Review

Ensuring Global Food Security by Improving Protein Content in Major Grain Legumes Using Breeding and ‘Omics’ Tools

Uday C. Jha 1,*, Harsh Nayyar 2, Swarup K. Parida 3, Rupesh Deshmukh 4*, Eric J. B. von Wettberg 5 and Kadambot H. M. Siddique 6,*

1 ICAR—Indian Institute of Pulses Research (IIPR), Kanpur 208024, India
2 Department of Botany, Panjab University, Chandigarh 160014, India; harshnayyar@hotmail.com
3 National Institute of Plant Genome Research, New Delhi 110067, India; swarup@nigpr.ac.in
4 National Agri-Food Biotechnology Institute, Punjab 140308, India; rupesh@nabi.res.in
5 Department of Plant and Soil Science, The University of Vermont, Burlington, VT 05405, USA; eric.bishop-von-wettberg@uvm.edu
6 The UWA Institute of Agriculture, The University of Western Australia, Perth, WA 6001, Australia
* Correspondence: u9811981@gmail.com (U.C.J.); kadambot.siddique@uwa.edu.au (K.H.M.S.)

Abstract: Grain legumes are a rich source of dietary protein for millions of people globally and thus a key driver for securing global food security. Legume plant-based ‘dietary protein’ biofortification is an economic strategy for alleviating the menace of rising malnutrition-related problems and hidden hunger. Malnutrition from protein deficiency is predominant in human populations with an insufficient daily intake of animal protein/dietary protein due to economic limitations, especially in developing countries. Therefore, enhancing grain legume protein content will help eradicate protein-related malnutrition problems in low-income and underprivileged countries. Here, we review the exploitable genetic variability for grain protein content in various major grain legumes for improving the protein content of high-yielding, low-protein genotypes. We highlight classical genetics-based inheritance of protein content in various legumes and discuss advances in molecular marker technology that have enabled us to underpin various quantitative trait loci controlling seed protein content (SPC) in biparental-based mapping populations and genome-wide association studies. We also review the progress of functional genomics in deciphering the underlying candidate gene(s) controlling SPC in various grain legumes and the role of proteomics and metabolomics in shedding light on the accumulation of various novel proteins and metabolites in high-protein legume genotypes. Lastly, we detail the scope of genomic selection, high-throughput phenotyping, emerging genome editing tools, and speed breeding protocols for enhancing SPC in grain legumes to achieve legume-based dietary protein security and thus reduce the global hunger risk.

Keywords: grain legume; protein; biofortification; molecular marker; QTL

1. Introduction

Alarming trends of anthropogenic climate change and environmental deterioration jeopardize global crop yields, resource distribution, and ecosystems, resulting in global food insecurity and undernourishment in the growing human population [1]. An estimated 840 million people globally will be undernourished by 2030 [2]. The COVID-19 pandemic will have compounded this figure, increasing the food-related hunger crisis. Dietary protein is an essential macronutrient for human growth and development, with infants requiring 1.52 g per kg body weight per day and adults recommended 0.80 g per kg body weight per day [3]. Apart from micronutrient deficiency, malnutrition from dietary protein deficiency causes ‘marasmus’, ‘kwashiorkor’ anemia, impaired immunity, and ‘environmental enteric dysfunction,’ most prevalent in developing and low-income countries, especially southern Asia and sub-Saharan Africa [4–6]. Most of the people residing in these regions predominantly consume maize, sorghum, and cassava in their daily diets, which are rich in...
starch but insufficient in protein [6,7]. Thus, many people, especially infants, inhabiting these regions do not consume the required daily protein, affecting their overall growth and development [5,6]. Notably, Europe imports 70% of the plant-based protein consumed by its human population [8], a trend that the increasing global human population will further exacerbate.

Breeding crops, especially legumes, with high-quality traits such as SPC is a promising approach for overcoming these challenges. Grain legumes are one of the richest sources of plant-based dietary protein, providing essential amino acids and supplying the increasing demand for protein-based human diets [9]. Grain legume seeds, popularly known as ‘poor man’s meat’, are the cheapest protein source [10–12]. In addition, legume-based protein could be instrumental in minimizing greenhouse gas emissions, helping to protect the environment [13]. Screening genetic variability for protein content in various legume germplasm and crop wild relatives is the first step to identifying high-protein grain legumes for the development of high-yielding, high-protein legumes. A classical genetics-based approach could identify the inheritance pattern of high-protein gene(s) in various legumes. Advances in genomics have enabled the dissection of the genetic architecture of QTLs/gene(s) in various legumes through biparental mapping and genome-wide association studies. Moreover, the availability of complete reference genome assemblies and pangenomes of various legumes could assist in underpinning high-protein genomic regions at the individual or species level. Likewise, advances in functional genomics have enabled the discovery of various candidate genes that improve legume protein content and their precise function. Proteomics and metabolomics can improve our understanding of various complex pathways, molecular networks, and metabolites underlying high-protein grain legumes. Non-destructive phenomics approaches could be instrumental for screening and identifying high-protein lines with high efficiency. Emerging technologies such as genomic selection, rapid generation advancement, and genome editing could be harnessed to improve SPC, eradicate malnutrition related to dietary protein deficiency, and meet the United Nations Sustainable Developmental Goal 2.

2. Grain Legumes as an Important Source of Dietary Protein

Grain legumes vary in their protein content, due to fundamental limitations on the components a seed must contain to be viable. Many grains legumes have 25–40% SPC, and it may be difficult to raise that number much beyond 40%. (See Table 1).

Table 1. Seed protein contents and deficient amino acids in major grain legumes.

| Crop       | Scientific Name | Range of Grain Seed Protein Content | References | Deficient Amino Acids                  |
|------------|-----------------|------------------------------------|------------|---------------------------------------|
| Chickpea   | *Cicer eretinum* L. | 17–22% before dehulling            | [14,15]    | Methionine, cysteine, threonine and valine [16] |
|            |                 | 25.3–28.9% after dehulling         |            |                                       |
| Lentil     | *Lens culinaris* Medik | 20.6% and 31.4%                    | [17]       | Methionine, cysteine [18]             |
| Lupin      | *Lupinus albus* L. | 35–44%                             | [19,20]    | Alanine, tryptophan [21]              |
Soybean 

Glycine max (L.) Merr. up to 40% [22,23] Methionine, cysteine, threonine, and lysine [24]

Common bean

Phaseolus vulgaris L. 20–30% [25,26] Methionine, cysteine [27]

Pigeonpea

Cajanus cajan (L.) Millsp. 20–22% [28] Methionine, cysteine, valine [29]

Faba bean

Vicia faba L. 26% to 41% [30,31] Methionine

Mung bean

Vigna radiata L. 20.97–31.32% [32] Methionine, cysteine

Cowpea

Vigna unguiculata L. Walp.) 14.8–25% [33–35] Methionine

Pea

Pisum sativum L. 13.7 to 30.7% [36–38] Methionine, cysteine and tryptophan [39] 37–39

Urd bean

Vigna mungo L. Hepper 25–28% [40,41] Methionine, cysteine

Lathyrus

Lathyrus sativus L. 8.6–34.6% [42] Methionine, cysteine

Chickpea (Cicer eretinum L.) SPC ranges from 17 to 22% before dehulling and 25.3 to 28.9% after dehulling [14,15] (see Table 1). Chickpea seed contains two main proteins—
globulin (11S legumin and 7S vicilin) and albumin—with low amounts of glutelins and prolamine; however, the seed is deficient in cysteine and methionine amino acids [43,44]. Despite rich sources of various essential amino acids, cysteine, methionine, valine, and threonine are the major limiting amino acids in chickpea [45]. Desi type chickpea has higher SPC than the kabuli type but no differences in essential amino acids [45].

Common bean (*Phaseolus vulgaris* L.) SPC ranges from 20 to 30% [25,26], and plays a pivotal role in mitigating protein-related malnutrition, especially in underdeveloped countries [46,47]. The major storage protein in common bean is phaseolin, accounting for 36–46% of total seed proteins [48], with 50–60% of the phaseolin belonging to the 7S vicilin class, insufficient in methionine, cysteine, and tryptophan essential amino acids [49–51].

Cowpea (*Vigna unguiculata* L. Walp.) is a ‘multi-functional’ grain legume widely used for human consumption. It helps mitigate the challenges of malnutrition in sub-Saharan Africa, and tropical and sub-tropical regions globally [52,53]. Cowpea SPC ranges from 15 to 25% [33,34] (see Table 1). Cowpea storage proteins are abundant in lysine and tryptophan but deficient in methionine and cysteine [53]. Globulins are the most abundant storage protein fraction of cowpea grain, followed by albumins, glutelins, and prolamin [54].

Faba bean (*Vicia faba* L.) SPC ranges from 26 to 41% [30,31,55,56], with abundant essential amino acids except for tryptophan, cysteine, and methionine [57]. More than 80% of the seed proteins comprise globulins (vicilin and legumin) [55]. Of the essential amino acids, faba bean seed is highest in lysine [58].

Lentil (*Lens culinaris* Medik) SPC ranges from 20 to 30% [59]. Like other legumes, lentil seed has a high globulin content (44–70% of storage protein, constituting 11S legumin and 7S vicilin and convicilin) and albumin (26–61% of lentil proteins) but low prolamin and glutelin levels [60,61].

White lupin (*Lupinus albus* L.) seeds are a rich reservoir of protein containing up to 44% [19,62], with two major classes of protein—albumin (15%) and globulin (85%) [63]. The globulin protein comprises α-, β-, γ-, and δ-conglutins [20]. Despite some allergenic effects in white lupin seed protein, they are low in antinutritive properties compared with other grain legumes such as pea and soybean [62,64]. Moreover, white lupin seed contains higher amounts of some important amino acids (lysine, phenylalanine, arginine, and leucine) than soybean, rendering it a high-demand grain legume from a nutritional point of view [65].

Soybean (*Glycine max* (L.) Merr.) is rich in protein, ranging from 35 to 45%. It is deficient in methionine [22,23] but has sufficient lysine to overcome the lysine deficiency of cereals [66]. In 2018, it was estimated that soybean alone contributed 70% of the global protein meal [67].

Mung bean (*Vigna radiata* L.) contains easily digestible protein and several essential micronutrients [68]. It is an excellent source of protein except for sulfur-containing amino acids (methionine and cysteine) [69]. Due to its ease of digestibility relative to other legumes [70] and low hypoallergenic properties, mung bean is used as a weaning food for infants [71]. Moreover, mungbean is a good meat substitute for vegetarians and those who cannot afford animal-based dietary protein [12].

Pea (*Pisum sativum* L.) is rich in protein, ranging from 13.7 to 30.7% [37]. Pea seed protein comprises legumin, vicilin, convicilin, and globulin-related proteins [37]. Vicilin is the most abundant protein (26.3–52.0% of total pea protein extract) [37]. Moreover, pea protein is in high demand in food industries due to its gluten-free quality and low allergenicity [72].

Pigeon pea (*Cajanus cajan* (L.) Millsp) seeds contain 20–22% protein and play an essential role in providing plant-based dietary protein to the vegetarian population in India, thus ensuring protein-based food security [73].

Urd bean (*Vigna mungo* L. Hepper) is another important grain legume rich in protein (up to 25%), comprising globulin (63%), albumin (12%), and glutelin (21%) [74]. Urd bean seeds are rich in glutamic acid, aspartic acid, and lysine but deficient in methionine and cysteine [74].
3. Harnessing Genetic Variability for Improving Seed Protein Content in Grain Legumes

Harnessing crop germplasm diversity is an economical way to improve important breeding traits, including SPC in grain legume crops [75–79]. Crop genetic resources are the key reservoir for exploring high-SPC genotypes in grain legumes. Considerable amounts of genetic variability for SPC have been captured in chickpea [78,80,81], such as 12.4–31.5% [82], 17–22% [83], and 14.6–23.2% [84]. Serrano et al. [84] identified several high SPC genotypes (LEGCA608, LEGCA609, LEGCA614, LEGCA619, LEGCA716) that could be used to improve chickpea SPC in elite cultivars.

Cowpea is a cheap source of protein for improving human nutrition. Boukar et al. [77] assessed a set of 1541 cowpea lines for genetic variability in grain protein content and mineral profiles. They reported a wide range of genetic variability for SPC (17.5–32.5%), including TVu-2508 (32.2%) [77]. Likewise, Weng et al. [85] screened 173 cowpea accessions collected from various parts of the world at two locations (Fayetteville and Alma, Arkansas). They also reported a substantial amount of genetic variability for SPC (22.8–28.9%), including PI 662992 (28.9%), PI 601085 (28.5%), PI 255765 (28.4%), PI 255774 (28.4%), and PI 666253 (28.4%) [85], which could be used to transfer the high SPC trait into high-yielding elite cowpea varieties. The nutritional profiles (including grain protein content) of 22 cowpea genotypes collected from various regions of eastern, southern, and western Africa were evaluated at two locations in South Africa [34]. Seed protein contents, measured using the combustion method, ranged from 23.16 to 28.13% [34]. The authors noted significant positive correlations between SPC and various mineral contents, indicating the possibility of simultaneously selecting these traits. Among the tested genotypes, 98K-5301 had high Ca and SPC [34]. Similarly, an evaluation of 21 cowpea genotypes identified high SPC in COVU-702 (27.7%) and HC-98-64 (27.9%) [86]. In another study, GonÇalves et al. [87] identified high SPC in Paulistinha (29.2%) among 18 tested cowpea genotypes. An evaluation of 30 Brazilian cowpea lines for protein, vitamin, and mineral content identified high SPC in MNC01-649F-2 (28.3%), BRŞ-Cauamé (27.8%), BRŞ-Paraguacu (27.7%), BRŞ-Marataoa (27.4%), Canapuzinho (25.0%), BRŞ-Tumucumaque (24.8%), and MNC01-631F-15 (24.6%) [88].

The SPC of selected common bean landraces ranged from 16.54 to 25.23%, while selected modern common bean cultivars ranged from 19.70 to 24.30% (Celmeli et al. [79]; see Table 2).

Table 2. List of various legume genotypes with improved seed protein content.

| Crop | Genotypes | Seed Protein Content | Source | References |
|------|------------|----------------------|--------|------------|
| Chickpea | ICC 5912 | 29.2% | ICRISAT, Patancheru, India | [78] |
| Common bean | LEGCA608, LEGCA609, LEGCA614, LEGCA619, LEGCA716 | >22% | Cordoba | [84] |
| Common bean | J-216, FJJP-43 | 222 (J/L-146) to 330 (J-216); 180 (G11027A) to 311 (FJJP-43) g kg⁻¹ | Mexican state of Jalisco and Durango | [89] |
| Common bean | LR05 | 25.23% | Food Safety and Agricultural Research Center, Akdeniz University | [79] |
| Common bean | 6-EX Accession 4049 | 23% | Santo Antônio de Goiás, Brazil | [90] |
| Cowpea | HC-6, HC-5, CP-21, LST-II-C-12, CP-16, COVU-702, HC-98-64 | 26.7–27.9% | India | [86] |
| Cowpea | TVu-2723, TVu-3658, TVu-2508 | 32.50% | Minjibir, Kano State, Nigeria | [77] |
| Cowpea | MNCO1-649F-2, BRŞ-Cauamé, BRŞ-Paraguacu, BRŞ-Marataoa | 27.4–28.3% | – | [88] |
| Cowpea | “Early Scarlet” and 09-204 | 26.9–27.4% | Arkansas State (Fayetteville, Alma, Hope) | [92] |
| Bengpl | 40% | Dokpong and Bamahu near Wa, Ghana, South Africa, Taung | | [93] |
| Bengpl | Pf662992, Pf601085, PI255765, PI255774, PI666253 | 28.4–28.9% | Tanzania, South Africa, Nigeria | [85] |
| Bengpl | Vuli, Mamlaka, IT90K-59, Ngoji, TVU13953, 98K-5301 | | South Africa, Nigeria | [34] |
Grasspea is an inherent climate-resilient grain legume with an excellent source of SPC. An evaluation of 37 grasspea genotypes identified IC127616 rich in SPC (32.2%) [95].

An analysis of 27 local mung bean landraces using the micro-Kjeldahl method identified significant genetic variability for SPC (17.2–29.9%), with the highest values in MGG30 (29.9%), NAGPURI (29.3%), and BSN1 (27.8%) [97]. Moreover, significant genetic variability for SPC (15.2–21%) rich in lysine, tryptophan, valine, leucine, isoleucine and phenylalanine amino acids was noted in *Vigna radiata var sublobata*, a wild species of mung bean [113].

Genetic variability for SPC in lentil ranges from 20 to 30% [76,114–117]. Likewise, lentil crop wild relatives (CWRs) have significant genetic variability for SPC, such as *L. orientalis* (18.3–27.75%) and *L. ervoides* (18.9–32.7%) [96], which could be used in breeding programs to improve SPC in elite lentil cultivars.

Pea has considerable genetic variability for SPC (19.3–25.2%) (Gottschalk et al. [75]; see Table 2). Wang and Daun [82] reported a SPC ranging from 201.6 to 266.6 g kg\(^{-1}\) DM in four elite pea cultivars. In large sets of pea accessions, the SPC was 22–32% [98] and 23–32% [118] using the Kjeldahl technique and 20.6–27.3% [119], 18.6–27.3% [120], 17–27% [121], 17.5–27.8% [122], and 19.3–30.3% [123] using the near-infrared technique. Several promising pea genotypes with high SPC have been identified: CDC Striker (up

| Crop | Genotypes | Seed Protein Content | Source | References |
|------|-----------|----------------------|--------|-----------|
| Grasspea | IC127616 | 32.20% | India | [95] |
| Lentil | *L. orientalis* | 18.3–27.75% | India, IIPR, Kanpur | [96] |
| | *L. ervoides* | 18.9–32.7% | India, IIPR, Kanpur | [96] |
| Mungbean | MGG330, Nagpuri | 29.9% and 29.3% | India | [97] |
| | Pea | >30% | Manitoba and Ontario, Canada | [98] |
| | Majoret | 240.4 g kg\(^{-1}\) | Institute of Field and Vegetable Crops (Smederevska Palanka, Serbia) | [72] |
| | NGB 101293 Jordan | 26.80% | – | [37] |
| | L1 | 317.63 g kg\(^{-1}\) | – | [101] |
| Soybean | D76-8070 Jordan | 450 g kg\(^{-1}\) | Central Experimental Farm (Ottawa, ON) Canada | [101] |
| | BARC-6, BARC-7, BARC-8, BARC-9 | – | Tennessee Agricultural Experiment Station, Tennessee, USA, USDA–ARS and the North Carolina Agricultural Research Service | [102] |
| | AC Proteus | – | Yeongnam Agricultural Research Institute (YARI), Milyang, Republic of Korea | [104] |
| | Lines developed from Kwangan- kong × Samnamkong and Danbaekkong × Samnam-kong | 34.3–44.4% and 35.8–49.6% | Yeongnam Agricultural Research Institute (YARI), Milyang, Republic of Korea | [104] |
| | 17D derived population and M23 derived lines | 53% | – | [105] |
| | High-pro 1’ developed from Wyandot × GASF98-114 | 401 g kg\(^{-1}\) | University of Missouri Fisher Delta Research Center, Portageville, MO USDA Agricultural Service and Ohio Agricultural Research and Development Centre Wooster | [107] |
| | ‘TN11-5102’ | 421 g kg\(^{-1}\) protein on a dry weight basis | University of Tennessee Agricultural Research | [108] |
| | PI407228 | 392.6–481.7 g kg\(^{-1}\) | Central Crops Research Station in Clayton, NC, Bradford Farm in Columbia, Sandhills Research Station in Jackson Springs, NC USDA Agricultural Service and Ohio Agricultural Research and Development Centre Wooster | [109] |
| | R11-7999 | 439 g kg\(^{-1}\) (dry weight) | Arkansas Agricultural Experiment Station | [110] |
| | Bioagro | – | – | [111] |
| | SI6-5540GT | 41.10% | University of Missouri–Fisher Delta Soybean Breeding Program | [112] |

Wang and Daun [82] reported a SPC ranging from 201.6 to 266.6 g kg\(^{-1}\) DM in four elite pea cultivars. In large sets of pea accessions, the SPC was 22–32% [98] and 23–32% [118] using the Kjeldahl technique and 20.6–27.3% [119], 18.6–27.3% [120, 17–27% [121], 17.5–27.8% [122], and 19.3–30.3% [123] using the near-infrared technique. Several promising pea genotypes with high SPC have been identified: CDC Striker (up
to 27.8%) [124–127], Ballet (up to 25.9%) [119,128], Solara (28.8%) [129], Caméor (29.9%), VavD265 (27.5%), and China (32%) [128].

Breeding for high SPC in soybean is a primary objective in soybean breeding programs; however, progress has been limited by the negative relationship between SPC and grain yield and oil content [24,130]. For example, Bandillo et al. [131] and Warrington et al. [132] reported a highly negative correlation between the soybean SPC allele and seed oil content, reducing oil content by 1% for every 2% increase in SPC.

High-protein soybean lines include Danbaegkong (48.9%) [133] and Kwangankong (44.7%) [134], and TN11-5102 selected from 5601T cultivar (421 g kg\(^{-1}\) protein on a dry weight basis) [108]. Apart from cultivated species, soybean CWRs (e.g., Glycine soja) are an important source of high-protein QTLs [135–137]. A population developed by incorporating exotic soybean germplasm exhibited significant genetic variability for SPC [138]. Wehrmann et al. [139] and Wilcox and Cavins [140] backcrossed the high-protein trait from Pando into Cutler 71, a high-yielding low-protein genotype soybean. Later, Cober and Voldeng [101] attempted to transfer the high-protein trait from AC Proteus to Maple Glen; however, the selected progenies exhibited higher protein content than Maple Glen but no yield advantage. Sathia, Seti, Kavre, and Soida Chiny soybean cultivars, collected in Nepal, had high SPC (up to 42–45%) compared to William 82 (39%) and higher arginine (5–10%) content than William 82 (7.4%) [141].

Hence, harnessing the available genetic variability for SPC requires the large-scale screening of land races, CWRs, and grain legume germplasm locked in gene banks across the globe.

### 4. Mendelian Inheritance of Seed Protein Content in Legumes

Several researchers have worked out the genetics of SPC based on Mendelian genetics in various grain legumes [142–144]. Considering pea storage proteins (legumin and convicilin), Matta and Gatehouse [145] mapped the *legumin* gene (*Lg-1*), behaving as a single Mendelian gene with five alleles on LG7, and the *convicilin* gene (*Cvc*), behaving as a single Mendelian gene on LG2 using seeds developed from 1238 × 1263, 110 × 807 and 110 × 851 F\(_2\) crosses. Subsequently, Mahmoud and Gatehouse [146] explained the monogenic inheritance of another pea SPC vicilin (*Vc-1*) gene controlled by two codominant genes located on LG7 using an F\(_2\) cross from 360 × 611.

Perez et al. [147] revealed the genetic basis of high and low SPC in pea using the genetics of seed size (round vs. wrinkled). They found that round-seeded pea plants (RR/RbRb) had low SPC with low albumin content, while those with recessive alleles (rr/rbrb) had high SPC and high albumin content [147]. High heritability of protein content and its control by a few gene(s) is an opportunity to improve protein content in cowpea [92]. Moreover, diallel crosses of six populations derived from two high-protein lines and two high-yielding soybean lines revealed a significant negative correlation between protein content and yield in the high protein × high protein population but a significant positive correlation between protein content and yield in the high yielding × high yielding population [148]. In pigeon pea, an analysis of F\(_1\) and F\(_2\) progenies derived from crosses involving four parents revealed a minimum of 3–4 genes controlling protein content [149]. The authors concluded that the low protein trait is partially dominant over the high protein trait.

Various studies have reported a significant effect of environment on SPC [150–152]. In soybean, this significant effect involved multiple genes and the quantitative nature of the SPC trait [150,151]. In chickpea, an F\(_2\) segregating population developed from ICC5912 (blue flowered) × ICC17109 (white flowered) revealed the quantitative nature of the SPC trait and its high negative correlation with seed yield and seed size [78]. A 5 × 5 half diallel cross of cowpea lines revealed the presence of additive and non-additive gene effects for SPC. High seed albumin, prolamin, and globulin were associated with positive effects of the dominant gene, while high SPC and glutelin content were associated with recessive genes [153]. In lentil, Kumar et al. [154] also reported the quantitative nature of the SPC trait.
High genetic variation in lentil seed storage protein resulted from high G × E interactions exhibiting moderate heritability (31.3%) [152].

5. QTL Mapping for Seed Protein Content

Advances in grain legume genomics have facilitated the identification of underlying QTLs controlling SPC using biparental mapping populations in various grain legumes [118,119,155–157].

Few studies have uncovered QTLs controlling SPC in chickpea. However, one study that phenotyped recombinant inbred lines (RILs) derived from ICC995 × ICC5912 across four environments and used a genotyping by sequencing approach delineated one major effect QTL q-3.2 for SPC that explained 44.3% of the phenotypic variation (PV) on LG3 [158].

In pea, using an F2-derived Wt10245 × Wt11238 mapping population, Irzykowska and Wolko [159] mapped five QTLs governing SPC on LG2, LG3, and LG7, explaining 13.1–25.8% PV. Subsequently, two F2 mapping populations developed from Wt11238 × Wt3557 and Wt10245 × Wt11238 revealed a QTL for protein content on LGVb flanked by cp, gp, and te markers [118]. Likewise, genotyping an Orb × CDC Striker RIL mapping population with SNP markers identified two SPC QTLs on LG1b, explaining 16% PV, and two on LG4a, explaining 10.2% PV, and genotyping a Carerra × CDC Striker RIL-based mapping population identified four SPC QTLs on LG7b, explaining 13% PV, and one on LG3b [160].

An evaluation of a Terese × K586 RIL population in five different environments identified 14 SPC QTLs located on LGI, LGIII, LGIV, LGV, LGVI, and LGVII [119]. The study identified the underlying candidate gene for the QTL on LGI as the Rgp gene (cell wall synthesis) and two underlying candidate genes for the QTL on LGV as Ls (GA biosynthesis) and Rbcs4 (encoding small Rubisco subunit) [119].

Obala et al. [157] gained insight into the genetic determinants controlling SPC in pigeonpea based on the results obtained from five F2 populations segregating for SPC (ICP 11605 × ICP 14209, ICP 8863 × ICP 11605, HPL 24 × ICP 11605, ICP 8863 × ICPL 87119, and ICP 5529 × ICP 11605). Fourteen major effect QTLs explaining 23.5% PV were found to be located on CcLG02, CcLG03, CcLG06 and CcLG11 [157].

In soybean, the SPC trait is controlled by multiple alleles and highly influenced by G × E interactions [150]. More than 300 QTLs contributing to SPC in soybean have been reported (http://www.soybase.org, (accessed on 10 May 2022)); [161] and reside across all chromosomes; however, major SPC QTLs are on chromosomes 5, 15, and 20. Diers et al. [155] first reported a major QTL governing high SPC on chromosome 20 in a population developed from crossing cultivated and wild soybean, which was later mapped to a 3 cM on LGI (Nichols et al., 2006) [156]. The location of this QTL was subsequently narrowed to 8.4 Mb [162], <1 MB [163], 77.4 kb [137], and even with only three candidate genes [131] on LG20. Likewise, another major SPC QTL, qSeedPro_15, was narrowed to 4 Mb (Zhang et al. [164]; see Table 3), overlapping the previously identified genomic region on chromosome 15 [24,131,155,165,166]. Zhang et al. [164] elucidated a possible candidate gene Glyma.15G049200 underlying the QTL. Genotyping recombinant inbred lines derived from the interspecific cross of Williams 82 × G. soja (PI 483460B) using Illumina Infinium BeadChip sequencing platform identified five SPC QTLs, mapped on chromosomes 6, 8, 13, 19, and 20, explaining 4.6–19.6% PV [167]. Of these identified QTLs, qPro_20 QTL was stable across the four tested environments.

SSR, DArT, and DArTseq analysis of five RIL-based mapping populations for high and low SPC and one high × high SPC identified two major QTLs controlling SPC on LG15 and LG20 in soybean [168]. Furthermore, bulk segregation analysis of four high × low SPC mapping populations unveiled novel SPC-controlling genomic regions on LG1, 8, 9, 14, 16, 17, 19, and 20 [168]. An assessment of soybean RILs developed from Linhafenqingdou × Meng 8206 in six different environments identified 25 SPC QTLs explaining up to 26.2% PV [169]. Of the identified QTLs, qPro-7-1 was highly stable across all tested environments.

Recently, Fliege et al. [137] cloned a major SPC governing QTL (cqSeed protein-003) and
elucidated the underlying causative candidate gene *Glyma.20G85100*, encoding a CCT domain protein. Thus, efforts are needed to fine map or clone major QTLs controlling SPC in other grain legumes to delineate the underlying candidate gene(s) and their function for genomic-assisted breeding to improve SPC in grain legumes.

### Table 3. List of seed protein content QTLs reported in various grain legumes.

| Crop                  | Mapping Population/Panel of Genotypes | QTL/Gene | Marker | LG    | PV%    | References |
|-----------------------|--------------------------------------|----------|--------|-------|--------|------------|
| Chickpea              | GWAS, 187                            | 4 QTLs 9 significant MTAs 6 candidate genes p<3.2 | SSR | LG3, 5 | 2.4-5.1 | [81]       |
|                        | GWAS, 336                            |          | SNP    |       |        |            |
|                        | ICC 995 × ICC3192, RIL (189)          |          | SNP    | LG3   | 44.3   | [158]      |
| Common bean           | Xana × Cornell 49242, RIL (104)       | SpA, SpB, SpE, Spf, Pha, SpF, SpG, Spk, SpM, SpC, SpD | AFLP, RAPD, SSR, SCAR | LG7, 4, 3, 1 |        | [171]      |
|                        | Xana × Cornell 49242, RIL (104)       | One QTL  | SSR    |       |        | [172]      |
| Ground nut            | TG26 × GPBD 4, RIL (146)              | 8 QTLs   | SSR    | LG1, 3, 4, 7, 8 | 1.5-10.7 | [173]      |
| Pea                   | 1238 × 1263, 110 × 807, 110 × 851 (F2) | Convicillin (Cv), Legumin (Lg-1) | Protein marker | LG2, 7 |        | [145]      |
|                        | 360 × 611 (F2)                        |          |        |       |        |            |
|                        | Wt10245 × Wt11238, F2 (114)          | prot1 prot2 prot3 prot4 prot5 | ABLP, RAPD, SSR, STS, CAPS | LG2, 5, 7 | 13.1-25.5 | [150]      |
|                        | Terese’ × Ks86, RIL (139)            | 14 QTLs  |        |       |        | [119]      |
|                        | Wt11238 × Wt13557 (FS), Wt10245 × Wt11238 (FS) | One QTL | ABLP, RAPD, STS, CAPS, SSR | LG7a, 5b |        | [118]      |
|                        | GWAS, 50                             | One significant SNP | SNP | LG1b, 4a | 16 | [160]      |
|                        | Orb × CDC Striker, Carrera × CDC Striker | 8 QTLs | SNP | LG3b, 7b |        | [157]      |
|                        | 1-2347-144 × CDC Meadow GWAS, 135 genotypes | 21 QTL |          |        |        | [160]      |
|                        | GWAS, 135                            | Chr3LG5_19453076 | SNP |       |        | [174]      |
|                        | 9 populations, RIL (1213)             | Chr3LG5_138235361 | SNP |       |        | [123]      |
|                        | ICP11605 × ICP 14209, ICP 8863 × ICP 11605, HPL 24 × ICP 11605, ICP 8863 × ICP 87119, ICP 5529 × ICP 11606 | 48 M-QTLs for SPC SNP | CccLG03, 11, 02, 06 | 0.7-23.5 | [157]      |
| Pigeonpea             | Parker × PI 468916                    | Two major quantitative trait locus (QTL) alleles | – | LG20 |        | [136]      |
| Soybean               | A3733 × PI 437088A, RIL (76)          | One QTL  | Satt496 and Satt239 | RAPD marker | OPJW13a | LG20 | [175]      |
|                        | Essex × Williams PI 97100 × Coker 237 |          |       |       |        |            |
|                        | N85-984-16 × TN05-99-99               |          |       |       |        |            |
|                        | 4 QTL for cysteine, 3 QTLs for methionine | Satt235, Satt252, Satt427, Satt436, Satt252, Satt564, Satt590 | D1a, 1G, 1M | – | [102]      |
|                        | A81356022 × PI 468916 G. soja (PI468916) × G. max (A81-356022) backcrossing | One QTL | SSR, AFLP | LG1 |        | [162]      |
|                        | G. max A81-356022 × G. soja PI468916, near isogenic lines | One QTL, 13 genes | SNP | LG20 |        | [162]      |
|                        | Magellan × PI 438499B ZYD09454 × Yudou12 |          |       |       |        |            |
|                        | 17 genomic regions, Glyma20g196680, Glyma20g21030, Glyma20g21080, Glyma20g19660, Glyma20g19630, Glyma20g21040 |          |       |       |        |            |
|                        | GWAS, 298                            | 8 QTLs   | SNP    | LG8, 9, 20 |        | [180]      |
|                        | IL-1964 (619 accessions), IL-1966 (977 accessions), MS-1996 (728 accessions), MS-2000 (934 accessions) |          | SNP    | LG20 |        | [163]      |
|                        | ZYD2738 × Jidou 12, F2, ZYD2738 × Jidou 9, F2 | 4 QTLs |        |       |        | [181]      |
|                        | Benning × Danbaekkong RO5-1415 × RO5-638 |            |        |       |        |            |
|                        | GWAS, 139                            | qPC19 and 8 significant genomic regions SNP | SNP | LG14, 15, 17, 20 | 55 | [182]      |
|                        | SD02-4-59 × A02-381100 (RIL), SD02-911 × SD00-1501 (RIL) | 8 QTLs | – | LG5, 8, 10, 14, 16, 19 | 10.3 | [183]      |
|                        | Danbaekkong × Glycine soja (PI468916) | wp allele, cpseed protein-003 | – | LG20 |        | [185]      |
|                        | G. max (Williams 82) × G. soja (PI 483460B) | 5 QTLs qPro_06, qPro_19, qPro_20, qPro_08, qPro_13 | SNP | LG6, 8, 13, 19, 20 | 4.6-19.6 | [167]      |
### Table 3. Cont.

| Crop | Mapping Population/Panel of Genotypes | QTL/Gene | Marker | LG | PV% | References |
|------|--------------------------------------|----------|--------|----|-----|------------|
| GWAS, 144 lines derived from four parents | Glyma.03G100800, Glyma.10G207300, Glyma.12G019300, Glyma.12G112900, Glyma.14G081600, Glyma.18G028600, Glyma.18G071110, Glyma.18G071300 | SNP | LG1, 2, 3, 4, 6, 7, 9, 10, 12, 14, 18 | 3.84-19.21 | [186] |
| – | 192 collinear protein QTLs, 13 candidate genes | – | – | – | [187] |
| Linhefenqingdou × Meng 8206 RIL (104) | 25 main effect QTLs | SNP | LG1, 4, 6, 7, 8, 9, 10, 13, 14, 17, 18, 19, 20 | 5.7-26.22 | [169] |
| GWAS, 621 accessions | Three genomic regions, 16 significant SNPs, Glyma.15g049100, Glyma.15g049200, Glyma.15g050100, Glyma.15g050600 | SNP | LG1, 4, 6, 7, 8, 9, 10, 13, 14, 17, 18, 19, 20 | – | [188] |
| GWAS, 185 | Three significant SNP markers (Kenfeng14 × Kenfeng15) × (Heinong48 × Kenfeng19) RIL (160) | 34 QTLs | SSR | LG12 | 2.65-13.83 | [189] |
| G15FN-12 mutant | – | SoySNP50K BeadChip | – | – | – | [190] |
| GWAS, 249 | 25 significant MTAs | SNP | LG2, 6, 7, 10, 13, 14, 16, 17, 18, 19 | – | [191] |
| AC Proteus Maple Arrow Fs, RIL | X3145-B-B-3-15 × 9063, F5, RIL; X3144-48-1-B-9063, Fs, RIL; X3144-48-1-B × AC Brant, Fs, RIL; X3145-B-B-3-15 × X3144-48-1-B, Fs, RIL | qPro_Gm02–3, qPro_Gm04–4, qPro_Gm06–1, qPro_Gm06–3, qPro_Gm06–6, qPro_Gm13–4, qPro_Gm15–3 | SNP | LG1, 8, 9, 14, 16, 17, 19, 20 | 15.5–60 | [192] |
| AC X790P × S18-R6′ and ‘AC X790P × S23-T5, RILs | 138 significant MTA, Glyma.07g175700 and Glyma.07g176000 | SNP | LG7, 13, 15 | 18-34 | [193] |
| GWAS, 211 | 50 QTLs, three candidate genes: Glyma.20G088000, Glyma.20G111100, Glyma.20p087600 | qPC-7-1, qPC-13-1, qPC-15-1 | SNP | LG3, 6, 13, 20 | – | [194] |
| (Kenfeng14 × Kenfeng15) × (Heinong48 × Kenfeng19) | 16 QTLs | SoySNP6k BeadChip | LG7 | – | – | [195] |
| G00-3213 × PI 594458A, 132 RILs | 13 significant MTA, Glyma.07g175700 and Glyma.07g176000 | SNP | LG1, 2, 4, 5, 6, 8, 12, 13, 15, 18 | 10.4–21.9 | [196] |
| GWAS, 165 | AC X790P × S18-R6′ and ‘AC X790P × S23-T5, RILs | qPro_Gm02–3, qPro_Gm04–4, qPro_Gm06–1, qPro_Gm06–3, qPro_Gm13–4, qPro_Gm15–3 | SNP | LG1, 2, 3, 5, 6, 8, 12, 14, 17, 19, 20 | 15.5-60 | [197] |
| RIL, 944, Primus × Protina, Gallec × Sigalia, Primus × Sigalia, Protina × Sigalia, Gallec × Protina, Gallec × Protina, Gallec × Sigalia, Primus × Sultana | qPY1, qPY2, qPY3, qPY4, qPY5, qPY6, qPY7 | SNP | LG5, 6, 7, 8, 16, 18, 19, 20 | – | – | [198] |
| RIL, 944, Primus × Protina, Gallec × Sigalia, Primus × Sigalia, Protina × Sigalia, Gallec × Protina, Gallec × Protina, Gallec × Sigalia, Primus × Sultana | qPY1, qPY2, qPY3, qPY4, qPY5, qPY6, qPY7 | SNP | LG5, 6, 7, 8, 16, 18, 19, 20 | 15.5-60 | [196] |
| Nanxiadou 25 × Tongdou 11, RIL (178) | 50 QTLs, three candidate genes: Glyma.20G088000, Glyma.20G111100, Glyma.20p087600 | SNP | LG1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 13, 15, 16, 20 | – | – | [197] |
| PI 468916 × A81-356022, BC | Glyma.20G088000, cpefseed protein-003 QTL | SNP | LG20 | – | – | [137] |
| 250, F2 | 3 QTLs | Infinium Soy6KSNP BeadChips | LG6, LG13, LG20 | – | – | [198] |

AFLP = Amplified fragment length polymorphism; SNP = Single nucleotide polymorphism, SCAR = Sequenced cleaved amplified region, CAPS = cleaved amplified polymorphic sequence, RAPD = Random Amplified polymorphic DNA, SSR = Simple Sequence Repeats, ISSR = Inter Simple Sequence Repeat.

### 6. Underpinning Genomic Region/Haplotypes Controlling High Protein Content through GWAS

Traditional biparental QTL mapping for obtaining genetic recombinants controlling complex traits such as protein content is limited due to the incorporation of only two parents in the crossing program. However, the increased capacity of next generation sequencing technology to derive single nucleotide polymorphism molecular markers in association with advanced phenotyping facilities has facilitated the development of numer-
ous genetic recombinants and identification of the underlying plausible candidate genomic regions controlling protein content in various grain legumes using GWAS [81,174,183,186]. Jadhav et al. [81] performed association mapping for SPC using SSR markers on a panel of 187 chickpea genotypes (desi, kabuli, and exotic). Nine significant marker trait associations (MTAs) for SPC were uncovered on LG1, LG2, LG3, LG4, and LG5, explaining 16.85% PV. A recent GWAS using high-throughput SNP markers on 140 chickpea genotypes subjected to drought and heat stress to shed light on MTAs with various nutrients uncovered 66 (non-stress), 46 (drought stress), and 15 (heat stress) MTAs for SPC [199], which could be used to identify high-protein lines for improving SPC in chickpea.

A GWAS relying on multilocation and multi-year phenotyping of a large set of pea germplasm representing diverse regions across the globe was undertaken to identify significant MTAs for agronomic and quality traits, including protein content [174]. Two significant MTAs controlling SPC were identified: Chr3LG5_138253621 and Chr3LG5_194530376. GWAS using 16,376 SNPs in 332 chickpea genotypes (desi and kabuli) delineated seven genomic loci controlling SPC and explaining 41% combined PV [170]. The authors also validated five SPC-controlling genes in a RIL-based mapping population ICC 12299 × ICC 4958, encoding cytidine (CMP), deoxycytidylate (dCMP) deaminases, ATP-dependent RNA helicase DEAD-box, and zinc finger protein. An earlier comprehensive GWAS of 298 soybean lines using Illumina Infinium and GoldenGate assays identified 17 significant genomic regions controlling SPC [180]. Among the SPC-controlling genomic regions, LG20 was important as it contained six candidate genes Glyma20g19680, Glyma20g21030, Glyma20g21080, Glyma20g19630, Glyma20g19620, and Glyma20g21040 in the 2.4 Mbp interval. Another GWAS performed on 139 soybean lines revealed eight significant regions contributing to SPC on LG5, LG8, LG10, LG14, LG16, LG19, and LG20 [183]. In addition, a major QTL qPC19 controlling SPC on LG19 in the 42.3 to 44.2 Mb interval explained 10.3% PV [183]. Likewise, an assay using SoySNP660k BeadChip in 144 soybean lines developed from four-way RILs identified eight candidate genes controlling SPC: Glyma.03G100800, Glyma.10G207300, Glyma.12G019300, Glyma.12G112900, Glyma.14G081600, Glyma.18G028600, Glyma.18G07110, and Glyma.18G071300 (Zhang et al. [186]; see Table 3). A comprehensive GWAS study in a collection of 877 soybean accessions, tested in five different environments in Midwest and southern USA using SoySNP50K iSelect BeadChip [188], identified significant trait-associated genomic regions for SPC that coincided with previous QTL/genomic regions identified on chromosomes 15 and 20 [161,166]. Three SNPs identified within 91 kb overlapped the 118 kb genomic region of meta-QTL controlling SPC and seed oil content previously reported by Van and McHale [161]. Some important candidate genes identified in these genomic regions—Glyma.15g049100 Glyma.15g049200, Glyma.15g049200, Glyma.15g050100, and Glyma.15g050600—participate in partitioning carbon and regulating protein content (Lee et al. [188]; see Table 3). The authors also elucidated eight novel genomic regions controlling methionine, cysteine, lysine, and threonine contents. A GWAS using whole genome sequencing data of 631 soybean accessions combined with a biparental QTL analysis uncovered a pleotropic gene GmSWEET39 (encoding sugar transporter) controlling SPC and seed oil content in soybean [164]. The authors also reported that a 2 bp (CC) deletion in Glyma.15g049200 underlying the GmSWEET39 allele rendered high seed oil content and low SPC.

A comprehensive association and linkage analysis surveyed 985 soybean accessions, including wild species, landraces, and old and modern cultivars, to capture haplotypic variation in the high SPC locus qProt-003 on chromosome 20 [200]. The study uncovered significant trait-associated genomic regions within a 173 kb linkage block containing three causal candidate genes: Glyma.20G084500, Glyma.20G085250, and Glyma.20G085100 [200]. Of these, Glyma.20G085100 (containing a 304 bp deletion and trinucleotide insertions) was tightly linked with the high protein content phenotype [200].
7. Functional Genomics Shedding Light on Causal Candidate Gene(s) Contributing Seed Protein Content in Grain Legumes

In the last decade, unprecedented advances in RNA sequencing have expedited functional genomics research, especially transcriptome analysis for discovering trait gene(s), in various grain legumes [197]. Numerous studies have elucidated various SPC-contributing candidate gene(s) and their functional roles in grain legumes; notably, cDNA cloning based functional characterization of genes encoding storage proteins such as pea seed albumin (PA1, PA1b) [201] and conglutin family in narrow leaf lupin [202]. Functional characterization of genes encoding storage protein in narrow leaf lupin by sequencing cDNA clones from developing seed identified 11 new storage protein (conglutin family)-encoding genes [202]. Transcriptome analysis via RNA-seq shed light on 16 conglutin genes encoding storage protein in the Tanjil cultivar of narrow leaf lupin [203]. Conglutin gene(s) expression is similar in lupin varieties of the same species but distinct between species [203]. In soybean, functional genomic analysis via gene expression profiling identified 329 differentially expressed genes underlying \( qSPC_20–1 \) and \( qSPC_20–2 \) QTL regions accounting for SPC using a QTL-seq approach [197]. Of the nine candidate genes underlying these QTL regions, Glyma.20G088000, Glyma.20G111100, and Glyma.20g087600 were functionally validated and identified as the most potential candidate genes controlling SPC [197]. RNAi technology—a robust functional genomic tool—offered novel insight into the regulatory role of Glyma.20g085100 harboring transposon insertion in the SPC-controlling genomic region of soybean [137]. Reduced expression of Glyma.20g085100 using RNAi enhanced the protein level in the low-protein Thorne soybean genotype [137]. Most functional genomics studies identifying SPC-controlling candidate genes with their putative function in major legumes have involved soybean; thus, studies should focus on elucidating candidate genes and deciphering the molecular mechanism for improving SPC via functional genomics in other grain legumes.

8. Proteomics and Metabolomics Shed Light on the Genetic Basis of High Seed Protein Content in Legumes

Proteomics helps us understand the entire set of proteins produced at a specific time under a particular set of conditions in an organism or cell [204]. This approach could be used to discover novel seed storage proteins and inquire about the molecular basis of enhancing SPC in various legumes [205]. A novel protein known as methionine-rich protein was discovered in soybean using a two-dimensional (2D) electrophoresis technique [205]. Later, a 2D-PAGE proteomic tool distinguished wild soybean (\( G. soja \)) from cultivated soybean based on high storage proteins (beta-conglycinin and glycinin) detecting 44 protein spots in wild soybean and 34 protein spots in cultivated soybean; thus, this helped in identifying high-protein soybean genotypes [206]. Combined SDS-PAGE and MALDI-TOF MS analysis in LG00-13260, PI 427138, and BARC-6 soybean genotypes revealed enhanced accumulation of beta-conglycinin and glycinins and thus high grain protein content compared to William 82 ([207]; see Table 4). A combined SDS-PAGE and MALDI-TOF MS analysis, comparing protein content in nine soybean accessions with William 82, revealed significant protein content differences in seed 11S storage globulins [208]. In common bean, proteome analysis of common bean deficient in seed storage proteins (phaseolin and lectins) revealed elevated sulfur amino acid content due to increased legumin, albumin 2, and defensin [209]. Santos et al. [210] characterized the protein content of 24 chickpea genotypes using a proteomics approach to explore genetic variability in storage protein. High-performance liquid chromatography analysis indicated the presence of sufficient genetic variability for SPC, with some genotypes rich in seven amino acids. In pea, a mature seed proteome map of a diverse set of 156 proteins identified novel storage proteins for enhanced SPC [211].
Table 4. Proteomic approach for investigating novel proteins for improving seed protein content in grain legumes.

| Crop                  | Protein Identified                                           | Approach Used                          | Reference | Genotype            |
|-----------------------|-------------------------------------------------------------|----------------------------------------|-----------|---------------------|
| Chickpea              | High amino acid content, 454 protein spots                  | Two-dimensional electrophoresis and mass spectrometry | [210]     | Flip97-171C, Elite   |
|                       |                                                              | High resolution liquid mass spectrometry |           |                     |
| Common bean           | Sulfur-containing amino acids, S-methylcysteine accumulation | Chromatography-tandem mass spectrometry | [212]     | –                   |
|                       | Sulfur-containing amino acids; enhanced concentration of cysteine and methionine | Mass spectrometry                     | [213]     | SARCl and SMARCIN-PN1 |
| Faba bean             | Amino acid metabolism                                       | iTRAQ                                  | [56]      | Cixidabaican         |
|                       | Legumin, vicilin, and convicilin                             | 1D SDS-PAGE, size-exclusion high-performance liquid chromatography | [214]     | Cartouche, NV657, NV734 |
| Narrow-leafed lupin   | 2760 protein identifications                                | LC-MS                                  | [215]     | P27255, Tandil, Unicrop |
| Pea                   | 156 proteins in Nepalse                                     | 2-D gels, MALDI-TOF MS, MALDI-TOF, two-dimensional gel electrophoresis | [211]     | Nepalse, Karve, Seti |
| Soybean               | High arginine content in Nepalse                            | MALDI-TOF; two-dimensional gel electrophoresis | [141]     | LG00-13260          |
|                       | High beta-conglycinin and glycinins                         | Two-dimensional electrophoresis SDS-PAGE | [207]     |                     |
|                       | High 11S storage globulins                                  | SDS-PAGE, MALDI-TOF, two-dimensional electrophoresis 2D-PAGE | [208]     | PI407788A           |
|                       | High storage protein                                         |                                        | [206]     | Wild soybean         |
|                       | Asparagine, free 3-cyanoalanine, and L-malic acid           | GC-TOF/MS                              | [216]     |                     |

An iTRAQ-based proteomics analysis of CX (low SPC) and LX (high SPC) faba bean genotypes revealed differentially abundant proteins involved in amino acid metabolism [56]. Furthermore, a KEGG analysis suggested that valine, leucine, histidine, and β-alanine metabolism were significantly enriched by differentially abundant proteins [56].

Likewise, metabolomic studies help us understand various metabolic pathways and metabolites controlling protein accumulation during seed development [217]. A meticulous amino acid profiling study using contrasting high and low SPC soybean lines revealed that the ability of embryos to assimilate nitrogen and synthesize storage proteins determines SPC accumulation [217]. Further, the authors reported that high SPC at maturity is related to increased accumulation of asparagine in developing cotyledons.

A metabolomics study using GC-TOF/MS in contrasting seed protein soybean lines showed a high abundance of metabolites (asparagine, aspartic acid, glutamic acid, free 3-cyanoalanine) that were positively associated with SPC and negatively associated with seed oil content [216]. However, various sugars (sucrose, fructose, glucose, mannose) had negative associations with seed protein and oil content [216]. Saboori-Robat et al. [218] undertook metabolite profiling of common bean genotypes differing in S-methylcysteine accumulation in seeds and found that S-methylcysteine accumulates as γ-glutamyl-S-methylcysteine during seed maturation, with a low accumulation of free methylcysteine.

Amino acid profiling of Valle Agricola, a nutritionally rich chickpea genotype cultivated in southern Italy, revealed that 66% of the total amino acids comprised glutamic acid, glutamine, aspartic acid, phenyl alanine, asparagine, lysine, and leucine, while ~40% comprised histidine, valine, isoleucine, leucine, methionine and threonine [219]. Further advances in metabolomics could improve our understanding of various cellular metabolism networks and pathways related to SPC in legumes. Thus, integrating various ‘omics’ tools and emerging novel breeding approaches could assist in developing protein-fortified grain legumes (see Figure 1).
9. Progress of Genetic Engineering and Scope of Genome Editing for Improving SPC in Grain Legumes

Numerous studies have been undertaken to improve the essential amino acid content in various grain legumes by manipulating amino acid encoding genes using genetic engineering [220–222]. Many examples of improved essential amino acid contents, especially sulfur-rich amino acids, by manipulating gene(s) in various legumes using transgenic technology are available. Chiaiese et al. [223] introduced an albumin transgene encoding methionine and cysteine-rich protein from sunflower seed into chickpea to improve seed methionine content. The transgenic chickpea seed accumulated more methionine than the control. Likewise, Molvig et al. [224] improved seed methionine content in narrow leaf lupin by introducing sunflower seed albumin transgene at the transgenic level. However, cysteine-rich storage proteins, especially conglutin delta, declined in narrow leaf lupin seed due to low expression of the cysteine-encoding gene (Tabe and [225]; see Table 5). Introducing Bertholletia excelsa methionine-rich 2S albumin gene into common bean enhanced seed methionine content by more than 20% over non-transgenic plants [220]. Improving sulfur-rich amino acids, such as methionine and cysteine, in soybean has been a research priority, made possible by introducing the 15 kDa [226], 27 kDa [227], and 11 kDa [221,228] δ-zein encoding protein genes from maize using genetic engineering.

Table 5. Selected list of grain legumes with improved seed protein content using a genetic engineering approach.

| Crop                        | Gene Source     | Gene Name                          | Function                                                                 | References          | Transformation Approach         |
|-----------------------------|-----------------|------------------------------------|--------------------------------------------------------------------------|---------------------|---------------------------------|
| Chickpea                    | Sunflower       | Sunflower seed albumin             | Increased methionine up to 90%                                          | [222]               | Agrobacterium tumefaciens       |
| Common bean                 | Brazilnut       | Brazilnut 2S albumin               | Increased methionine by 14–23%                                          | [220]               | Particle bombardment            |
| Narrow-leafed lupin         | Sunflower       | Sunflower seed albumin             | Increased methionine by 90%                                            | [224,225]           | Agrobacterium tumefaciens       |
| Arabidopsis                 | Serine acetyltransferase |                                  | 26-fold increase in free cysteine                                    | [229]               | Agrobacterium tumefaciens       |
| Soybean                     | Maize           | 15 kDa δ-zein                      | Increased methionine from 15.49 to 18.57% and cysteine from 26.97 to 29.33% | [226]               | Agrobacterium tumefaciens       |
|                             | Maize           | 27 kDa γ-zein                      | Methionine                                                               | [227]               | Particle bombardment            |
|                             | Maize           | 11 kDa δ-zein                      | Increased methionine by 16% and cysteine by 66%                        | [221]               | Agrobacterium tumefaciens       |
|                             | Maize           | MB-16                               | Increased methionine by 15–19%                                         | [230]               | Biolistic                       |
| Soybean                     | Soybean plastid ATP | sultrylase isoform 1               | Increased cysteine by 37–52% and methionine by 15–19%                 | [228]               | Agrobacterium tumefaciens       |
|                             | Maize           | 11 kDa δ-zein                      | Increased sulfur amino acids                                           | [222]               | Agrobacterium tumefaciens       |
|                             | Soybean         | Glyma.20g085100                     | Enhance protein content                                                 | [137]               | RNAi technology                 |

Figure 1. Integrated ‘omics’ and emerging novel breeding approach for improving protein content in grain legumes.
Despite some successes introducing transgenes to enhance SPC in grain legumes at the transgenic level, transgenic regulatory or governing bodies do not allow or restrict the use of these genetically engineered improved grain legumes commercially due to health and environmental safety issues. To overcome these stringent issues related to genetically modified crops, rapidly evolving genome editing technologies could help develop enhanced-protein grain legumes without introducing foreign genes. Using genome editing technologies, various crop plants have improved quality traits, such as increased fragrance and low gluten, starch, or oleic acid contents (for details, see [231]). However, the use of genome editing for SPC fortification in grain legumes is limited; future studies could adopt these powerful technologies to improve SPC by editing various gene(s), such as those encoding essential sulfur-rich amino acids or improving storage proteins.

10. Whole Genome Resequencing and Pangenome Sequencing for Elucidating Novel Structural Variants Related to High SPC across the Genome

Current breakthroughs in genome sequencing technologies have facilitated the sequencing of the global germplasm of various crops, including legumes, to underpin novel structural variants (SVs) such as presence/absence and copy number variations prevailing at the genome level [232,233]. An analysis combining association and biparental mapping using WGRS data of 631 soybean genotypes discovered a pleiotropic sugar transporter QTL gene \( \text{GmSWEET39} \) on chromosome 15 controlling SPC and seed oil content [164]. The authors suggested that deletion of 2 bp CC in the underlying causative \( \text{Glyma.15G049200} \) gene reduced SPC and enhanced seed oil content. Likewise, a pangenomic approach can describe the full complement of genes in the ‘core genome’ and ‘accessory genome’ to capture structural variation (not available in ‘single reference genome assembly’) at the species level [232]. Pangenome assemblies have been reported in chickpea [233], pigeon pea [234], soybean [235] and mungbean [236]. Thus, future construction and annotation of pangenes for different grain legumes could reveal missing information on SPC structural variations in the available reference genome assemblies, expediting the development of grain legumes with enriched protein.

11. Non-Destructive Phenomics Approach for Quantifying High Protein Content in Grain Legumes

Several high-throughput phenotyping approaches have been developed to bridge the genotyping and phenotyping gap for various quality traits, including protein content [237–239]. Advances in high-throughput non-destructive phenotyping approaches such as hyperspectral technologies, near-infrared reflectance spectroscopy, and nuclear magnetic resonance have enabled the phenotyping of various biochemical attributes in cereal and legume seeds, including protein content, with high accuracy and efficiency [237–241]. For example, Raman spectroscopy has been used to measure SPC in soybean [237]. Earlier, near-infrared reflectance spectroscopy was used to screen high-protein soybean genotypes [242,243]. Thus, non-destructive high-throughput phenotyping approaches could save time when screening high-SPC lines.

12. Genomic Selection and Rapid Generation Advances for Selecting High SPC Lines to Increase Genetic Gain

Unprecedented advances in genome-wide molecular marker development allow the use of genomic selection (GS) for predicting the genetic merit of progenies with complex traits without observing their phenotypic values from large target populations by developing a prediction model and calculating genomic-assisted breeding values in a ‘training population’ with known phenotypic observation [244]. The benefit of GS for improving genetic gain could be harnessed by increasing selection intensity \( i \) and selection accuracy \( r \), and reducing the breeding cycle length \( L \) in the breeder’s equation: \( \Delta G = R = h^2S = \sigma_a \times i \times r/L. \) \( \Delta G = \) genetic gain, \( R = \) response to selection, \( h^2 = \) heritability, \( \sigma_a = \) additive genetic variance. Notable instances of using GS as a substitute for phenotypic selection for complex traits include grain yield under moisture stress in chickpea [245], common
bean [246], cowpea (Ravelombola et al., 2021) [247], and pea [248,249] and cooking time in common bean [250]. However, GS has limited application for selecting high SPC genotypes in legumes [251]. A rrBLUP model was used to predict SPC in 306 pea genotypes derived from three RILs, tested in three autumn seasons in northern and central Italy, to determine any advantage of GS over phenotypic selection for SPC [251]. The mean predictive ability of GS for SPC was 0.53. Future studies could use GS to improve SPC and select various grain legume progenies with high SPC without phenotyping.

Likewise, the emerging benefits of speed breeding techniques could be harnessed by using optimum light intensity, photoperiod and temperature to enhance the rate of photosynthesis, resulting in early flowering and plant maturity, thus shortening the breeding cycle [252]. Speed breeding protocols have been established in chickpea, lupin, lentil, pea, soybean, and faba bean [253–257]. Further optimization of speed breeding protocols could fast-track improvements in various traits of breeding importance, including SPC, in grain legumes for sustaining global food security.

13. Fundamental Constraints on Seed Protein Content

As the offspring of plants, seeds are subject to several fundamental trade-offs that impact their size and composition. Seeds have fundamental required components, such as cell walls, and some amount of carbohydrates, lipids, and nucleic acids to make a viable embryo. Consequently, there are limits to potential selection on protein content. For example, long term selection on maize seed oil content has shown limits to the power of selection (e.g., [258]). Over the past two or more decades, ecologists have increasingly conceptualized these trade-offs as part of an economic spectrum, which influences the range of traits observed in leaves [259,260], stems [261,262] and roots [263]. As a dispersal unit, seeds are able to travel farther if they are smaller, but establish more readily if larger [264]. In many individual legume crops, wild relatives have presumably been under millenia of selection for these trade-offs in seed size and composition, limiting genetic variation and architecture. However, few researchers have linked evolutionary and ecological limits on seed composition to efforts at breeding, nor looked carefully at how they impact seed protein content. Seed size is generally an important co-variate in seed protein content, although among legumes its role differs somewhat among grain legumes.

Recent elegant work in chickpea suggests that these constraints are in fact real, and shape contemporary genetic diversity in seed size and composition. Chickpea has a QTL hotspot for seed size, leaf size, drought responses, and other “Vigour” traits. Nguyen and colleagues have recently fine mapped this QTL [265,266] showing it to be due to variation in a TIFY gene, which mutant studies in Arabidopsis have shown to impact seed size. Natural variation at this locus suggests it contributes significantly to a seed-size number trade-off, among parents that also differ in seed protein content.

14. Conclusions and Future Perspective

The increasing human population is facing increasing malnutrition-related problems such as dietary protein deficiency, especially in underprivileged and developing countries. Supplying protein-rich legumes improved through plant breeding and molecular breeding approaches could minimize the rising challenge of hunger and malnutrition-related problems. Moreover, improved grain legume dietary protein could be an important and economically viable alternative to high-cost animal-based dietary protein. Protein biofortification of major grain legumes will help satisfy the daily needs of human dietary protein in underprivileged and developing countries. Accurate characterization of various crop gene pool and landrace haplotypes with genetic variation for SPC needs urgent attention to accelerate SPC improvement in legumes. Harnessing the benefits of pre-breeding approaches could play a pivotal role in introgressing gene(s)/QTLs regulating high protein content from CWRs into high-yielding low-protein elite legume cultivars [96]. Recent advances in genomics, genome-wide association mapping, and whole genome resequencing approaches and the availability of complete genome and pangenome sequences in vari-
ous legume crops could help underpin the causative alleles/QTLs/haplotypes/candidate genes controlling high protein at the genome level, enabling genomics-assisted selection for improving protein concentration in grain legumes. Likewise, functional genomics, proteomics, and metabolomics could enrich our understanding of the complex molecular networks controlling improved protein content in various grain legumes. Selecting protein-rich grain legume genotypes in assessed germplasm or segregating progenies is challenging as most protein-estimating processes are based on destructive methods. Thus, high-throughput non-destructive methods are important for selecting high-protein legume genotypes. Likewise, genomic selection and rapid generation advances could be important for selecting high-protein progenies and rapidly developing protein-dense legumes. To overcome the challenges of transgenic technology, genome editing will help us manipulate and edit genes(s) governing high protein content at specific locations on legume genomes to enhance SPC. Capitalizing on these modern breeding tools, we should be able to identify grain legumes with improved protein content without compromising yield, as these two traits have a strong inverse relationship [123]. Hence, the amalgamation of approaches could help combat the growing protein-based malnutrition and lower the hunger risk, ensuring sustainable human growth globally.

Author Contributions: Conceptualization, U.C.J. and K.H.M.S.; writing—original draft preparation, U.C.J., H.N., S.K.P., R.D. and E.J.B.v.W. E.J.B.v.W. also wrote some sections for this manuscript; overall manuscript editing, K.H.M.S. and R.D. All authors have read and agreed to the published version of the manuscript.

Funding: APC was provided by The University of Western Australia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the article.

Acknowledgments: U.C.J. acknowledge support from ICAR, New Delhi, India.

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

References
1. Hasegawa, T.; Fujimori, S.; Takahashi, K.; Yokohata, T.; Masui, T. Economic implications of climate change impacts on human health through undernourishment. Clim. Chang. 2016, 136, 189–202. [CrossRef]
2. FAO; IFAD; UNICEF; WFP; WHO. State of Food Security and Nutrition in the World 2020: Transforming Food Systems for Affordable Healthy Diets; Food & Agriculture Organization: Rome, Italy, 2020. [CrossRef]
3. Wu, G. Dietary protein intake and human health. Food Funct. 2016, 7, 1251–1265. [CrossRef] [PubMed]
4. Millward, D.J.; Jackson, A.A. Protein/energy ratios of current diets in developed and developing countries compared with a safe protein/energy ratio: Implications for recommended protein and amino acid intakes. Public Health Nutr. 2004, 7, 387–405. [CrossRef] [PubMed]
5. Müller, O.; Krawinkel, M. Malnutrition and health in developing countries. CMAJ 2005, 173, 279–286. [CrossRef]
6. Borresen, E.C.; Zhang, L.; Trehan, I.; Nealon, N.J.; Maleta, K.M.; Manary, M.J.; Ryan, E.P. The nutrient and metabolite profile of 3 complementary legume foods with potential to improve gut health in rural Malawian children. Curr. Dev. Nutr. 2017, 1, e001610. [CrossRef]
7. Trehan, I.; Benzoni, N.S.; Wang, A.Z.; Bollinger, L.B.; Ngoma, T.N.; Chimimba, U.K.; Stephenson, K.B.; Agapova, S.E.; Maleta, K.M.; Manary, M.J. Common beans and cowpeas as complementary foods to reduce environmental enteric dysfunction and stunting in Malawian children: Study protocol for two randomized controlled trials. Trials 2015, 16, 1–12. [CrossRef]
8. Rubiales, D.; Mikic, A. Introduction: Legumes in Sustainable Agriculture. CRC Crit. Rev. Plant Sci. 2015, 34, 2–3. [CrossRef]
9. Multari, S.; Stewart, D.; Russell, W.R. Potential of fava bean as future protein supply to partially replace meat intake in the human diet. Compr. Rev. Food Sci. Food Saf. 2015, 14, 511–522. [CrossRef]
10. Hall, C.; Hillen, C.; Garden, R.J. Composition, nutritional value, and health benefits of pulses. Cereal Chem. 2017, 94, 11–31. [CrossRef]
11. Singh, B.; Singh, J.P.; Shevkani, K.; Singh, N.; Kaur, A. Bioactive constituents in pulses and their health benefits. J. Food Sci. Technol. 2017, 54, 858–870. [CrossRef]
12. Hou, D.; Yousafl, L.; Xue, Y.; Hu, J.; Wu, J.; Hu, X.; Feng, N.; Shen, Q. Mung bean (Vigna radiata L.): Bioactive polyphenols, polysaccharides, peptides, and health benefits. Nutrients 2019, 11, 1238. [CrossRef] [PubMed]
13. Di Paola, A.; Rulli, M.C.; Santini, M. Human food vs. Animal feed debate. A thorough analysis of environmental footprints. *Land Use Policy* **2017**, *67*, 652–659. [CrossRef]

14. Hulse, J.H. Nature, composition and utilization of pulses. In *Uses of Tropical Grain Legumes*. In Proceedings of the Consultants Meeting, Patancheru, India, 27–30 March 1989; ICARDA: Patancheru, India, 1991; pp. 11–27.

15. Badshah, A.; Khan, M.; Bibi, N.; Khan, M.; Ali, S.; Ashraf Chaudry, M. Quality studies of newly evolved chickpea cultivars. *Adv. Food Sci.* **2003**, *25*, 95–99.

16. Manan, F.; Hussain, T.; Iqbal, P. Proximate composition and minerals constituents of important cereals and pulses grown in NWFP. *Pak. J. Sci. Res.* **1984**, *36*, 45–49.

17. Urbano, G.; Porres, J.M.; Frais, J.; Vidal-Valverde, C. Nutritional value. In *Lentil: An Ancient Crop for Modern Times*; Yadav, S.S., McNeil, D.L., Stevenson, P.C., Eds.; Springer: Dordrecht, The Netherlands, 2007; pp. 47–93.

18. Khazaei, H.; Subedi, M.; Nickerson, M.; Martínez-Delgado, M.; Martínez-Villaluenga, C.; Frias, J.; Vandenberg, A. Seed protein of lentils: Current status, progress, and food applications. *Foods* **2019**, *8*, 391. [CrossRef] [PubMed]

19. Lucas, M.M.; Stoddard, F.L.; Annicchiarico, P.; Frias, J.; Martinez-Villaluenga, C.; Sussmann, D.; Duranti, M.; Seger, A.; Zander, P.M.; Pueyo, J.J. The future of lupin as a protein crop in Europe. *Front. Plant Sci.* **2015**, *6*, 705. [CrossRef] [PubMed]

20. Duranti, M.; Consorini, A.; Magni, C.; Sessa, F.; Scarafoni, A. The major proteins of lupin seed: Characterisation and molecular properties for use as functional and nutraceutical ingredients. *Trends Food Sci. Technol.* **2008**, *19*, 624–633. [CrossRef]

21. Duranti, M.; Cerletti, P. Amino acid composition of seed proteins of *Lupinus albus*. *J. Agric. Food Chem.* **1979**, *27*, 977–978. [CrossRef]

22. Hou, A.; Chen, P.; Alloatti, J.; Mozzoni, L.; Zhang, B.; Shi, A. Genetic variability of seed sugar content in worldwide soybean germplasm collections. *Crop Sci.* **2009**, *49*, 903–912. [CrossRef]

23. Sharma, S.; Kaur, M.; Goyal, R.; Gill, B.S. Physical characteristics and nutritional composition of some new soybean (*Glycine max* (L.) Merril) genotypes. *J. Food Sci. Technol.* **2014**, *51*, 551–557. [CrossRef]

24. Patil, G.; Mian, R.; Vuong, T.; Pantalone, V.; Song, Q.; Shen, P.; Shannon, G.J.; Carter, T.C.; Nguyen, H.T. Molecular mapping and genomics of soybean seed protein: A review and perspective for the future. *Theor. Appl. Genet.* **2017**, *130*, 1975–1991. [CrossRef] [PubMed]

25. Shellie-Dessert, K.; Bliss, F. Genetic improvement of food quality factors. Common beans. In *Research for Crop Improvement*; van Schoonhoven, A., Voyset, O., Eds.; CAB International: Wallingford, UK, 1991; pp. 649–677.

26. Sathe, S.K.; Desphande, S.S.; Salunkhe, D.K.; Rackis, J.J. Dry beans of *Phaseolus* (*Phaseolus vulgaris* L.)/walnut). *J. Food Sci.* **2004**, *69*, 114–118. [CrossRef]

27. Taylor, M.; Chapman, R.; Beyaert, R.; Hernández-Sebastiáa, C.; Marsolais, F. Seed storage protein deficiency improves sulfur amino acid content in common bean (*Phaseolus vulgaris* L.): Redirection of sulfur from γ-glutamyl-S-methyl-cysteine. *J. Agric. Food Chem.* **2008**, *56*, 5647–5654. [CrossRef] [PubMed]

28. Saxena, K.B.; Kumar, R.V.; Rao, P.V. Pigeonpea nutrition and its improvement. *J. Crop Prod.* **2002**, *5*, 227–260. [CrossRef]

29. Bressani, R.; Gómez-Brenes, R.A.; Elias, L.G. Nutritional quality of pigeon pea protein, immature and ripe, and its supplementary value for cereals. *Arch. Latinoam. Nutr.* **1986**, *36*, 108–116. [PubMed]

30. Picard, J. Some results dealing with breeding for protein content in *Vicia faba* L. In *Protein Quality from Leguminous Crops*; Station d’Amélioration des Plantes INRA: Dijon, France, 1997.

31. Crépon, K.; Marget, P.; Peyronnet, C.; Carroué, B.; Arese, P.; Duc, G. Nutritional value of faba bean (*Vicia faba* L.) seeds for food and feed. *Field Crops Res.* **2010**, *115*, 329–339. [CrossRef]

32. Anwar, F.; Latif, S.; Przybylski, R.; Sultana, B.; Ashraf, M. Chemical composition and antioxidant activity of seeds of different cultivars of mung bean. *J. Food Sci. Technol.* **2007**, *72*, S503–S510. [CrossRef]

33. Horax, R.; Hettiarachchy, N.S.; Chen, P.; Jalaluddin, M. Preparation and characterization of protein isolate from cow pea (*Vigna unguiculata* (L) walp). *J. Food Sci.* **2004**, *69*, 114–118. [CrossRef]

34. Gerrano, A.S.; Jansen van Rensburg, W.S.; Venter, S.L.; Sharige, N.G.; Amelework, B.A.; Shimelis, H.A.; Labuschagne, M.T. Selection of cowpea genotypes based on grain mineral and total protein content. *Acta Agric. Scand. Sect. B Soil Plant Sci.* **2019**, *69*, 155–166. [CrossRef]

35. Martos-Fuentes, M.; Sánchez-Navarro, V.; Ruiz-Hernández, M.V.; Weiss, J.; Egea-Gilabert, C.; Zornoza, R.; Faz, A.; Fernández, J.A.; Egea-Cortes, M. Genetic and growth conditions determine the protein content in cowpea (*Vigna unguiculata*). In *Proceedings of the EUCARPIA International Symposium on Protein Crops*: V Meeting AEL [V JORNADAS DE LA AEL]; Pontevedra, Spain, 4–7 May 2015; pp. 143–144. [CrossRef]

36. Mossié, J.; Huet, J.C.; Baudet, J. Changements de la composition en acides aminés des graines de pois en fonction de leur taux d’azote. *Sci. Aliment.* **1987**, *7*, 301–324.

37. Titzikas, E.N.; Vincken, J.P.; de Groot, J.; Gruppen, H.; Visser, R.G. Genetic variation in pea seed globulin composition. *J. Agric. Food Chem.* **2006**, *54*, 425–433. [CrossRef] [PubMed]

38. Duc, G.; Agrama, H.; Bao, S.; Berger, J.; Bourion, V.; De Ron, A.M.; Gowda, C.L.L.; Mikic, A.; Millot, D.; Singh, K.B.; et al. Breeding annual grain legumes for sustainable agriculture: New methods to approach complex traits and target new cultivar ideotypes. *Crit. Rev. Plant Sci.* **2015**, *34*, 381–411. [CrossRef]

39. Özer, S.; Tümer, E.; Baloch, F.S.; Karaköy, T.; Toklu, F.; Özkan, H. Variation for nutritional and cooking properties among Turkish field pea landraces. *J. Food Agric. Environ.* **2012**, *10*, 324–329.
40. Daba, S.D.; Morris, C.F. Pea proteins: Variation, composition, genetics, and functional properties. *Cereal Chem.* 2022, 99, 8–20. [CrossRef]

41. Zia-Ul-Haq, M.; Ahmad, S.; Bukhari, S.A.; Amarowicz, R.; Ercisli, S.; Jaafer, H.Z. Compositional studies and biological activities of some mash bean (*Vigna mungo* (L.) Hepper) cultivars commonly consumed in Pakistan. *Biol. Res.* 2014, 47, 1–14. [CrossRef] [PubMed]

42. Ramya, K.R.; Tripathi, K.; Pandey, A.; Barpate, S.; Gore, P.G.; Raina, A.P.; Khawar, K.M.; Swain, N.; Sarker, A. Rediscovering the Potential of Multifaceted Orphan Legume Grasspea-a Sustainable Resource with High Nutritional Values. *Front. Nutr.* 2022, 8, 826208. [CrossRef]

43. Siharan, K.; Khetarpaul, N. Protein quality traits of vegetable and field peas: Varietal differences. *Plant Foods Hum. Nutr.* 1994, 45, 11–22. [CrossRef]

44. Wallace, T.C.; Murray, R.; Zelman, K.M. The nutritional value and health benefits of chickpeas and hummus. *Nutrients* 2016, 8, 766. [CrossRef]

45. Khan, M.A.; Akhtar, N.; Ullah, I.; Jaffery, S. Nutritional evaluation of desi and kabuli chickpeas and their products commonly consumed in Pakistan. *Int. J. Food Sci. Nutr.* 1995, 46, 215–223. [CrossRef]

46. Shimelis, E.A.; Rakshit, S.K. Effect of processing on antinutrients and in vitro protein digestibility of kidney bean (*Phaseolus vulgaris* L.) varieties grown in East Africa. *Food Chem.* 2007, 103, 161–172. [CrossRef]

47. Drewnowski, A. The cost of US foods as related to their nutritive value. *Am. J. Clin. Nutr.* 2010, 92, 1181–1188. [CrossRef] [PubMed]

48. Lioli, I. Variation of the storage protein phaseolin in common bean (*Phaseolus vulgaris* L.) from the Mediterranean area. *Euphytica* 1989, 44, 151–155. [CrossRef]

49. Bollini, R.; Vitale, A. Genetic variability in charge microheterogeneity and polypeptide composition of phaseolin, the major storage protein of *Phaseolus vulgaris*; and peptide maps of its three major subunits. *Physiol. Plant.* 1981, 52, 96–100. [CrossRef]

50. Gepts, P.; Bliss, F.A. Phaseolin variability among wild and cultivated common beans (*Phaseolus vulgaris*) from Colombia. *Econ. Bot.* 1986, 40, 469–478. [CrossRef]

51. Johnson, S.; Grayson, G.; Robinson, L.; Chahade, R.; McPherson, A. Biochemical and crystallographic data for phaseolin, the major protein of *Phaseolus vulgaris*. *Biochemistry* 2002, 41, 4839–4843. [CrossRef]

52. Phillips, R.D.; McWatters, K.H.; Chinnan, M.S.; Hung, Y.C.; Beuchat, L.R.; Sefa-Dedeh, S.; Sakyi-Dawson, E.; Ngoddy, P.; Nnyanyelugo, D.; Enwere, J.; et al. Utilization of cowpeas for human food. *Field Crop. Res.* 2003, 82, 193–213. [CrossRef]

53. Timko, M.P.; Singh, B.B. Cowpea, a multifunctional legume. In *Genomics of Tropical Crop Plants*; Springer: New York, NY, USA, 2008; pp. 227–258.

54. Jayathilake, C.; Visvanathan, R.; Deen, A.; Bangamuwage, R.; Jayawardana, B.C.; Nammi, S.; Liyanage, R. Cowpea: An overview on its nutritional facts and health benefits. *J. Sci. Food Agric.* 2018, 98, 4793–4806. [CrossRef]

55. Warsame, A.O.; O’Sullivan, D.M.; Tosi, P. Seed storage proteins of faba bean (*Vicia faba* L.): Current status and prospects for genetic improvement. *J. Agric. Food Chem.* 2018, 66, 12617–12626. [CrossRef]

56. Hao, P.; Zhu, Y.; Feng, Q.; Jin, Z.; Wu, F. Differences in Grain Microstructure and Proteomics of a Broad Bean (*Vicia faba* L.) Landrace Cixidabaican in China Compared with Lingxiyicun Introduced from Japan. *Plants* 2021, 10, 1385. [CrossRef]

57. Lisiewaka, Z.; Kmiecik, W.; Slupski, J. Content of amino acids in raw and frozen broad beans (*Vicia faba* var. major) seeds at milk maturity stage, depending on the processing method. *Food Chem. 2007*, 105, 1468–1473. [CrossRef]

58. Samaei, S.P.; Ghorbani, M.; Tagliazucchi, D.; Martini, S.; Gotti, R.; Themelis, T.; Tesini, F.; Gianotti, A.; Toschi, T.G.; Babini, E. Functional, nutritional, antioxidant, sensory properties and comparative peptidomic profile of faba bean (*Vicia faba* L.) seed protein hydrolysates and fortified apple juice. *Food Chem.* 2020, 330, 127120. [CrossRef] [PubMed]

59. Salaria, S.; Boatwright, J.L.; Thavarajah, P.; Kumar, S.; Thavarajah, D. Protein Biofortification in Lentils (*Lens culinaris* Medik.) Toward Human Health. *Front. Plant Sci.* 2022, 13, 869713. [CrossRef] [PubMed]

60. Saint-Clair, P.M. Responses of Lens Esculenta Moench to Controlled Environmental Factors; Wangeningen University: Wangeningen, The Netherlands, 1972.

61. Suliema, M.A.; Hassan, A.B.; Osman, G.A.; El Tyeb, M.M.; El Kh, E.A.; El Tina, A.H.; Babiker, E.E. Changes in Total Protein Digestibility, Fractions Content and Structure During Cooking of Lentil Cultivars. *Pak. J. Nutr.* 2008, 7, 193–213. [CrossRef]

62. Kurlovich, B.S., Ed.; OY International Express: St. Petersburg, Russia, 2010; pp. 227–258.

63. Petterson, D.S. Composition and food uses of lupins. In *Lupin as Crop Plants. Biology, Production and Utilization*; Kurlovich, B.S., Ed.; OY International Express: St. Petersburg, Russia, 2010; pp. 227–258.

64. Guillamon, E.; Rodriguez, J.; Burbano, C.; Muzquiz, M.; Pedrosa, M.M.; Cabanillas, B.; Crespo, J.F.; Sancho, A.I.; Mills, E.N.C.; Cuadrado, C. Characterization of lupin major allergens (*Lupinus albus* L.). *Mol. Nutr. Food Res.* 2010, 54, 1668–1676. [CrossRef] [PubMed]

65. Janusz, P. White lupin (*Lupinus albus* L.)—nutritional and health values in human nutrition—a review. *Czech J. Food Sci.* 2017, 35, 95–105. [CrossRef]

66. Lajalo, F.M.; Genevese, M. Nutritional significance of lectins and enzyme inhibitors from legumes. *J. Agric. Food Chem.* 2002, 50, 6592–6598. [CrossRef]
67. American Soybean Association, 2019 SOYSTATS; A Reference Guide to Soybean Facts and Figures; American Soybean Association: Saint Louis, MO, USA, 2019.

68. Kolláróvá, K.; Vatehová, Z.; Slováková, L.U.; Lišková, D. Interaction of galactoglucomannan oligosaccharides with auxin in mung bean primary root. Plant Physiol. Biochem. 2010, 48, 401–406. [CrossRef]

69. Yi-Shen, Z.; Shuai, S.; Fitzgerald, R. Mung bean proteins and peptides: Nutritional, functional and bioactive properties. Food Nutr. Res. 2018, 62. [CrossRef]

70. Mubarak, A.E. Nutritional composition and antinutritional factors of mung bean seeds (Phaseolus aureus) as affected by some home traditional processes. Food Chem. 2005, 89, 489–495. [CrossRef]

71. Bazaz, R.; Baba, W.N.; Masoodi, F.A.; Yaqoob, S. Formulation and characterization of hypo allergic weaning foods containing potato and sprouted green gram. J. Food Meas. Charact. 2016, 10, 453–465. [CrossRef]

72. Barac, M.; Cabrilo, S.; Pesic, M.; Stanojevic, S.; Zilic, S.; Macej, O.; Ristic, N. Profile and functional properties of seed proteins from six pea (Pisum sativum) genotypes. Int. J. Mol. Sci. 2010, 11, 4973–4990. [CrossRef] [PubMed]

73. Varshney, R.K.; Pennettsa, R.V.; Dutta, S.; Kulwal, P.L.; Saxena, R.K.; Datta, S.; Sharma, T.R.; Rosen, B.; Carrasquilla-Garcia, N.; Farmer, A.D.; et al. Pigeonpea genomics initiative (PGI): An international effort to improve crop productivity of pigeonpea ( Cajanus cajan L.). Mol. Breed. 2010, 26, 393–408. [CrossRef] [PubMed]

74. Mahajan, R.; Malhotra, S.P.; Singh, R. Characterization of seed storage proteins of urdbean (Vigna mungo). Plant Foods Hum. Nutr. 1988, 38, 163–173. [CrossRef] [PubMed]

75. Gottschalk, W.; Mueller, H.P.; Wolff, G. The genetic control of seed protein production and composition (peas). Egypt J. Gene Cytol. 1975, 4, 453–468.

76. Erskine, W.; Williams, P.C.; Nkouk, H. Genetic and environmental variation in the seed size, protein, yield, and cooking quality of lentils. Field Crops Res. 1985, 12, 153–161. [CrossRef]

77. Boukar, O.; Massawe, F.; Muranaka, S.; Franco, J.; Maziya-Dixon, B.; Singh, B.; Fatokun, C. Evaluation of cowpea germplasm lines for protein and mineral concentrations in grains. Plant Genet. Resour. 2011, 9, 515–522. [CrossRef]

78. Gaur, P.M.; Singh, M.K.; Samineni, S.; Sajja, S.B.; Jukanti, A.K.; Kamatam, S.; Varshney, R.K. Inheritance of protein content and its relationships with seed size, grain yield and other traits in chickpea. Euphytica 2016, 209, 253–260. [CrossRef]

79. Celmeli, T.; Sari, H.; Cangi, H.; Sari, D.; Adak, A.; Eker, T.; Toker, C. The nutritional content of common bean (Phaseolus vulgaris L.) landraces in comparison to modern varieties. Agronomy 2018, 8, 166. [CrossRef]

80. Pandir, R.P.S.; Reddy, K.N.; Mengesha, M.H. ICRISAT Chickpea Germplasm Catalog: Evaluation and Analysis; ICRISA: Patancheru, India, 1988; Volume 94, pp. 89–100.

81. Jadhav, A.A.; Rayate, S.J.; Mhase, L.B.; Thudi, M.; Chitikineni, A.; Harer, P.N.; Jadhav, A.S.; Varshney, R.K.; Kulwal, P.L. Marker-trait association study for protein content in chickpea (Cicer arietinum L.). J. Genet. 2015, 94, 279–286. [CrossRef]

82. Wang, N.; Daun, J.K. Effect of variety and crude protein content on nutrients and certain antinutrients in field peas (Pisum sativum). J. Sci. Food Agric. 2000, 80, S11–S26. [CrossRef]

83. Jukanti, A.K.; Gaur, P.M.; Gowda, C.L.L.; Chibbar, R.N. Nutritional quality and health benefits of chickpea (Cicer arietinum L.). A review. Br. J. Nutr. 2012, 108, S11–S26. [CrossRef]

84. Serrano, C.; Carbas, B.; Castanho, A.; Soares, A.; Patto, M.C.; Brites, C. Characterisation of nutritional quality traits of a chickpea (Cicer arietinum) germplasm collection exploited in chickpea breeding in Europe. Crop Pasture Sci. 2017, 68, 1031–1040. [CrossRef]

85. Weng, Y.; Qin, J.; Eaton, S.; Yang, Y.; Ravelombola, W.S.; Shi, A. Evaluation of seed protein content in USDA cowpea germplasm. Hort. Sci. 2019, 54, 814–817. [CrossRef]

86. Gupta, P.; Singh, R.; Malhotra, S.; Boora, K.S.; Singal, H.R. Characterization of seed storage proteins in high protein genotypes of cowpea [Vigna unguiculata (L.) Walp.] J. Food. Mol. Biol. Plants 2010, 16, 53–58. [CrossRef] [PubMed]

87. GonÇalves, F.V.; Medici, L.O.; Fonseca, M.P.S.; Pimentel, C.; Gaziola, S.A.; Azevedo, R.A. Protein, phytate and minerals in grains of commercial cowpea genotypes. An. Bras. Ciências Animais 2020, 92. [CrossRef]

88. Carvalho, A.F.U.; de Sousa, M.B.; de Lima, G.P.G.; et al. Nutritional ranking of 30 Brazilian genotypes of cowpeas including determination of antioxidant capacity and vitamins. J. Food Compos. Agron. 2018, 26, 81–88. [CrossRef]

89. Guzmán-Maldonado, S.H.; Acosta-Gallegos, J.; Paredes-López, O. Protein and mineral content of a novel collection of wild and weedy common bean (Phaseolus vulgaris L.). J. Sci. Food Agric. 2000, 80, 1874–1881. [CrossRef]

90. Rezende, A.A.; Pacheco, M.T.B.; Silva, V.S.N.D.; Ferriera, T.A.P.D.C. Nutritional and protein quality of dry Brazilian beans (Phaseolus vulgaris L.). Food Sci Technol. 2017, 38, 421–427. [CrossRef]

91. Mendes, F.A.; Leitão, S.T.; Correia, V.; Mecha, E.; Rubiales, D.; Bronze, M.R.; Vaz Patto, M.C. Portuguese Common Bean Natural Variation Helps to Clarify the Genetic Architecture of the Legume’s Nutritional Composition and Protein Quality. Plants 2022, 11, 26. [CrossRef]

92. Ravelombola, W.S.; Shi, A.; Weng, Y.; Motes, D.; Chen, P.; Srivastava, V.; Wingfield, C. Evaluation of total seed protein content in eleven Arkansas cowpea (Vigna unguiculata (L.) Walp.) lines. Annu. J. Plant Sci. 2016, 7, 2288–2296. [CrossRef]

93. Dakora, F.D.; Belane, A.K. Evaluation of protein and micronutrient levels in edible cowpea (Vigna Ungueiculata L. Walp.) leaves and seeds. Front. Sustain. Food Syst. 2019, 3, 70. [CrossRef]

94. Khazaei, H.; Vandenberg, A. Seed Mineral Composition and Protein Content of Faba Beans (Vicia faba L.) with Contrasting Tannin Contents. Agronomy 2020, 10, 511. [CrossRef]
148. Shannon, J.G.; Wilcox, J.R.; Probst, A.H. Estimated grains from selection for protein and yield in the F4 generation of six soybean populations. Crop Sci. 1972, 12, 824–826. [CrossRef]
149. Dahiya, B.S.; Brar, J.S.; Bhullar, B.S. Inheritance of protein content and its correlation with grain yield in pigeonpea (Cajanus cajan (L.) Millsp.). Qual. Plant. 1977, 27, 327–334. [CrossRef]
150. Burton, J.W. Breeding soybeans for improved protein quantity and quality. In Proceedings of the 3rd World Soybean Research Conference; Shibles, R., Ed.; Westview Press: Boulder, CO, USA, 1985, pp. 361–367.
151. Wilcox, J.R. Breeding soybeans for improved oil quantity and quality. In World Soybean Research Conference III: Proceedings; Shibles, R., Ed.; Westview Press: Boulder, CO, USA, 1985; pp. 380–386.
152. Gautam, N.K.; Bhardwaj, R.; Yadav, S.; Suneha, P.; Tripathi, K.; Ram, B. Identification of lentil (Lens culinaris Medik.) germplasm rich in protein and amino acids for utilization in crop improvement. Ind. J. Genet. 2018, 78, 9. [CrossRef]
153. Tchiagam, L.B.N.; Bell, J.M.; Nassourou, A.M.; Njintang, N.Y.; Youmbi, E. Genetic analysis of seed proteins contents in cowpea (Vigna unguiculata). Afr. J. Biotechnol. 2011, 10, 3077–3086.
154. Kumar, S.; Gupta, P.; Chouki, H.; Siddique KH, M. “Efficient breeding of pulse crops”. In Accelerated Plant Breeding; Gosal, S.S., Wani, S.H., Eds.; Springer International Publishing: Cham, Switzerland, 2020; Volume 3, pp. 1–30.
155. Diers, B.W.; Keim, P.; Fehr, W.R.; Shoemaker, R.C. RFLP analysis of soybean seed protein and oil content. Theor. Appl. Genet. 1992, 83, 608–612. [CrossRef]
156. Nichols, D.M.; Glover, K.D.; Carlson, S.R.; Specht, J.E.; Diers, B.W. Fine mapping of a seed protein QTL on soybean linkage group I and its correlated effects on agronomic traits. Crop Sci. 2006, 46, 834–839. [CrossRef]
157. Obala, J.; Saxena, R.K.; Singh, V.K.; Kale, S.M.; Garg, V.; Kumar, C.V.; Saxena, K.B.; Tongoon, P.; Sibiya, J.; Varshney, R.K. Seed protein content and its relationships with agronomic traits in pigeonpea is controlled by both main and epistatic effects QTLs. Sci. Rep. 2020, 10, 1–17. [CrossRef]
158. Wang, R.; Gangola, M.P.; Irvine, C.; Gaur, P.M.; Bäga, M.; Chibbar, R.N. Co-localization of genomic regions associated with seed morphology and composition in a desi chickpea (Cicer arietinum L.) population varying in seed protein concentration. Theor. Appl. Genet. 2019, 132, 1263–1281. [CrossRef] [PubMed]
159. Irzykowska, L.; Wilko, B. Interval mapping of QTLs controlling yield-related traits and seed protein content in Pisum sativum. J. Appl. Genet. 2004, 45, 297–306. [PubMed]
160. Gali, K.K.; Liu, Y.; Sindhu, A.; Diapari, M.; Shumugam, A.S.K.; Argansoa, G.; Daba, K.; Caron, C.; Lachagari, R.V.B.; Tar’An, B.; et al. Construction of high-density linkage maps for mapping quantitative trait loci for multiple traits in field pea (Pisum sativum L.). BMC Plant Biol. 2018, 18, 1–25. [CrossRef]
161. Van, K.; McHale, L.K. Meta-analyses of QTLs associated with protein and oil contents and compositions in soybean [Glycine max (L.) Merr.] seed. Int. J. Mol. Sci. 2017, 18, 1180. [CrossRef] [PubMed]
162. Bolon, Y.T.; Joseph, B.; Cannon, S.B.; Graham, M.A.; Diers, B.W.; Farmer, A.D.; May, G.D.; Muehlbauer, G.J.; Specht, J.E.; Tu, Z.J.; et al. Complementary genetic and genomic approaches help characterize the linkage group I seed protein QTL in soybean. BMC Plant Biol. 2010, 10, 41. [CrossRef]
163. Vaughn, J.N.; Nelson, R.L.; Song, Q.; Cregan, P.B.; Li, Z. The genetic architecture of seed composition in soybean is refined by genome-wide association scans across multiple populations. G3 Genes Genomes Genet. 2014, 4, 2283–2294. [CrossRef] [PubMed]
164. Zhang, H.; Goettel, W.; Song, Q.; Jiang, H.; Hu, Z.; Wang, M.L.; An, Y.-Q.C. Selection of GmSWEET39 for oil and protein improvement in soybean. PLoS Genet. 2020, 16, e1009114. [CrossRef]
165. Pathan, S.M.; Vuong, T.; Clark, K.; Lee, J.D.; Shannon, J.G.; Roberts, C.A.; Ellersieck, M.R.; Burton, J.W.; Cregan, P.B.; Hyten, D.L.; et al. Genetic mapping and confirmation of quantitative trait loci for seed protein and oil contents and seed weight in soybean. Crop Sci. 2013, 53, 765–774. [CrossRef]
166. Kim, M.; Schultz, S.; Nelson, R.L.; Diers, B.W. Identification and fine mapping of a soybean seed protein QTL from PI 407788A on chromosome 15. Crop Sci. 2016, 56, 219–225. [CrossRef]
167. Patil, G.; Vuong, T.D.; Kale, S.; Valliyodan, B.; Deshmukh, R.; Zhu, C.; Wu, X.; Bai, Y.; Yungbluth, D.; Lu, F.; et al. Dissecting genomic hotspots underlying seed protein content, and sucrose content in an interspecific mapping population of soybean using high-density linkage mapping. Plant Biotechnol. J. 2018, 16, 1939–1953. [CrossRef]
168. Samanfar, B.; Cober, E.R.; Charette, M.; Tan, L.H.; Bekele, W.A.; Morrison, M.J.; Kilian, A.; Belzile, F.; Molnar, S.J. Genetic Analysis of High Protein Content in ‘AC Proteus’ Related Soybean Populations Using SSR, SNP, DArT and DArTseq Markers. Sci. Rep. 2019, 9, 16657. [CrossRef] [PubMed]
169. Karikari, B.; Li, S.; Bhat, J.A.; Cao, Y.; Kong, J.; Yang, J.; Gai, J.; Zhao, T. Genome-wide detection of major and epistatic effect QTLs for seed protein and oil content in soybean under multiple environments using high-density bin map. Int. J. Mol. Sci. 2019, 20, 979. [CrossRef] [PubMed]
170. Upadhayaya, H.D.; Bajaj, D.; Narromiya, L.; Das, S.; Kumar, V.; Gowda, C.L.L.; Sharma, S.; Tyagi, A.K.; Parida, S.K. Genome-wide scans for delineation of candidate genes regulating seed-protein content in chickpea. Front. Plant Sci. 2016, 7, 302. [CrossRef] [PubMed]
171. Campa, A.; Pañeda, A.; Pérez-Vega, E.; Giraldez, R.; Ferreira, J.J. Mapping and use of seed protein loci for marker-assisted selection of growth habit and photoperiod response in Nufna bean (Phaseolus vulgaris L.). Euphytica 2011, 179, 383–391. [CrossRef]
172. Casanas, F.; Pérez-Vega, E.; Almirall, A.; Plans, M.; Sabaté, J.; Ferreira, J.J. Mapping of QTL associated with seed chemical content in a RIL population of common bean (Phaseolus vulgaris L.). Euphytica 2013, 192, 279–288. [CrossRef]
173. Sarvamangala, C.; Gowda, M.V.C.; Varshney, R.K. Identification of quantitative trait loci for protein content, oil content and oil quality for groundnut (Arachis hypogaea L.). Field Crop. Res. 2011, 122, 49–59. [CrossRef]

174. Gali, K.K.; Sackville, A.; Tafesse, E.G.; Lachagari, V.R.; McPhee, K.; Hybl, M.; Mikić, A.; Smýkal, P.; McGee, R.; Burstin, J.; et al. Genome-Wide Association Mapping for Agronomic and Seed Quality Traits of Field Pea (Pisum sativum L.). Front. Plant Sci. 2019, 10, 1538. [CrossRef]

175. Chung, J.; Babka, H.L.; Graef, G.L.; Staswick, P.E.; Lee, D.J.; Cregan, P.B.; Shoemaker, R.C.; Specht, J.E. The seed protein, oil, and yield QTL on soybean linkage group I. Crop Sci. 2003, 43, 1053–1067. [CrossRef]

176. Hyten, D.L.; Pantalone, V.R.; Sams, C.E.; Saxton, A.M.; Landau-Ellis, D.; Stefaniak, T.R.; Schmidt, M.E. Seed quality QTL in a prominent soybean population. Theor. Appl. Genet. 2004, 109, 552–561. [CrossRef]

177. Fasoula, V.A.; Harris, D.K.; Boerma, H.R. Validation and designation of quantitative trait loci for seed protein, seed oil, and seed weight from two soybean populations. Crop Sci. 2004, 44, 1218–1223. [CrossRef]

178. Panthee, D.R.; Pantalone, V.R.; West, D.R.; Saxton, A.M.; Sams, C.E. Quantitative trait loci for seed protein and oil concentration, and seed size in soybean. Crop Sci. 2005, 45, 2015–2022. [CrossRef]

179. Lu, W.; Wen, Z.; Li, H.; Yuan, D.; Li, J.; Zhang, H.; Huang, Z.; Cui, S.; Du, W. Identification of the quantitative trait loci (QTL) underlying water soluble protein content in soybean. Theor. Appl. Genet. 2013, 126, 425–433. [CrossRef] [PubMed]

180. Wang, X.; Jiang, G.L.; Song, Q.; Cregan, P.B.; Scott, R.A.; Zhang, J.; Yen, Y.; Brown, M. Quantitative trait locus analysis of seed sulfur-containing amino acids in two recombinant inbred line populations of soybean. Euphytica 2015, 201, 293–305. [CrossRef]

181. Brzostowski, L.F.; Pruski, T.I.; Specht, J.E.; Diers, B.W. Impact of seed protein alleles from three soybean sources on seed quality for groundnut (Arachis hypogaea L.). Theor. Appl. Genet. 2017, 130, 2315–2326. [CrossRef] [PubMed]

182. Zhang, K.; Liu, S.; Li, W.; Liu, S.; Li, X.; Fang, Y.; Zhang, J.; Wang, Y.; Xu, S.; Zhang, J.; et al. Identification of QTNs Controlling Seed Protein Content in Soybean Using Multi-Locus Genome-Wide Association Studies. Front. Plant Sci. 2018, 9, 1690. [CrossRef]

183. Yuan, W.; Wu, Z.; Zhang, Y.; Yang, R.; Wang, H.; Kan, G.; Yu, D. Genome-Wide Association Mapping for Agronomic and Seed Quality Traits of Field Pea (Pisum sativum L.). Front. Plant Sci. 2019, 10, 1538. [CrossRef]

184. Zhang, K.; Liu, S.; Li, W.; Liu, S.; Li, X.; Fang, Y.; Zhang, J.; Wang, Y.; Xu, S.; Zhang, J.; et al. Identification of QTNs Controlling Seed Protein Content in Soybean Using Multi-Locus Genome-Wide Association Studies. Front. Plant Sci. 2018, 9, 1690. [CrossRef]

185. Hyten, D.L.; Pantalone, V.R.; Sams, C.E.; Saxton, A.M.; Landau-Ellis, D.; Stefaniak, T.R.; Schmidt, M.E. Seed quality QTL in a prominent soybean population. Theor. Appl. Genet. 2004, 109, 552–561. [CrossRef]

186. Zhang, K.; Liu, S.; Li, W.; Liu, S.; Li, X.; Fang, Y.; Zhang, J.; Wang, Y.; Xu, S.; Zhang, J.; et al. Identification of QTNs Controlling Seed Protein Content in Soybean Using Multi-Locus Genome-Wide Association Studies. Front. Plant Sci. 2018, 9, 1690. [CrossRef]

187. Huang, S.; Yu, J.; Li, Y.; Wang, J.; Wang, X.; Qi, H.; Xu, M.; Qin, H.; Yin, Z.; Mei, H.; et al. Identification of soybean genes related to soybean seed protein content based on quantitative trait loci collinearity analysis. J. Agric. Food Chem. 2019, 67, 258–274. [CrossRef]

188. Lee, S.; Van, K.; Sung, M.; Nelson, R.; LaMantia, J.; McHale, L.K.; Mian, M.A. Genome-wide association study of seed protein, oil and amino acid contents in soybean from maturity groups I to IV. Theor. Appl. Genet. 2019, 132, 1639–1659. [CrossRef]

189. Li, D.; Zhao, X.; Han, Y.; Li, W.; Xie, F. Genome-wide association mapping for seed protein and oil contents using a large panel of soybean accessions. Genomics 2019, 111, 90–96. [CrossRef] [PubMed]

190. Prenger, E.M.; Ostezan, A.; Mian, M.A.; Stupar, R.M.; Glenn, T.; Li, Z. Identification and characterization of a fast-neutron-induced mutant with elevated seed protein content in soybean. Theor. Appl. Genet. 2019, 132, 2965–2983. [CrossRef] [PubMed]

191. Qin, J.; Shi, A.; Song, Q.; Li, S.; Wang, F.; Cao, Y.; Ravelombola, W.; Song, Q.; Yang, C.; Zhang, M. Genome-wide association study and genomic selection of amino acid concentrations in soybean seeds. Front. Plant Sci. 2019, 10, 1445. [CrossRef]

192. Whiting, R.M.; Torabi, S.; Lukens, L.; Eskandari, M. Genomic regions associated with important seed quality traits in food-grade soybeans. BMC Plant Biol. 2020, 20, 1–4. [CrossRef] [PubMed]

193. Li, X.; Wang, P.; Zhang, K.; Liu, S.; Qi, Z.; Fang, Y.; Wang, Y.; Tian, X.; Song, J.; Wang, J.; et al. Fine mapping QTL and mining genes for protein content in soybean by the combination of linkage and association analysis. Theor. Appl. Genet. 2021, 134, 1095–1122. [CrossRef]

194. Arnold, B.; Menke, E.; Mian, M.A.; Song, Q.; Buckley, B.; Li, Z. Mining QTLs for elevated protein and other major seed composition traits from diverse soybean germplasm. Mol. Breed. 2021, 41, 1–8. [CrossRef]

195. Yuan, W.; Wu, Z.; Zhang, Y.; Yang, R.; Wang, H.; Kan, G.; Yu, D. Genome-wide association studies for sulfur-containing amino acids in soybean seeds. Euphytica 2021, 217, 155. [CrossRef]

196. Zhu, X.; Leiser, W.L.; Hahn, V.; Würschem, T. Identification of QTL for seed yield and agronomic traits in 944 soybean (Glycine max) RILs from a diallel cross of early-maturing varieties. Plant Breed. 2021, 140, 254–266. [CrossRef]

197. Wang, J.; Mao, L.; Zeng, Z.; Yu, X.; Lian, J.; Feng, J.; Yang, W.; An, J.; Wu, H.; Zhang, M.; et al. Genetic mapping high protein content QTL from soybean ‘Nanjixiaozu 25’and candidate gene analysis. BMC Plant Biol. 2021, 21, 388. [CrossRef]

198. Ravelombola, F.; Chen, P.; Vuong, T.; Nguyen, H.; Mian, R.; Acuña, A.; Flores-Palacios, L.; Wu, C.; Harrison, D.; De Oliveira, M.; et al. Genetics of seed protein and oil inherited from “BARC-7” soybean in two F2-derived mapping populations. J. Crop Improv. 2022, 1–17. [CrossRef]
223. Chiaiese, P.; Ohkama-Ohtsu, N.; Molvig, L.; Godfrey, R.; Dove, H.; Hocart, C.; Fujiwara, T.; Higgins, T.J.V.; Tabe, L.M. Sulphur and nitrogen nutrition influence the response of chickpea seeds to an added, transgenic sink for organic sulphur. *J. Exp. Bot.* 2004, 55, 1889–1901. [CrossRef] [PubMed]

224. Molvig, L.; Tabe, L.M.; Eggum, B.O.; Moore, A.E.; Craig, S.; Spencer, D.; Higgins, T.J.V. Enhanced methionine levels and increased nutritive value of seeds of transgenic lupins (*Lupinus angustifolius* L.) expressing a sunflower seed albumin gene. *Proc. Natl. Acad. Sci. U.S.A.* 1997, 94, 8393–8398. [CrossRef]

225. Tabe, L.M.; Droux, M. Limits to sulfur accumulation in transgenic lupin seeds expressing a foreign sulfur-rich protein. *Plant Physiol.* 2002, 128, 1137–1148. [CrossRef]

226. Dinkins, R.D.; Reddy, M.S.S.; Meurer, C.A.; Yan, B.; Trick, H.; Thibaud-Nissen, F.; Finer, J.J.; Parrott, W.A.; Collins, G.B. Increased sulfur amino acids in soybean plants over expressing the maize 15 kDa zein protein. *In Vitro Cell. Dev. Biol. Plant* 2001, 37, 742–747. [CrossRef]

227. Li, Z.; Meyer, S.; Essig, J.S.; Liu, Y.; Schapaugh, M.A.; Muthukrishnan, S.; Hainline, B.E.; Trick, H.N. High-level expression of maize gamma-zein protein in transgenic soybean (*Glycine max*). *Mol. Breed.* 2005, 16, 11–20. [CrossRef]

228. Kim, W.S.; Sun-Hyung, J.; Oehrle, N.W.; Jez, J.M.; Krishnan, H.B. Overexpression of ATP sulfurylase improves the sulfur amino acid content, enhances the accumulation of Bowman–Birk protease inhibitor and suppresses the accumulation of the β-subunit of β-cytochrome c in soybean seeds. *Sci. Rep.* 2020, 10, 1–13. [CrossRef]

229. Tabe, L.; Wirtz, M.; Molvig, L.; Droux, M.; Hell, R. Overexpression of serine acetyltransferase produced large increases in O-acetylserine and free cysteine seeds of a grain legume. *J. Exp. Bot.* 2010, 61, 721–733. [CrossRef] [PubMed]

230. Zhang, Y.; Schernthaner, J.; Labbé, N.; Hefford, M.A.; Zhao, J.; Simmonds, D.H. Improved protein quality in transgenic soybean expressing a de novo synthetic protein, MB-16. *Transgenic Res.* 2014, 3, 1–13. [CrossRef] [PubMed]

231. Ku, H.K.; Ha, S.H. Improving nutritional and functional quality by genome editing of crops: Status and perspectives. *Front. Plant Sci.* 2020, 11, 1514. [CrossRef] [PubMed]

232. Della Coletta, R.; Qui, Y.; Ou, S.; Hufford, M.B.; Hirsch, C.N. How the pan-genome is changing crop genomics and improvement. *Genome Biol.* 2021, 22, 1–19. [CrossRef]

233. Varshney, R.K.; Roorkiwal, M.; Sun, S.; Bajaj, P.; Chitikineni, A.; Thudi, M.; Singh, N.P.; Du, X.; Upadhyaya, H.D.; Khan, A.W.; et al. A chickpea genetic variation map based on the sequencing of 3366 genomes. *Nature* 2021, 599, 622–627. [CrossRef] [PubMed]

234. Zhao, J.; Bayer, P.E.; Ruperao, P.; Saxena, R.K.; Khan, A.W.; Golicz, A.A.; Nguyen, H.T.; Batley, J.; Edwards, D.; Varshney, R.K. Trait associations in the pangenome of pigeon pea (*Cajanus cajan*). *Plant Biotechnol. J.* 2020, 18, 1946–1954. [CrossRef]

235. Liu, C.; Wang, Y.; Peng, J.; Fan, B.; Xu, D.; Wu, J.; Cao, Z.; Gao, Y.; Wang, X.; Li, S.; et al. High-quality genome assembly and pan-genome studies facilitate genetic discovery in mungbean and its improvement. *Plant Commun.* 2022. [CrossRef]

236. Lee, H.; Cho, B.K.; Kim, M.S.; Lee, W.H.; Tewari, J.; Bae, H.; Sohn, S.I.; Chi, H.Y. Prediction of crude protein and oil content of soybeans using Raman spectroscopy. *Sens. Actuators B Chem.* 2013, 185, 694–700. [CrossRef]

237. Jasinski, S.; Lécureuil, A.; Durandet, M.; Bernard-Moulin, P.; Guerche, P. Arabidopsis seed content QTL mapping using high-throughput phenotyping: The assets of near infrared spectroscopy. *Front. Plant Sci.* 2016, 7, 1682. [CrossRef]

238. Sun, D.; Cen, H.; Weng, H.; Lan, W.; Abdalla, A.; El-Manawy, A.I.; Zhu, Y.; Zhao, N.; Fu, H.; Tang, J.; et al. Using hyperspectral analysis as a potential high throughput phenotyping tool in GWAS for protein content of rice quality. *Plant Methods* 2019, 15, 54. [CrossRef] [PubMed]

239. Zhang, S.; Hao, D.; Zhang, S.; Zhang, D.; Wang, H.; Du, H.; Kan, G.; Yu, D. Genome-wide association mapping for protein, oil and water-soluble protein contents in soybean. *Mol. Genet. Genom. Genomics* 2021, 91, 91–102. [CrossRef] [PubMed]

240. Roth, L.; Barendregt, C.; Bétrix, C.A.; Hund, A.; Walter, A. High-throughput field phenotyping of soybean: Spotting an ideotype. *Remote Sens. Environ.* 2022, 269, 112797. [CrossRef]

241. Choung, M.G.; Baek, I.Y.; Kang, S.T.; Han, W.Y.; Shin, D.C.; Moon, H.P.; Kang, K.H. Determination of protein and oil contents in soybean seed by near infrared reflectance spectroscopy. *Korean J. Crop Sci.* 2001, 46, 106–111. [CrossRef]

242. Choung, M.G.; Baek, I.Y.; Kang, S.T.; Han, W.Y.; Shin, D.C.; Moon, H.P.; Kang, K.H. Non-destructive method for selection of soybean lines containing high protein and oil by near infrared reflectance spectroscopy. *Korean J. Crop Sci.* 2001, 46, 401–406. [CrossRef]

243. Meeuwissen, T.H.; Hayes, B.J.; Goddard, M.E. Prediction of Total Genetic Value Using Genome-wide Dense Marker Maps. *Genetics* 2001, 157, 1819–1829. [CrossRef]

244. Li, Y.; Ruperao, P.; Batley, J.; Edwards, D.; Khan, T.; Colmer, T.D.; Pang, J.; Siddique, K.H.M.; Sutton, T. Investigating drought tolerance in chickpea using genome-wide association mapping and genomic selection based on whole-genome resequencing data. *Front. Plant Sci.* 2018, 9, 190. [CrossRef]

245. Keller, B.; Ariza-Suarez, D.; De la Hoz, J.; Aparicio, J.S.; Portilla-Benavides, A.E.; Buendia, H.E.; Mayor, V.M.; Studer, B.; Raatz, B. Genomic prediction of agronomic traits in common bean (*Phaseolus vulgaris* L.) under environmental stress. *Front. Plant Sci.* 2020, 11, 1001. [CrossRef]

246. Ravelombola, W.; Shi, A.; Huynh, B.L. Loci discovery, network-guided approach, and genomic prediction for drought tolerance index in a multi-parent advanced generation intercross (MAGIC) cowpea population. *Hort. Res.* 2021, 8, 24. [CrossRef]

247. Annicchiarico, P.; Nazzicari, N.; Laouar, M.; Ami-Alami, I.; Romani, M.; Pecetti, L. Development and proof-of-concept application of genome-enabled selection for pea grain yield under severe terminal drought. *Int. J. Mol. Sci.* 2020, 21, 2414. [CrossRef]
249. Annicchiarico, P.; Nazzicari, N.; Pecetti, L.; Romani, M.; Ferrari, B.; Wei, Y.; Brummer, E.C. GBS-based genomic selection for pea grain yield under severe terminal drought. *Plant Genome* 2017, 10, 2. [CrossRef] [PubMed]

250. Diaz, S.; Ariza-Suarez, D.; Ramdeen, R.; Aparicio, J.; Arunachalam, N.; Hernandez, C.; Diaz, H.; Ruiz, H.; Piepho, H.P.; Raatz, B. Genetic architecture and genomic prediction of cooking time in common bean (*Phaseolus vulgaris* L.). *Front. Plant Sci.* 2021, 11, 2257. [CrossRef] [PubMed]

251. Crosta, M.; Nazzicari, N.; Ferrari, B.; Pecetti, L.; Russi, L.; Romani, M.; Cabassi, G.; Cavalli, D.; Marocco, A.; Annicchiarico, P. Pea Grain Protein Content Across Italian Environments: Genetic Relationship with Grain Yield, and Opportunities for Genome-Enabled Selection for Protein Yield. *Front. Plant Sci.* 2021, 12, 718713. [CrossRef]

252. Watson, A.; Ghosh, S.; Williams, M.J.; Cuddy, W.S.; Simmonds, J.; Rey, M.D.; Asyraf Md Hatta, M.; Hinchliffe, A.; Steed, A.; Reynolds, D.; et al. Speed breeding is a powerful tool to accelerate crop research and breeding. *Nat. Plants* 2018, 4, 23–29. [CrossRef] [PubMed]

253. Crosser, J.S.; Pazos-Navarro, M.; Bennett, R.G.; Tschirren, S.; Edwards, K.; Erskine, W.; Creasy, R.; Ribalta, F.M. Time to flowering of temperate pulses in vivo and generation turnover in vivo–in vitro of narrow-leaf lupin accelerated by low red to far-red ratio and high intensity in the far-red region. *Plant Cell. Tissue Organ Cult.* 2016, 127, 591–599. [CrossRef]

254. Samineni, S.; Sen, M.; Sajja, S.B.; Gaur, P.M. Rapid generation advance (RGA) in chickpea to produce up to seven generations per year and enable speed breeding. *Crop J.* 2020, 8, 164–169. [CrossRef] [PubMed]

255. Cazzola, F.; Bermejo, C.J.; Gatti, I.; Cointry, E. Speed breeding in pulses: An opportunity to improve the efficiency of breeding programs. *Crop Pasture Sci.* 2021, 72, 165–172. [CrossRef]

256. Cazzola, F.; Bermejo, C.J.; Guindon, M.F.; Cointry, E. Speed breeding in pea (*Pisum sativum* L.), an efficient and simple system to accelerate breeding programs. *Euphytica* 2020, 216, 178. [CrossRef]

257. Jähne, F.; Hahn, V.; Würschum, T.; Leiser, W.L. Speed breeding short-day crops by LED-controlled light schemes. *Theor. Appl. Genet.* 2020, 133, 2335–2342. [CrossRef]

258. Moose, S.P.; Dudley, J.W.; Rocheford, T.R. Maize selection passes the century mark: A unique resource for 21st century genomics. *Trends Plant Sci.* 2004, 9, 358–364. [CrossRef]

259. Wright, I.J.; Reich, P.B.; Westoby, M.; Ackerly, D.D.; Baruch, Z.; Bongers, F.; Cavender-Bares, J.; Chapin, T.; Cornelissen, J.H.; Diemer, M.; et al. The worldwide leaf economics spectrum. *Nature* 2004, 428, 821–827. [CrossRef] [PubMed]

260. Reich, P.B. The world-wide ‘fast–slow’ plant economics spectrum: A traits manifesto. *J. Ecol.* 2014, 102, 275–301. [CrossRef]

261. Baraloto, C.; Timothy Paine, C.E.; Poorter, L.; Beauchene, J.; Bonal, D.; Domenach, A.M.; Hérault, B.; Patiño, S.; Roggy, J.C.; Chave, J. Decoupled leaf and stem economics in rain forest trees. *Ecol. Lett.* 2010, 13, 1338–1347. [CrossRef]

262. Chave, J.; Coomes, D.; Jansen, S.; Lewis, S.L.; Swenson, N.G.; Zanne, A.E. Towards a worldwide wood economics spectrum. *Ecol. Lett.* 2009, 12, 351–366. [CrossRef]

263. Weemstra, M.; Mommer, L.; Visser, E.J.; van Ruijven, J.; Kuyper, T.W.; Mohren, G.M.; Sterck, F.J. Towards a multidimensional root trait framework: A tree root review. *New Phytol.* 2016, 211, 1159–1169. [CrossRef] [PubMed]

264. Sadas, V.O. Evolutionary aspects of the trade-off between seed size and number in crops. *Field Crops Res.* 2007, 100, 125–138. [CrossRef]

265. Nguyen, D.T.; Hayes, J.E.; Atieno, J.; Li, Y.; Baumann, U.; Pattison, A.; Bramley, H.; Hobson, K.; Roorkiwal, M.; Varshney, R.K.; et al. The genetics of vigour-related traits in chickpea (*Cicer arietinum* L.): Insights from genomic data. *Theor. Appl. Genet.* 2022, 135, 107–124. [CrossRef] [PubMed]

266. Nguyen, D.T.; Hayes, J.E.; Harris, J.; Sutton, T. Fine Mapping of a Vigor QTL in Chickpea (*Cicer arietinum* L.) Reveals a Potential Role for Ca4_TIFY4B in Regulating Leaf and Seed Size. *Front. Plant Sci.* 2022, 13, 829566. [CrossRef]