahington, DC.

3. Cao G, Sofic E, Prior RL. 1996. Antioxidant capacity of tea and common vegetables. *Journal of Agricultural and Food Chemistry* 44:3426−31

4. Howard LR, Pandjaitan N, Morelock T, Gil MI. 2002. Antioxidant capacity and phenolic content of spinach as affected by genetics and growing season. *Journal of Agricultural and Food Chemistry* 50:5891−96

5. USDA-ARS. 2020. FoodData Central, 2020.

6. Bunea A, Andjelkovic M, Socaciu C, Bobis O, Neacsu M, et al. 2008. Total and individual carotenoids and phenolic acids content in fresh, refrigerated and processed spinach (*Spinacia oleracea* L.). *Food Chemistry* 108:649−56

7. Yosefi Z, Tabaraki R, Gharneh HAA, Mehrabi AA. 2010. Variation in antioxidant activity, total phenolics, and nitrate in spinach. *International Journal of Vegetable Science* 16:233−42

8. Ko SH, Park JH, Kim SY, Lee SW, Chun SS, et al. 2014. Antioxidant effects of spinach (*Spinacia oleracea* L.) supplementation in hyperlipidemic rats. *Preventive Nutrition and Food Science* 19:19−26

9. Correll JC, Bluhm BH, Feng C, Lamour K, Toit LJ, et al. 2011. Spinach: better management of downy mildew and white rust through genomics. *European Journal of Plant Pathology* 129:193−205

10. Ribera A, Bai Y, Wolters AMA, Treuren R, Kik C. 2020. A review on the genetic resources, domestication and breeding history of spinach (*Spinacia oleracea* L.). *Euphytica* 216:48

11. Kadereit G, Mavrodiev EV, Zacharias EH, Sukhorukov AP. 2010. Molecular phylogeny of Atripliceae (Chenopodioideae, Chenopodiaceae): Implications for systematics, biogeography, flower and fruit evolution, and the origin of $C_4$ photosynthesis. *American Journal of Botany* 97:1664−87

12. Fuentes-Bazan S, Mansion G, Borsch T. 2012. Towards a species level tree of the globally diverse genus *Chenopodium* (Chenopodiaceae). *Molecular Phylogenetics and Evolution* 62:359−74

13. Stevens P. 2020. Angiosperm Phylogeny Website. Version 14, July 2017. http://www.mobot.org/MOBOT/research/APweb/welcome.html

14. Arumuganathan K, Earle ED. 1991. Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* 9:208−18

15. Morelock TE, Correll JC. 2007. Spinach. In Vegetables I. Handbook of Plant Breeding, eds. Prohens J, Nuez F. New York: Springer. pp 189−218 https://doi.org/10.1007/978-0-387-30443-4_6

16. Koike ST, Cahn M, Cantwell M, Fennimore S, Lestrange M, et al. 2011. *Spinach Production in California*. Publication Number: 7212. University of California, Agriculture and Natural Resources. https://doi.org/10.3733/ucanr.7212

17. Khattak JZK, Torp AM, Andersen SB. 2006. A genetic linkage map of *Spinacia oleracea* and localization of a sex determination locus. *Euphytica* 148:311−18

18. Janick J, Stevenson EC. 1955. Genetics of the monoecious character in spinach. *Genetics* 40:429−37

19. Iizuka M, Janick J. 1971. Sex chromosome variation in *Spinacia oleracea* L. *Journal of Heredity* 62:349−52

20. Yamamoto K, Oda Y, Haseda A, Fujito S, Mikami T. 2014. Molecular evidence that the genes for dioecism and monoecism in *Spinacia oleracea* L. are located at different loci in a chromosomal region. *Heredity* 112:317−24

21. Okazaki Y, Takahata S, Hirakawa H, Suzuki Y, Onodera Y, et al. 2019. Molecular evidence for recent divergence of X- and Y-linked gene pairs in *Spinacia oleracea* L. *PLoS One* 14:e0214949

22. She H, Xu Z, Zhang H, Li G, Wu J, et al. 2021. Identification of a male-specific region (MSR) in *Spinacia oleracea*. *Horticultural Plant Journal* 7:341−46

23. Komai F, Masuda K. 2004. Plasticity in sex expression of spinach (Spinacia oleracea) regenerated from root tissues. *Plant Cell, Tissue and Organ Culture* 78:285−87

24. Janick J, Stevenson EC. 1955. Environmental influences on sex expression in monoecious lines of spinach. *Proceeding Am. Soc. Hortic. Sci.* 75:416−22

25. Thompson AE. 1955. Methods of producing first-generation hybrid seed in spinach. *Cornell Agric. Exp. Stn. Mem.* 336:1−48

26. Chailakhyan MK, Khrianin VN. 1987. Hormonal regulation of sex expression and age-related changes. In *Sexuality in Plants and Its Hormonal Regulation*, ed. Thimann KV. NY: Springer, New York. pp. 33−58 https://doi.org/10.1007/978-1-4612-4734-0_3

27. Rubatzky VE, Yamaguchi M. 1997. Spinach, table beets, and other vegetable Chenopods. In *World Vegetables*. MA: Springer, Boston. pp. 457−73 https://doi.org/10.1007/978-1-4615-6015-9_21

28. Andersen SB, Torp AM. 2011. Spinacia. In *Wild Crop Relatives: Genomic and Breeding Resources*, ed. Kole C. Heidelberg: Springer, Berlin. pp. 273−76 https://doi.org/10.1007/978-3-642-20450-0_13

29. Sneep J. 1983. The domestication of spinach and the breeding history of its varieties. *Euphytica* Supplement 2:1−27

30. Xu C, Jiao C, Sun H, Cai X, Wang X, et al. 2017. Draft genome of spinach and transcriptome diversity of 120 *Spinacia* accessions. *Nature Communications* 8:15275

31. Shi A, Qin J, Mou B, Correll J, Weng Y, et al. 2017. Genetic diversity and population structure analysis of spinach by single-nucleotide polymorphisms identified through genotyping-by-sequencing. *PLoS One* 12:e0188745

32. Fujito S, Takahata S, Suzuki R, Hoshino Y, Ohmido N, et al. 2015. Evidence for a common origin of homomorphic and heteromorphic sex chromosomes in distinct Spinacia species. *G3 Genes|Genomes|Genetics* 5:1663−73

33. Gyawali S, Bhattarai G, Shi A, et al. 2021. Genetic diversity and population structure of *Spinacia turkestanica*, a wild progenitor of cultivated spinach, *Spinacia oleracea*. *HortScience* 56:S200. (Abstr.)

34. Cai X, Sun X, Xu C, Sun H, Wang X, et al. 2021. Reference genome and resequencing of 305 accessions provide insights into spinach evolution, domestication and genetic basis of agronomic traits. *bioRxiv* In Press

35. Bhattarai G, Shi A, Kandel DR, Solís-Gracia N, da Silva JA, et al. 2021. Genome-wide simple sequence repeats (SSR) markers discovered from whole-genome sequence comparisons of multiple spinach accessions. *Scientific Reports* 11:9999

36. Ribera A, van Treuren R, Kik C, Bai Y, Wolters AMA, et al. 2021. On the origin and dispersal of cultivated spinach (*Spinacia oleracea* L.). *Genetic Resources and Crop Evolution* 68:1023−32

37. Meyer RS, Purugganan MD. 2013. Evolution of crop species: genetics of domestication and diversification. *Nature Reviews Genetics* 14:840−52

38. Yang Z, Li G, Tieman D, Zhu G. 2019. Genomics approaches to domestication studies of horticultural crops. *Horticultural Plant Journal* 5:240−46

39. Matsumura H, Hsiao MC, Lin YP, Toyoda A, Taniai N, et al. 2020. Long-read bitter gourd (*Momordica charantia*) genome and the genomic architecture of nonclassic domestication. *Proceedings of the National Academy of Sciences of the United States of America* 117:14543−51

40. Wei T, van Treuren R, Liu X, Zhang Z, Chen J, et al. 2021. Whole-genome resequencing of 445 *Lactuca* accessions reveals the domestication history of cultivated lettuce. *Nature Genetics* 53:752−60

41. van Treuren R, de Groot L, Hisoriev H, Khassanov F, Farzaliyev V, et al. 2020. Acquisition and regeneration of *Spinacia turkestanica* Iljin and *S. tetrandra* Steven ex M. Bieb. to improve a spinach gene bank collection. *Genetic Resources and Crop Evolution* 67:549−59

42. Bhattarai G, Shi A, Feng C, Dhillon B, Mou B, et al. 2020. Genome wide association studies in multiple spinach breeding populations refine downy mildew race 13 resistance genes. *Frontiers in Plant Science* 11:563187

43. Khattak JZK, Christiansen JL, Torp AM, Andersen SB. 2007. Genic microsatellite markers for discrimination of spinach cultivars. *Plant Breeding* 126:454−56

44. Hu J, Mou B, Vick BA. 2007. Genetic diversity of 38 spinach (*Spinacia oleracea* L.) germplasm accessions and 10 commercial hybrids assessed by TRAP markers. *Genetic Resources and Crop Evolution* 54:1667−74

45. Kuwahara K, Suzuki R, Ito Y, Mikami T, Onodera Y. 2014. An analysis of genetic differentiation and geographical variation of spinach germplasm using SSR markers. *Plant Genetic Resources* 12:185−90

46. Xu C, Jiao C, Zheng Y, Sun H, Liu W, et al. 2015. *De novo* and comparative transcriptome analysis of cultivated and wild spinach. *Scientific Reports* 5:17706

47. Whisson DL, Steele Scott N. 1985. Nuclear and mitochondrial DNA have sequence homology with a chloroplast gene. *Plant Molecular Biology* 4:267−273

48. Timmis JN, Scott NS. 1983. Sequence homology between spinach nuclear and chloroplast genomes. *Nature* 305:65−67

49. Scott NS, Timmis JN. 1984. Homologies between nuclear and plastid DNA in spinach. *Theoretical and Applied Genetics* 67:279−88

50. Stren DB, Palmer JD. 1986. Tripartite mitochondrial genome of spinach: Physical structure, mitochondrial gene mapping, and locations of transposed chloroplast DNA sequences. *Nucleic Acids Research* 14:5651−66

51. Schmitz-Linneweber C, Maier RM, Alcaraz JP, Cottet A, Herrmann RG, et al. 2001. The plastid chromosome of spinach (*Spinacia oleracea*): complete nucleotide sequence and gene organization. *Plant Molecular Biology* 45:307−15

52. Ito M, Ohmido N, Akiyama Y, Fukui K, Koba T. 2000. Characterization of spinach chromosomes by condensation patterns and physical mapping of 5S and 45S rDNAs by FISH. *Journal of the American Society for Horticultural Science* 125:59−62

53. Groben R, Wricke G. 1998. Occurrence of microsatellites in spinach sequences from computer databases and development of polymorphic SSR markers. *Plant Breeding* 117:271−274

54. Feng C, Bluhm BH, Correll JC. 2015. Construction of a spinach bacterial artificial chromosome (BAC) library as a resource for gene identification and marker development. *Plant Molecular Biology Reporter* 33:1996−2005

55. Göl Ş, Göktay M, Allmer J, Doğanlar S, Frary A. 2017. Newly developed SSR markers reveal genetic diversity and geographical clustering in spinach (*Spinacia oleracea*). *Molecular Genetics and Genomics* 292:847−855

56. Li S, Wang B, Guo Y, Deng C, Gao W. 2018. Genome-wide characterization of microsatellites and genetic diversity assessment of spinach in the chinese germplasm collection. *Breeding Science* 68:455−464

57. Dohm JC, Minoche AE, Holtgräwe D, Capella-Gutiérrez S, Zakrzewski F, et al. 2014. The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*). *Nature* 505:546−549

58. Minoche AE, Dohm JC, Schneider J, Holtgräwe D, Viehöver P, et al. 2015. Exploiting single-molecule transcript sequencing for eukaryotic gene prediction. *Genome Biology* 16:184

59. Irish BM, Correll JC, Feng C, Bentley T, de los Reyes BG. 2008. Characterization of a resistance locus (*Pfs-1*) to the spinach downy mildew pathogen (*Peronospora farinosa* f. sp. *spinaciae*) and development of a molecular marker linked to *Pfs-1*. *Phytopathology* 98:894−900

60. Collins K, Zhao K, Jiao C, Xu C, Cai X, et al. 2019. SpinachBase: A central portal for spinach genomics. *Database* 2019:baz072

61. Hulse-Kemp AM, Bostan H, Chen S, Ashrafi H, Stoffel K, et al. 2021. An anchored chromosome-scale genome assembly of spinach improves annotation and reveals extensive gene rearrangements in euasterids. *The Plant Genome* 14:e20101

62. Hirakawa H, Toyoda A, Itoh T, Suzuki Y, Nagano AJ, et al. 2021. A spinach genome assembly with remarkable completeness, and its use for rapid identification of candidate genes for agronomic traits. *DNA Research* 28:dsab004

63. Shi A, Correll J, Feng C, et al. 2019. Progress at developing genetic and molecular resources to improve spinach production and management. *HortScience* 54:S137. (Abstr.)

64. Sneep J. 1958. The present position of spinach breeding. *Euphytica* 7:1−8

65. Janick J. 2015. Hybrids in Horticultural Crops. In *Concepts and Breeding of Heterosis in Crop Plants*, eds. Larnkey KR, Staub JE. Madison, Wisconsin: Crop Science Society of America, pp. 45−56 https://doi.org/10.2135/cssaspecpub25.c4

66. Toit L. 2018. Spinach seed production in the Pacific Northwest. *2018 International Spinach Conference*, Murcia, Spain https://spinach.uark.edu/spain-presentations/

67. McLeod P, Morelock TE, Goode MJ. 1991. Preference, developmental time, adult longevity and fecundity of green peach aphid (Homoptera: *Aphididae*) on spinach. *Journal of Entomological Science* 26:95−98

68. Mou B. 2008. Leafminer resistance in spinach. *HortScience* 43:1716−19

69. Greenhut RF. 2018. *Developing baby leaf spinach with reduced cadmium accumulation*. MS Thesis. University of California, Davis.

70. Knepper C, Mou B. 2014. Initial Evaluation of Spinach Collections for Drought Tolerance. *HortScience* 49:S295. (Abstr.)

71. Liu Z, Lu T, Feng C, Zhang H, Xu Z. 2021. Fine mapping and molecular marker development of the *Fs* gene controlling fruit spines in spinach (*Spinacia oleracea* L.). *Theoretical and Applied Genetics* 134:1319−28

72. Mou B. 2019. 'USDA Red' spinach. *HortScience* 54:2070−72

73. Correll JC, Morelock TE, Black MC, Brandenberger LP, Dainello FJ. 1994. Economically important diseases of spinach. *Plant Disease* 78:653−60

74. Smith PG, Webb RE, Luhn CH. 1962. Immunity to race 2 of Spinach downy mildew. *Phytopathology* 52:597−99

75. Zink FW, Smith PG. 1958. A second physiologic race of spinach downy mildew. *Plant Disease Report* 42

76. Smith PG, Webb RE, Millett AM, Luhn CH. 1961. Downy mildew on spinach: A second race of fungus has been found on Califlay variety in the coastal valley area of California. *California Agriculture* 15:5

77. Eenink AE. 1976. Resistance in spinach to downy mildew. *In Proceedings of Eucarpia Meeting. Leafy Vegetables*. Wageningen, Holland. pp. 53−54

78. Jones RK. 1982. Occurrence of race 3 of Peronospora effusa on spinach in Texas and identification of sources of resistance. *Plant Disease* 66:1078−79

79. Brandenberger LP, Correll JC, Morelock TE . 1991. Identification of and cultivar reactions to a new race (race 4) of *Peronospora farinosa* f. sp. *spinaciae* on spinach in the United States. *Plant Disease* 75:630−34

80. Irish BM, Correll JC, Koike ST, Schafer J, Morelock TE. 2003. Identification and cultivar reaction to three new races of the spinach downy mildew pathogen from the United States and Europe. *Plant Disease* 87:567−72

81. Irish BM, Correll JC, Koike ST, Morelock TE. 2007. Three new races of the spinach downy mildew pathogen identified by a modified set of spinach differentials. *Plant Disease* 91:1392−96

82. Plantum. 2018. Denomination of Pfs: 17, a new race of downy mildew in spinach. http://www.naktuinbouw.com/about-naktuinbouw/news/denomination-pfs-17-new-race-downy-mildew-spinach. Accessed 10 Feb 2021

83. Feng C, Saito K, Liu B, Manley A, Kammeijer K, et al. 2018. New races and novel strains of the spinach downy mildew pathogen *Peronospora effusa*. *Plant Disease* 102:613−18

84. Feng C, Correll JC, Kammeijer KE, Koike ST. 2014. Identification of new races and deviating strains of the spinach downy mildew pathogen *Peronospora farinosa* f. sp. *spinaciae*. *Plant Disease* 98:145−52

85. Plantum. 2021. Denomination of Pe : 18 and 19, two new races of downy mildew in spinach. https://plantum.nl/denomination-of-pe-18-and-19-two-new-races-of-downy-mildew-in-spinach/. Accessed 25 May 2021

86. Lyon R, Correll J, Feng C, Bluhm B, Shrestha S, et al. 2016. Population structure of *Peronospora effusa* in the southwestern United States. *PLoS One* 11:e0148385

87. Kandel SL, Mou B, Shishkoff N, Shi A, Subbarao KV, et al. 2019. Spinach downy mildew: Advances in our understanding of the disease cycle and prospects for disease management. *Plant Disease* 103:791−803

88. Bhattarai G, Feng C, Dhillon B, Shi A, Villarroel-Zeballos M, et al. 2020. Detached leaf inoculation assay for evaluating resistance to the spinach downy mildew pathogen. *European Journal of Plant Pathology* 158:511−20

89. Dhillon B, Feng C, Villarroel-Zeballos MI, et al. 2020. Sporangiospore viability and oospore production in the spinach downy mildew pathogen, Peronospora effusa. *Plant Disease* 104:2634−41

90. Brandenberger LP, Correll JC, Morelock TE, McNew RW. 1994. Characterization of resistance of spinach to white rust (*Albugo occidentalis*) and downy mildew (*Peronospora farinosa* f. sp. *spinaciae*). *Phytopathology* 84:431−37

91. Sullivan MJ, Damicone JP, Payton ME. 2002. The effects of temperature and wetness period on the development of spinach white rust. *Plant Disease* 86:753−58

92. Leskovar DI, Kolenda K. 2002. Strobilurin + acibenzolar-S-methyl controls white rust without inducing leaf chlorosis in spinach. *Annals of Applied Biology* 140:171−75

93. Vakalounakis DJ, Doulis AG. 2013. First record of white rust, caused by *Albugo occidentalis*, on spinach in Greece. *Plant Disease* 97:1253

94. Correll JC, Feng CD, Liu B. 2017. First report of white rust (*Albugo occidentalis*) of spinach in mexico. *Plant Disease* 101:511

95. Soylu S, Kara M, Ku rt, Uysal A, Shin HD, et al. 2018. First report of white blister rust disease caused by Albugo occidentalis on Spinach in Turkey. *Plant Disease* 102:826

96. du Toit LJ. 2004. Management of diseases in seed crops. In *Encyclopedia of Plant and Crop Science*, ed. Goodman RM. New York: Routledge. pp. 675−77 http://doi.org/10.1081/E-EPCS120019947

97. Gatch EW, du Toit LJ. 2017. Limestone-mediated suppression of Fusarium wilt in spinach seed crops. *Plant Disease* 101:81−94

98. Gyawali S, du Toit LJ, Shi A, Correll JC. 2019. Genome wide association studies of Fusarium wilt resistance in spinach (*Spinacia oleracea* L.). *Phytopathology* 109:S2.83 (Abstr.)

99. du Toit LJ, Derie ML, Hernandez-Perez P. 2005. Verticillium wilt in spinach seed production. *Plant Disease* 89:4−11

100. Liu B, Stein L, Cochran K, du Toit LJ, Feng C, et al. 2020. Characterization of leaf spot pathogens from several spinach production areas in the United States. *Plant Disease* 104:1994−2004

101. du Toit LJ, Derie ML. 2001. 2001. *Stemphylium botryosum* pathogenic on spinach seed crops in Washington. *Plant Disease* 85:920

102. Hernandez-Perez P, Du Toit LJ. 2006. Seedborne *Cladosporium variabile* and *Stemphylium botryosum* in spinach. *Plant Dis* 90:137−45

103. Mou B, Koike ST, Du Toit LJ. 2008. Screening for resistance to leaf spot diseases of spinach. *HortScience* 43:1706−10

104. Shi A, Mou B, Correll J, Koike ST, Motes D, et al. 2016. Association analysis and identification of SNP markers for Stemphylium leaf spot (*Stemphylium botryosum* f. sp. *spinacia*) resistance in spinach (*Spinacia oleracea*). *American Journal of Plant Sciences* 07:1600−11

105. Liu B, Bhattarai G, Shi A, et al. 2020. Evaluation of resistance of USDA spinach germplasm to *Stemphylium vesicarium*. *Phytopathology* 110:S2.178 (Abstr.)

106. Bhattarai G, Shi A, Avila CA, Stein LA. 2020. GWAS on USDA Spinach Germplasm Evaluated for Downy Mildew in the Texas' Wintergarden Production Area. 55:S227. (Abstr.)

107. Bates ML. 1984. Root rot of hydroponically grown spinach caused by *Pythium aphanidermatum* and *P. dissotocum*. *Plant Disease* 68:989−91

108. Gold SE, Stanghellini ME. 1985. Effects of temperature on *Pythium* root rot of spinach grown under hydroponic conditions. *Phytopathology* 75:333−37

109. Liu B, Feng C, Matheron ME, Correll JC. 2018. Characterization of foliar web blight of spinach, caused by *Pythium aphanidermatum*, in the desert southwest of the United States. *Plant Disease* 102:608−12

110. Onodera Y, Yonaha I, Masumo H, Tanaka A, Niikura S, et al. 2011. Mapping of the genes for dioecism and monoecism in *Spinacia oleracea* L.: Evidence that both genes are closely linked. *Plant Cell Reports* 30:965−71

111. Chan-Navarrete R, Dolstra O, van Kaauwen M, Lammerts van Bueren ET, Linden CG, et al. 2016. Genetic map construction and QTL analysis of nitrogen use efficiency in spinach (*Spinacia oleracea* L.). *Euphytica* 208:621−36

112. Fujita N, Ayukawa Y, Fuke M, eraoka T, Watanabe K, et al. 2017. Rapid sex identification method of spinach (*Spinacia oleracea* L.) in the vegetative stage using loop-mediated isothermal amplification. *Planta* 245:221−26

113. Qian W, Fan G, Liu D, Zhang H, Wang X, et al. 2017. Construction of a high-density genetic map and the *X/Y* sex-determining gene mapping in spinach based on large-scale markers developed by specific-locus amplified fragment sequencing (SLAF-seq). *BMC Genomics* 18:1

114. Feng C, Bluhm B, Shi A, Correll JC. 2018. Development of molecular markers linked to three spinach downy mildew resistance loci. *Euphytica* 214:174

115. She H, Qian W, Zhang H, Liu Z, Wang X, et al. 2018. Fine mapping and candidate gene screening of the downy mildew resistance gene RPF1 in Spinach. *Theoretical and Applied Genetics* 131:2529−2541

116. Bhattarai G, Yang W, Shi A, Feng C, Dhillon B, et al. 2021. High resolution mapping and candidate gene identification of downy mildew race 16 resistance in spinach. *BMC Genomics* 22:478

117. Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW. 2003. Genome-wide analysis of NBS-LRR-encoding genes in arabidopsis. *The Plant Cell* 15:809−34

118. Michelmore RW, Meyers BC. 1998. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Research* 8:1113−30

119. Hulbert SH, Webb CA, Smith SM, Sun Q. 2001. Resistance gene complexes: evolution and utilization. *Annual Review of Phytopathology* 39:285−312

120. Bhattarai G. 2019. *Genetic resistance to the downy mildew pathogen and mapping the RPF resistance loci in spinach*. Doctoral Dissertation. University of Arkansas, Fayetteville, USA. https://scholarworks.uark.edu/etd/3442

121. Bhattarai G, Shi A, Correll JC, Feng C. 2020. Field evaluation and genome-wide association analysis of downy mildew resistance in spinach. *HortScience* 55:S227. (Abstr.)

122. Olaoye D. 2021. *Resistance Screening and Association Mapping for Resistance to the Downy Mildew Pathogen of Spinach*. Graduate Theses and Dissertations. University of Arkansas, Fayetteville, USA. https://scholarworks.uark.edu/etd/4107

123. Ma J, Shi A, Mou B, Evans M, Clark J, et al. 2016. Association mapping of leaf traits in spinach (*Spinacia oleracea* L.). *Plant Breeding* 135:399−404

124. Chitwood J, Shi A, Mou B, Evans M, Clark J, et al. 2016. Population structure and association analysis of bolting, plant height, and leaf erectness in spinach. *HortScience* 51:481−86

125. Shi A, Mou B. 2016. Genetic diversity and association analysis of leafminer (*Liriomyza langei*) resistance in spinach (*Spinacia oleracea*). *Genome* 59:581−88

126. Shi A, Mou B, Correll JC. 2016. Association analysis for oxalate concentration in spinach. *Euphytica* 212:17−28

127. Shi A, Mou B, Correll J, Motes D, Weng Y, et al. 2016. SNP association analysis of resistance to Verticillium wilt (*Verticillium dahliae* Kleb.) in spinach. *Australian Journal of Crop Science* 10:1188−96

128. Qin J, Shi A, Mou B, Grusak MA, Weng Y, et al. 2017. Genetic diversity and association mapping of mineral element concentrations in spinach leaves. *BMC Genomics* 18:941

129. Awika HO, Cochran K, Joshi V, Bedre R, Mandadi KK, et al. 2020. Single-marker and haplotype-based association analysis of anthracnose (*Colletotrichum dematium*) resistance in spinach (*Spinacia oleracea*). *Plant Breeding* 139:402−18

130. Awika HO, Marconi TG, Bedre R, Mandadi KK, Avila CA. 2019. Minor alleles are associated with white rust (*Albugo occidentalis*) susceptibility in spinach (*Spinacia oleracea*). *Horticulture Research* 6:129

131. Awika HO, Bedre R, Yeom J, Marconi TG, Enciso J, et al. 2019. Developing growth-associated molecular markers via high-throughput phenotyping in spinach. *Plant Genome* 12:190027

132. Kunz DDR, Awika HO, Avila CA. 2020. Vitamin C content in spinach germplasm: phenotypic diversity and genome-wide association mapping. *HortScience* 55:S351. (Abstr.)

133. Gyawali S, Bhattarai G, du Toit LJ, et al. 2021. Genome wide association studies of Fusarium wilt resistance in wild spinach, *Spinacia turkestanica*. In American Phytopathological Society Annual Meeting 2021 https://www.apsnet.org/meetings/annual/PlantHealth2021/Pages/default.aspx

134. Kandel SL, Hulse-Kemp AM, Stoffel K, Koike ST, Shi A, et al. 2020. Transcriptional analyses of differential cultivars during resistant and susceptible interactions with *Peronospora effusa*, the causal agent of spinach downy mildew. *Scientific Reports* 10:6719

135. Zia B. 2021. *Genetic Resistance to the Downy Mildew Pathogen and Breeding towards Durable Disease Management in Spinach*. Doctoral Dissertation. University of Arkansas, Fayetteville, USA.

136. Bhattarai G, Shi A. 2021. Comparative transcriptome and proteome analysis of resistant and susceptible spinach in response to downy mildew pathogen. *HortScience* 56:S103. (Abstr.)

137. Mou B. 2008. Evaluation of oxalate concentration in the U.S. spinach germplasm collection. *HortScience* 43:1690−93

138. Bsc SN. 1999. Oxalate content of foods and its effect on humans. *Asia Pacific Journal of Clinical Nutrition* 8:64−74

139. Jonvik KL, Nyakayiru J, Pinckaers PJ, Senden JMG, van Loon LJC, et al. 2016. Nitrate-rich vegetables increase plasma nitrate and nitrite concentrations and lower blood pressure in healthy adults. *The Journal of Nutrition* 146:986−93

140. Murphy KM, Campbell KG, Lyon SR, Jones SS. 2007. Evidence of varietal adaptation to organic farming systems. *Field Crops Research* 102:172−77

141. Kaminishi A, Kita N. 2006. Seasonal change of nitrate and oxalate concentration in relation to the growth rate of spinach cultivars. *HortScience* 41:1589−95

142. Bai Y, Pavan S, Zheng Z, Zappel NF, Reinstädler A, et al. 2008. Naturally occurring broad-spectrum powdery mildew resistance in a Central American tomato accession is caused by loss of *Mlo* function. *Molecular Plant-Microbe Interactions* 21:30−39

143. Pessina S, Lenzi L, Perazzolli M, Campa M, Dalla Costa L, et al. 2016. Knockdown of MLO genes reduces susceptibility to powdery mildew in grapevine. *Horticulture Research* 3:16016

144. Raza A, Razzaq A, Mehmood SS, Zou X, Zhang X, et al. 2019. Impact of climate change on crops adaptation and strategies to tackle its outcome: a review. *Plants* 8:34

145. Gornall J, Betts R, Burke E, Clark R, Camp J, et al. 2010. Implications of climate change for agricultural productivity in the early twenty-first century. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* 365:2973−89

146. Battisti DS, Naylor RL. 2009. Historical warnings of future food insecurity with unprecedented seasonal heat. *Science* 323:240−44

147. Kumar R, Bohra A, Pandey AK, Pandey MK, Kumar A. 2017. Metabolomics for plant improvement: Status and prospects. *Frontiers in Plant Science* 8:1302

148. Fernie AR, Schauer N. 2009. Metabolomics-assisted breeding: a viable option for crop improvement? *Trends in Genetics* 25:39−48

149. Zhao J, Bayer PE, Ruperao P, Saxena RK, Khan AW, et al. 2020. Trait associations in the pangenome of pigeon pea (*Cajanus cajan*). *Plant Biotechnology Journal* 18:1946−54

150. Gao L, Gonda I, Sun H, Ma Q, Bao K, et al. 2019. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nature Genetics* 51:1044−51

151. Bayer PE, Golicz AA, Scheben A, Batley J, Edwards D. 2020. Plant pan-genomes are the new reference. *Nature Plants* 6:914−20

152. Varshney RK, Bohra A, Yu J, Graner A, Zhang Q, et al. 2021. Designing future crops: genomics-assisted breeding comes of age. *Trends in Plant Science* 26:631−49

153. Araus JL, Kefauver SC, Zaman-Allah M, Olsen MS, Cairns JE. 2018. Translating high-throughput phenotyping into genetic gain. *Trends in Plant Science* 23:451−66

154. Yang W, Feng H, Zhang X, Zhang J, Doonan JH, et al. 2020. Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives. *Mol Plant* 13:187−214