agronomy

MDPI

Article

Seed Protein Genetics Linked with Nitrogen and Phosphorus Translocation Efficiency in Soybean

Qingsong Zhao [1,2,†], Niannian Ma [1,†], Ruirui Li [1], Yongjia Zhong [1], Xinxin Li [1] and Hong Liao [1,*]

[1] Root Biology Center, College of Resources and Environment, Fujian Agriculture and Forestry University, Fuzhou 350002, China
[2] Hebei Laboratory of Crop Genetics and Breeding, National Soybean Improvement Center Shijiazhuang Sub-Center, Huang-Huai-Hai Key Laboratory of Biology and Genetic Improvement of Soybean, Ministry of Agriculture and Rural Affairs, Institute of Cereal and Oil Crops, Hebei Academy of Agricultural and Forestry Sciences, Shijiazhuang 050035, China
* Correspondence: hliao@fafu.edu.cn
† These authors contributed equally to this work.

Abstract: Soybean (*Glycine max* (L.) Merr.) is an important nutritional crop with high seed protein content. Production of high protein concentrations relies on sufficient nutrient supplies, especially of nitrogen (N) and phosphorus (P). Although the genetic basis for seed quality traits has been well studied, little information exists on any genetic connections between seed quality and nutrient supplies in soybean. Here, a recombinant inbred line (RIL) population of 179 progeny was generated using HC6 and JD17 as parents contrasting in seed quality and N and P translocation efficiencies. Seed protein and N and P translocation efficiencies were higher in HC6 than in JD17. Meanwhile, positive correlations were observed between seed protein content and translocation efficiency of N and P in RILs, implying that high N and P translocation efficiencies might facilitate seed protein accumulation. A genetic map was constructed using 5250 SNP markers covering a genetic distance of 3154.83 cM. A total of 6 loci for quality and 13 loci for N and P translocation efficiency were detected. Among them, two fragments on chromosome 6 and chromosome 20 contained multiple significant markers for both quality and N and P translocation efficiencies, with the respective observed LOD values ranging from 2.98 to 5.61, and 3.01 to 11.91, while the respective PVE values ranged from 8.2% to 13.9%, and 8.3% to 28.0%. Interestingly, one significant locus on chromosome 20 appears to be the product of a transposable element (TE) InDel in *Glyma.20G085100*, with progeny lacking the TE also exhibiting higher N and P translocation efficiencies, along with higher seed protein contents. Taken together, these results provide genetic evidence that increasing N and P translocation efficiencies may lead to increasing protein contents in soybean seeds. Furthermore, a TE InDel may be used as a genetic marker for breeding elite soybean cultivars with high protein content and N and P translocation efficiencies.

Keywords: soybean; translocation efficiency; nutrient; quality; protein; QTL

Citation: Zhao, Q.; Ma, N.; Li, R.; Zhong, Y.; Li, X.; Liao, H. Seed Protein Genetics Linked with Nitrogen and Phosphorus Translocation Efficiency in Soybean. *Agronomy* **2023**, *13*, 598. https://doi.org/10.3390/agronomy13020598

Academic Editor: Yulong Ren

Received: 30 December 2022
Revised: 12 February 2023
Accepted: 15 February 2023
Published: 19 February 2023

1. Introduction

Soybean (*Glycine max* (L.) Merr.) is one of the most economically important crops globally, owing largely to high seed protein and oil contents with 40% protein and 20% oil in seeds on a dry basis [1]. Over recent decades, as much as 56% of dietary protein consumed in human food and animal feed has been produced by soybean seeds [2,3], and global soybean production has reached 367.8 million metric tons. On the other hand, world protein meal consumption has reached 243.6 million metric tons and accounts for 70% [4,5]. Therefore, further clarification of seed protein genetics and methods for releasing elite soybean cultivars rich in seed protein are critical.

Genetic and molecular processes connected with soybean seed protein content have been well reported. In previous experiments, protein was found to be controlled more by

genetics than by environment, with heritability above 0.8. Meanwhile, seed protein content negatively correlated with oil content, which also had a heritability of over 0.8 [4]. With these high heritabilities for complex protein and oil synthesis processes, quantitative trait locus (QTL) and genome-wide association study (GWAS) approaches may then be effective means for detecting candidate loci for seed protein content to propagate in breeding programs. As of September 2022, 241 loci associated with seed protein content have been detected in biparental populations and listed on the Soybase website [6]. To verify genetic loci associations with seed protein content, a meta-analysis was conducted, and 55 meta-QTLs for protein and oil contents were putatively identified [7]. An additional set of loci were associated with protein accumulation using GWAS methods [8–11]. Among significant chromosomal purlieus, LG15, LG19, and LG 20 are important due to the presence of major or multiple loci within their bounds. Recently, individual genes with specific functions in seed trait development underlying protein or oil content have been cloned and analyzed, such as *GmSWEET39* on chromosome 15 [12] and *POWR1* on chromosome 20 [13]. However, neither protein nor oil can accumulate without nutrients, especially nitrogen (N) and phosphorus (P) [14]. To date, genetic associations between seed quality and nutrient supply have yet to be evaluated.

Nitrogen is generally considered to be the most limiting factor for plant growth and development in various ecosystems [15]. As a component of proteins, enzymes, nucleic acids, and plant growth regulators, this element is involved in many physiological and biochemical processes in plants [16,17]. Phosphorus, another crucial plant macronutrient, participates in plant cells as a structural constituent of essential biomolecules involved in energy metabolism [18] and in the formation of key macromolecules, such as nucleic acids and phospholipids [19]. The essential role of P in many aspects of cellular metabolism is also evident from the large amounts of P that are stored in seeds to enable embryo development, germination, and seedling growth [20].

Soybean seeds are rich in protein and oils, and, therefore, require more N and P than other crops. Most plant seeds contain lower concentrations of N and P than soybean seeds. For example, cotton seeds may hold average N concentrations of around 38 $mg \cdot g^{-1}$ [21], and rice seed P concentrations can average around 3.0 $mg \cdot g^{-1}$ [22]. Meanwhile, soybean seeds can contain 61 $mg \cdot g^{-1}$N [23] and 4.8 $mg \cdot g^{-1}$P [24]. Without sufficient N and P supplies, soybean plants produce fewer seeds, but also may decrease protein and oil contents in the seeds that are produced. It is very important to understand the relationship between soybean protein content and nitrogen and phosphorus translocation efficiency as defined by the ability to move nutrients from leaves to seeds to cultivate high-protein soybean cultivars.

As the most direct and largest source of nutrients, leaves provide most of the N and P to seeds [25]. Therefore, nutrient translocation efficiency is crucial for allocating sufficient resources to seed formation [26,27]. Researchers have found that most N in wheat grains (~80%) comes from leaves [28]. Senescent leaves are primary sources of N and P resorption, and the proportion of P re-absorbed is higher than that of N [29]. Nitrogen stored in soybean leaves begins to flow out from the R6 stage, and is predominantly transferred to seeds for use in synthesis of storage proteins [30]. Remobilization from senescing leaves is a key factor for efficient utilization of N [31]. Similarly, when P is scarce, plants remobilize P from old leaves to maintain plant growth and metabolism [32,33]. Both QTLs and molecular factors of nutrient translocation efficiency have been reported. Genetic loci of mineral element accumulation in soybean seeds have also been reported [34,35], including three QTL_S for P content on chromosomes 7, 12, and 17 that were detected using 916 SSR markers and 92 $F_{2:4}$ lines [24]. An advantageous gene for P absorption and utilization efficiency identified on chromosome 11 and named *PE1* also appears to improve P allocation to pods [27].

Although the genetic components of seed production, protein and oil quality, and N and P translocation efficiency have been studied separately, any existing relationships between them have rarely been reported and remain largely mysterious. In-depth obser-

vations are required to further decipher genetic and molecular connection basis between seed protein and translocation efficiencies of N and P. In this study, we aimed to analyze the genetic basis of three sets of traits in soybean using a RIL population bred from parents with significant differences in seed protein content and N and P translocation efficiency. The results provide the foundation for soybean marker-assisted breeding, and provide a theoretical reference for improvement in soybean seed quality.

2. Materials and Methods

2.1. Plant Materials and Growth Conditions of Field Trials

In this study, the variety HC6, with relatively high protein content and high N and P translocation efficiencies, was crossed with JD17, of relatively low protein content and low N and P translocation efficiencies, to develop a population comprising 176 F_7-derived recombinant inbred lines (RILs) using the single seed descent (SSD) method [27,36,37]. The RIL population was used to construct a genetic linkage map, and detect QTLs for N and P translocation efficiencies and quality traits. The RILs and parents were planted in a randomized completed block design (RCBD), with three replications in 2020 at Yangzhong Experimental Farm (118.20° E, 26.17° N) of Fujian Agriculture and Forestry University, Fuzhou City, China. The RILs and the two parental lines were sown in single row plots 1.2 m long, with 10 plants retained in each row spaced 0.5 m apart. Irrigation and pest control in the field followed local practices. Before sowing, the physical and chemical properties of the surface 20 cm of soils were measured as follows: pH, 4.64 ± 0.05; organic matter (%), 2.63 ± 0.03; available N, P (Bray I) and K (mg/kg), 123.21 ± 6.11, 179.30 ± 15.13, and 159.21 ± 12.95, respectively.

2.2. Plant Sampling and Trait Evaluation

Three representative soybean plants of each plot were harvested at the R6 stage. For each plot, the leaves and seeds were sampled separately and dried at 60 °C, then measured and weighed per plant. After ash digestion of leaves and seeds, the content of N and P was assessed using a continuous flow analyzer as previously reported [38]. Ten N- and P-related traits were measured, including N concentration of leaves (NcL), N concentration of seeds (NcS), P concentration of leaves (PcL), P concentration of seeds (PcS), total N content of leaves (TNL), total P content of leaves (TPL), N translocation efficiency (NTE), P translocation efficiency (PTE), difference in N concentration between seeds and leaves (NcD), and difference in P concentration between seeds and leaves (PcD). The translocation efficiencies NTE and PTE were calculated as follows:

$$\mathrm{N/PTE} = \frac{\mathrm{TN/PS}}{\mathrm{TN/PS} + \mathrm{TN/PL}}$$

where TN/PS and TN/PL were the total nutrient (either N or P) content in seeds and leaves per plant (mg·plant^{-1}), respectively. The total N (P) content means N or P of leaves (seeds) per plant, while concentrations means N or P content in unit biomass of leaves or seeds. Ten g of dry seeds from each plot was sampled, and protein and oil content was determined using near-infrared spectrometry (NIS) using a prior method [39]. Flower color at full flowering stage was used to verify the reliability of the genetic map for the investigated RIL population.

2.3. Genotyping by High-Throughput Sequencing

Genomic DNA was extracted from the two parents using the commercial extraction kit DNeasy Plant Mini Kit (QIAGEN, Venlo, Germany). GBS libraries were constructed and sequenced on an Illumina sequencing platform. In short, genomic DNA was fragmented to 100–300 bp sizes and sequenced on a Hiseq X10 PE150 sequencing system after terminal repair, splicing, isolation and purification, amplification, and enrichment. Following extraction, GBS libraries were digested by the EcoRI and NIaIII enzymes according to Elshire's protocol [40]. Genomic DNA was then extracted, and 100 ng of DNA was used

for each plant in digestion reactions with EcoRI and NIaIII (New England Biolabs, Ipswich, MA, USA) in 96-well plates. Digested products were mixed with A1 and A2 adaptors for addition to the two DNA ends. After mixing the 96-well library, 400–600 bp DNA fragments were screened out and DNA was purified using a PCR kit (NEB, Beijing, China). Finally, pooled library concentrations were adjusted to 10 nmol prior to sequencing with PE125 on a HiSeq4000 (Illumina, San Diego, CA, USA). Sequencing of parental and progeny GBS libraries was performed by Genedenovo Biotechnology Co., Ltd. (Guangzhou, China) using a HiSeq X10 NGS platform (Illumina, San Diego, CA, USA).

2.4. SNP Identification and Bin Map Construction

Using the Burrows–Wheeler Aligner (BWA), the parameter "MEM 4-K 32-M" was set to compare the 400–600 bp double-terminal sequence with the *Glycine Max* Wm82.A2.v1 reference genome in the Genome Analysis Toolkit (GATK) (-Window, 4–filter "QD < 2.0 | | FS > 60.0 | | MQ < 40.0, "-g_filter" GQ < 20 ") for SNP calling. Genotypes of all samples were determined using GATK's Unified Genotyper [41]. The physical position of each SNP was further determined by using the software tool ANNOVAR [42]. Polymorphic markers consistent with the JD17 genotype or HC6 genotype were recorded as "a" or "b", respectively, and ambiguous segregation patterns were recorded as missing data "−".

Offspring genotypes were screened for differences from parental genotypes to ensure that 1 kb sequences holding polymorphism markers were otherwise similar to at least 95% of individual plant genotypes using Chi-square ($\chi 2$) testing ($p \leq 0.05$). Highest average depth of sequencing was used to select sequences for further analysis in cases of multiple overlapping sections. A genetic linkage map was constructed using Join Map 4.1 for the derived markers. Colocalization analysis was performed using the Auto QTLs method in MapChart 2.32, with colocalized markers being those around graph maxima, and within the inner and outer intervals, defined as positions where the graph falls below the maximum minus the "inner threshold" and the "outer threshold", respectively [43].

2.5. QTL Detection and Statistical Analysis

MapQTL 6.0 was used to map markers for nutrient traits. MapQTL6.0 uses Interval Mapping (IM) and the multiple-QTL model (MQM) [44]. The position of each QTL was first checked by the interval mapping (IM) option, and then the markers with peak logarithm of odds (LOD) values were selected as cofactors using subsequent multiple QTL mapping (MQM) QTL detection. A LOD score threshold of 2.5 was set for declaring the presence of putative QTLs in given genomic regions. MapChart 2.32 software was used to draw the final QTL Map. Trait data were subjected to variance and QTL analysis. Analysis of variance (ANOVA) and Pearson correlation analysis were implemented using R software [45]. Genetic correlation was calculated using the softwares Plink and CGTA. Broad sense heritability (h^2b) was estimated for each trait in QTL ICIMapping V4.1 software using the following formula: h^2b = VG/(VG + VE), where VG is the genetic variance, and VE is the environmental variance [37].

3. Results

3.1. Comparison of Seed Protein and Oil Contents and Nutrient Translocation Efficiency Traits between the Parental Genotypes

In order to explore genetic loci regulating nitrogen (N) and phosphorus (P) translocation efficiency, along with seed protein and oil contents, the contrasting genotypes JD17 and HC6 were selected as parents to construct recombinant inbred lines (RILs). The results showed that seed protein content was higher and seed oil content lower in parent HC6 than in parent JD17 ($p < 0.05$) (Figure 1A). Interestingly, N and P translocation efficiencies all displayed similar trends as seed contents of protein (Figure 1A, B), indicating that higher N and P translocation efficiencies from leaves to seeds might lead to higher seed protein content. In correlation analysis using the population of 179 RILs, significant positive correlations were found between seed protein content and both N and P translocation efficiencies

(Figure S1), with *r* values of 0.34 (***) and 0.29 (***) observed for N and P translocation efficiency, respectively. At the same time, there was a negative correlation between oil and N and P translocation efficiencies, with *r* values of −0.16 (*) and −0.08 (*NS*), respectively. For genetic correlation analysis, *rG* between protein and N and P translocation efficiency was 0.77 (***) and 0.76 (***), respectively. *rG* between oil and N and P translocation efficiency was −0.40 (*) and −0.30, respectively (Table S1). Taken together, the N and P positive correlations with protein content and negative correlations with oil content indicate that seed protein content is intimately related to N and P translocation efficiencies. Therefore, RILs derived from these two soybean accessions were used to further study the genetic loci that control soybean N and P translocation efficiencies and seed protein content.

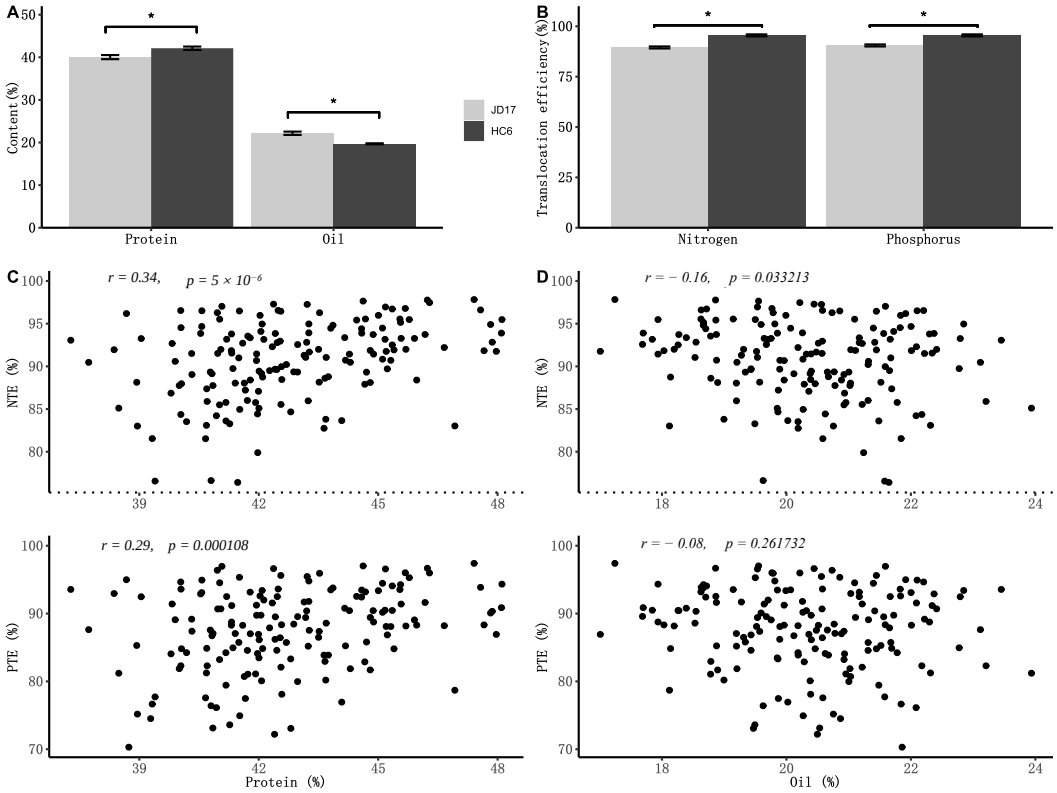

Figure 1. Comparison of seed protein and oil contents and nutrient translocation efficiency. (**A**) Protein and oil content of JD17 and HC6; (**B**) Nitrogen and phosphorus translocation efficiency of JD17 and HC6; (**C**) Correlation analysis between seed protein and oil contents and nitrogen translocation efficiency; (**D**) Correlation analysis between seed protein and oil contents and phosphorus translocation efficiency. NTE represents nitrogen translocation efficiency, and PTE represents phosphorus translocation efficiency. Asterisks denote significance of differences according to Student's *t* tests (* $p < 0.05$).

3.2. Phenotypic and Genetic Variation among RILs

To further study the genetic basis of N and P translocation efficiency traits, a total of 10 N and P traits were evaluated along with seed protein and oil contents using a population of 179 RILs derived from soybean parents HC6 and JD17 (Table 1). Considerable genetic variation was detected among the parents and RILs in all measured traits. The average CVs of leaf traits (NcL, PcL, TNL, and TPL) were greater than for seed traits (Pro, Oil, NcS, PcS). Extensive transgressive heritability among the 179 RILs was observed for each observed trait. All average values of traits fell between parent values, while maximum and minimum values lay beyond parental bounds, suggesting that both parents might contribute to phenotypic variation (Figure 2). According to histograms of traits, all traits followed continuous distribution (Figure 2). According to kurtosis and skew values

observed across the population, the peak of TNL and TPL fell in the left portion of the distribution, while the peak of NTE and PTE fell in the right portion of the distribution. Results above indicate that these traits are inherited as multiple loci in a quantitative fashion (Table 1, Figure 2). The ANOVA test showed that lines were the main reason for variation (Table S2). In addition, the broad sense heritability of the observed traits ranged from 0.73 to 0.91 (Table 1), suggesting that genetic variation accounted for most of the observed phenotypic variation, and location effects were relatively minor.

Table 1. Phenotypic variation and genetic analysis of nutrient and quality traits in 179 soybean recombination inbred lines.

Traits	Parents		RILs							
	JD17	HC6	Min.	Max.	Mean	SD	CV (%)	Kurt.	Skew.	h^2_b
Pro	40.04	42.13	37.28	50.84	42.76	2.45	5.72	0.36	0.49	0.90
Oil	22.18	19.71	16.28	24.28	20.39	1.38	6.76	−0.07	−0.18	0.88
NcS	66.73	70.28	52.47	79.86	65.81	4.79	7.28	−0.09	0.12	0.89
PcS	6.89	7.99	5.63	8.78	7.36	0.55	7.47	0.07	−0.03	0.91
NcL	21.81	16.25	10.15	24.55	16.03	2.90	18.08	0.11	0.49	0.75
PcL	2.65	2.25	1.39	4.41	2.40	0.53	22.26	1.15	0.76	0.75
TNL	189.98	154.96	18.23	449.58	116.82	82.19	70.36	3.74	1.80	0.84
TPL	23.12	21.51	2.46	52.19	16.57	9.78	59.01	1.35	1.20	0.82
NcD	44.92	54.02	28.77	63.31	49.84	5.70	36.18	0.36	−0.11	0.87
PcD	4.24	5.74	2.50	6.87	4.98	0.74	55.48	0.55	−0.20	0.85
NTE	0.84	0.93	0.64	0.98	0.90	0.06	6.83	2.75	−1.49	0.74
PTE	0.82	0.92	0.58	0.97	0.87	0.07	8.05	1.92	−1.19	0.73

RILs: recombinant inbred lines. NcL, PcL: nitrogen and phosphorus concentrations in leaves (mg/g); TNL, TPL: nitrogen and phosphorus total content in leaves (mg); NcS, PcS: nitrogen and phosphorus concentrations in seeds (mg/g); NTE, PTE: nitrogen and phosphorus translocation efficiency; NcD, PcD: nitrogen and phosphorus concentration difference between seeds and leaves. Pro, Oil: protein and oil content in seeds (%).

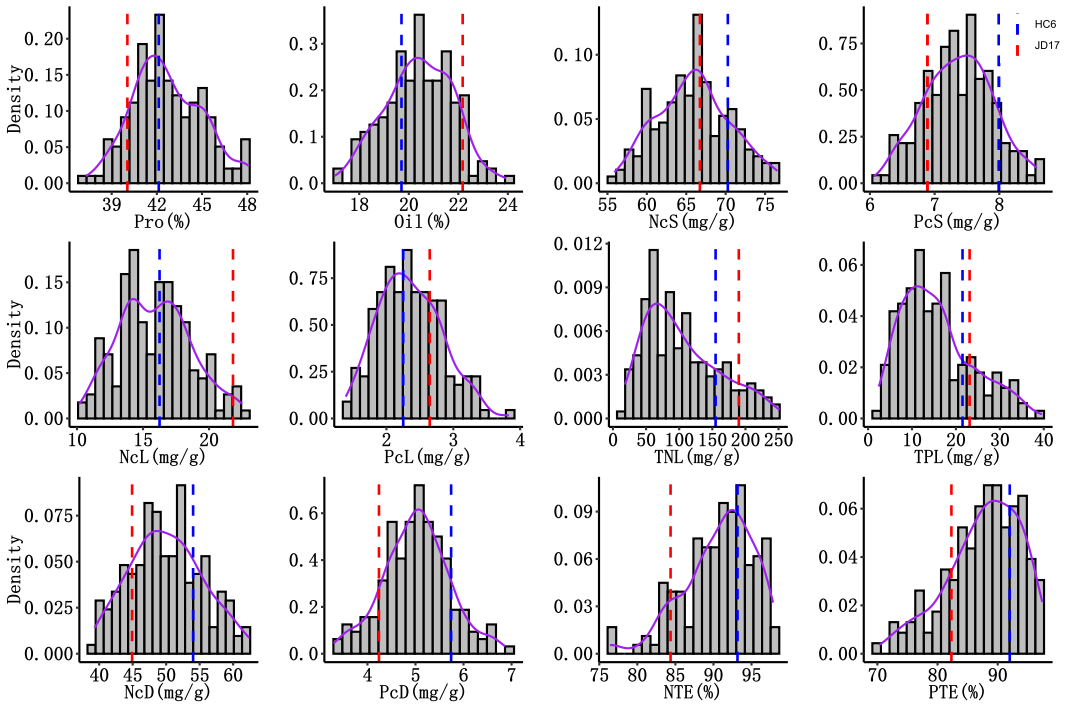

Figure 2. Frequency distribution of quality and N and P translocation efficiency traits among RIL population containing 179 lines. The red and blue dashed lines represent the value of traits for JD17 and HC6, respectively. Pro, Oil: protein and oil content in seeds (%); NcS, PcS: nitrogen and phosphorus concentrations in seeds (mg/g); NcL, PcL: nitrogen and phosphorus concentrations in leaves (mg/g); TNL, TPL: nitrogen and phosphorus total content in leaves (mg); NcD, PcD: nitrogen and phosphorus concentration difference between seeds and leaves; NTE, PTE: nitrogen and phosphorus translocation efficiency.

3.3. Construction of Genetic Linkage Map

A total of 5250 SNPs with polymorphism between parents were used to construct a RIL population genetic map (Table 2) covering a total genetic distance of 3154.833 cM at an average coverage distance of 0.642 cM. Chromosome 5 had the largest average coverage distance of 1.139 cM, which held 132 molecular markers over a genetic distance of 150.356 cM. The average coverage distance of chromosome 18 was the smallest of those observed at 0.29 cM, containing 500 molecular markers over a genetic distance of 144.744 cM (Table 2). The above results indicate that the genetic map constructed in this study has a small average coverage distance in a high resolution, and uniformly distributed context.

Table 2. Genetic and physical map with recombination rates of population containing 179 RILs derived from the cross HC6 × JD17.

	Number of Markers	Start (bp)	Stop (bp)	Genetic Distance (cM)	Physical Size (bp)	Average Coverage Distance (cM)	Recombination Rate (cM/Mb)
Chr.01	256	289,316	54,169,376	139.683	53,880,060	0.546	2.592
Chr.02	302	331,980	46,721,845	223.733	46,389,865	0.741	4.823
Chr.03	187	162,171	42,036,871	135.455	41,874,700	0.724	3.235
Chr.04	262	560,706	51,913,854	144.989	51,353,148	0.553	2.823
Chr.05	132	74,227	42,218,304	150.356	42,144,077	1.139	3.568
Chr.06	264	105,140	51,169,402	157.940	51,064,262	0.598	3.093
Chr.07	205	176,838	44,351,629	174.225	44,174,791	0.850	3.944
Chr.08	229	133,699	47,621,113	187.701	47,487,414	0.820	3.953
Chr.09	369	24,932	47,374,864	157.685	47,349,932	0.427	3.330
Chr.10	294	166,537	51,028,020	165.823	50,861,483	0.564	3.260
Chr.11	278	28,686	34,363,362	168.111	34,334,676	0.605	4.896
Chr.12	152	105,680	39,977,366	128.184	39,871,686	0.843	3.215
Chr.13	294	935,103	43,521,095	179.917	42,585,992	0.612	4.225
Chr.14	281	291,967	47,851,291	169.878	47,559,324	0.605	3.572
Chr.15	234	1,974,828	50,562,456	114.326	48,587,628	0.489	2.353
Chr.16	280	12,481	37,757,414	150.733	37,744,933	0.538	3.993
Chr.17	248	1,847,759	41,500,612	161.743	39,652,853	0.652	4.079
Chr.18	500	5391	57,905,267	144.744	57,899,876	0.289	2.500
Chr.19	226	81,698	49,599,441	156.191	49,517,743	0.691	3.154
Chr.20	257	2,789,679	47,717,330	143.416	44,927,651	0.558	3.192
Min	132	5391	34,363,362	114.326	3,433,4676	0.289	2.353
Max	500	2,789,679	57,905,267	223.733	5,789,9876	1.139	4.896
Mean	262.5			157.741	45,963,104	0.642	3.490
Total	5250			3154.833	919,262,094		

The genetic distances and reference physical distances were compared to assess the quality of our linkage map, with the result that genetic distance and physical distance exhibited good collinearity in both heterochromatic and euchromatinic regions of chromosomes (Figure S2). Overall, physical locations were better than genetic distance at finding candidate genes near QTLs. To determine whether the genetic maps constructed in this study were capable of QTL mapping, the flower color quality trait was also mapped in the RIL population (Figure 3). Here, the primary QTL for flower color was identified with an LOD value of 13.25 and a PVE value of 30% on chromosome 13 near the physical location of 16,983,990 bp, where the *W* gene responsible for flower color has previously been reported [46]. Taken together, these results indicate that the genetic linkage map constructed here would be useful for further QTL mapping of seed quality and nutrient translocation efficiency traits.

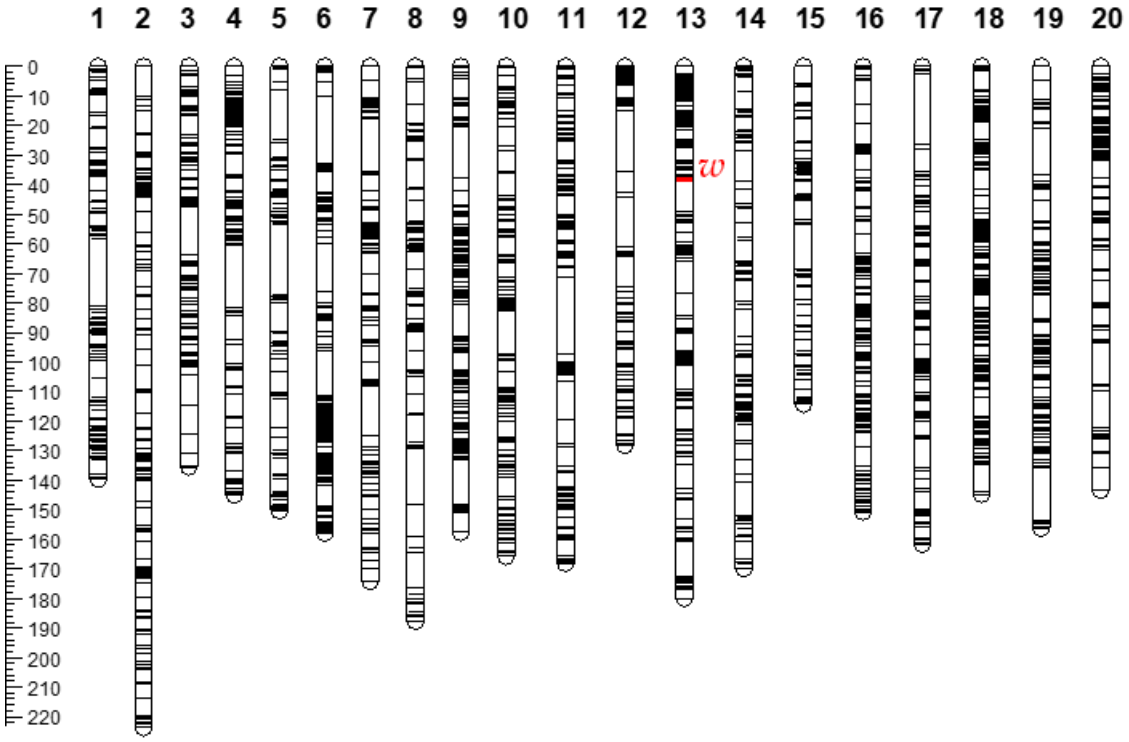

Figure 3. Genetic map of RIL population developed from a cross between HC6 × JD17. Genetic map was constructed in Join Map 4.1 and drawn with MapChart 2.3. The scale on the left represents the genetic distance between bin markers using centimorgans as the unit. Letter "*w*" represents the flower locus. Numbers at the top represent the chromosome number of soybean.

3.4. QTL Identification for N and P Translocation Efficiency and Seed Quality Traits

Using the suitable high-resolution genetic map described above, QTLs were identified for translocation efficiencies of N and P, along with seed protein and oil contents (Table 3). In the end, a total of 41 loci were detected for 12 traits, including two protein QTLs, three oil QTLs, two NTE QTLs, three PTE QTLs, five NcD QTLs, two PcD QTLs, and twenty-two other nutrient-related QTLs. The LOD values ranged from 2.5 to 11.91, and the PVE values ranged from 6.4 to 28.0. A marker for difference in P between seeds and leaves, *qPcD20*, had the highest LOD and PVE values of all markers for all measured traits, and the additive effect (Add) was −0.37.

Table 3. Putative QTLs for quality and N and P translocation efficiency traits detected in a population derived from an HC6×JD17 cross containing 179 RILs.

Trait	QTL	Chr	Position (cM)	Locus/Interval	LOD	Add	PVE (%)
NcL	*qNcL10*	10	137.076	10_45243194	2.8	0.75	7.1
	qNcL13	13	88.639	13_27737833	2.51	0.72	6.4
	qNcL16	16	120.329	16_32651304	2.68	0.74	6.8
	qNcL18	18	102.611	18_55034166	2.84	0.77	7.2
PcL	*qPcL6*	6	116.324–127.161	06_18481680–06_45055022	5.29	0.18	13.2
	qPcL18	18	28.701	18_3930525	2.98	0.14	7.6
TNL	*qTNL6*	6	118.453	06_19585884	9.9	−27.94	23.6
	qTNL10	10	137.076	10_45243194	7.64	24.55	18.8
	qTNL13	13	25.37	13_10785045	2.71	−15.08	7.1

Table 3. *Cont.*

Trait	QTL	Chr	Position (cM)	Locus/Interval	LOD	Add	PVE (%)
TPL	*qTPL6*	6	118.453	06_19585884	6.85	−3.42	16.8
	qTPL10.1	10	7.45	10_4289419	3.23	−2.40	8.3
	qTPL10.2	10	78.479	10_18967468	2.98	−2.32	7.7
	qTPL10.3	10	137.076	10_45243194	5.69	3.09	14.2
	qTPL11	11	125.659	11_31189747–11_32037532	2.67	−2.32	6.9
	qTPL13	13	71.913	13_22901190–13_26865258	2.58	−2.41	6.7
	qTPL19	19	40.956	19_5501220	2.59	−2.20	6.7
NcS	*qNcS6*	6	111.393–124.78	06_16674718–06_44270622	4.96	1.64	12.2
	qNcS10	10	138.232	10_45550230	4.25	−1.50	10.6
	qNcS18	18	9.891	18_472964	2.98	−1.30	7.5
	qNcS20	20	8.311	20_31160346	5.1	−1.67	12.6
PcS	*qPcS6*	6	111.393	06_16674718	6.48	0.22	15.9
	qPcS20	20	10.02	20_31536419	9.46	−0.26	22.4
NTE	*qNTE6*	6	117.276	06_18864382	2.99	1.34	8.4
	qNTE10	10	142.797	10_45640909–10_46681050	5.92	−1.99	15.8
PTE	*qPTE6*	6	116.83	06_18854552	2.98	1.77	8.2
	qPTE10	10	149.904	10_47584960	4.48	−2.15	12.0
	qPTE20	20	8.311	20_31160346	3.01	−1.84	8.3
NcD	*qNcD6*	6	111.393	06_16674718	3.72	1.71	9.5
	qNcD10.1	10	97.544	10_39222550	2.94	−1.52	7.6
	qNcD10.2	10	138.232	10_45550230	8.03	−2.42	19.4
	qNcD18	18	12.043	18_883579	2.68	−1.47	7.0
	qNcD20	20	9.646	20_31274128–20_31536419	6.65	−2.27	16.4
PcD	*qPcD2*	2	220.034	02_45085208	2.64	−0.18	7.0
	qPcD12	12	0.504	12_286008	2.5	−0.18	6.7
	qPcD20	20	10.02	20_31536419	11.91	−0.37	28.0
Pro	*qPro6*	6	112.474–118.676	06_17296433–06_20707785	5.61	0.89	13.9
	qPro20	20	10.02	20_31536419	8.87	−1.08	21.0
Oil	*qOil5*	5	150.356	05_42079283	4.55	0.47	11.3
	qOil6	6	118.229–124.498	06_19369334–06_43439776	4.23	−0.45	10.6
	qOil9	9	31.217	09_1558594–09_2811050	4.03	0.52	10.1
	qOil20	20	10.243	20_31540803	5.49	0.51	13.5

ADD values of >0 and <0 represent increasing effects of the QTLs derived from JD17 and HC6, respectively.

Two QTLs were detected for seed protein content (Table 3). One was *qPro6*, with an LOD of 5.61, an Add of 0.89, and a PVE of 13.9, and was located on chromosome 6 in the interval 06_17296433–06_20707785. The other was *qPro20*, with an LOD of 8.87, an Add of −1.08, and a PVE of 21.0, and was located on chromosome 20 near 20_31536419. The Add value of *qPro20* was less than 0, indicating that the favorable gene comes from HC6 while the favorable gene of *qPro6* comes from JD17, according to the negative Add value. There were four QTLs detected for seed oil content with LOD values ranging from 4.03 to 5.49, and PVE values of 10.1 to 13.5 (Table 3). Interestingly, two of the four detected seed oil markers were found near the two protein loci, while the other two seed oil QTLs were found on chromosome 5 and chromosome 9. All favorable seed oil genes originated with JD17, except for *qOil6* on chromosome 6, with the favorable gene from HC6.

For N translocation efficiency, two significant QTLs loci were detected (Table 3). One on chromosome 6, *qNTE6*, was detected with an LOD of 2.99, an Add of 1.34, and a PVE of 8.4. The other, *qNTE10*, was found on chromosome 10 and had an LOD of 5.92, an Add of −1.99, and a PVE of 15.8. Here, again, JD17 provided the favorable gene for locus *qNTE6* on chromosome 6, and HC6 provided the favorable gene for locus *qNTE10*. For P translocation efficiency, the three significant QTLs, *qPTE6*, *qPTE10*, and *qPTE20*, were detected on chromosomes 6, 10, and 20, respectively (Table 3). The LOD values for PTE loci ranged from 2.98 to 4.48, and the PVE values ranged from 8.2 to 12.0. According to their Add values, PTE favorable genes were contributed by JD17 for *qPTE10* and *qPTE20*, and by HC6 for *qPTE6*.

Five QTLs were detected for the differences in N concentration between seed and leaves (Table 3). Single representatives were detected on chromosomes 6, 18, and 20 and were named according to the conventions used above along with two other QTLs identified on chromosome 10 that were labelled *qNcD10.1* and *qNcD10.2*. The NcD LOD values ranged from 2.68 to 8.03, while the PVE values ranged from 7.0 to 16.4. All of the NcD favorable genes came from HC6, except for *qNcD6*, which was another protein- and N-related locus on chromosome 6 with the favorable allele contributed by JD17. *qNcD10.2* had the greatest LOD and PVE values. For difference in P concentration between seed and leaves, three significant QTLs were detected that were located on chromosomes 2, 12, and 20. The LOD values ranged from 2.50 to 11.91, and the PVE values ranged from 7.0 to 28.0. All favorable PcD genes came from HC6. For the other nutrient traits, four NcS QTLs, two PcS QTLs, four NcL QTLs, two PcL QTLs, three TNL QTLs and six TPL QTLs were detected and placed on chromosomes 6, 10, 11, 12, 16, 18, and 19. The LOD values ranged from 2.51 to 9.9, and the PVE values ranged from 6.4 to 23.6. Among them, *qTNL6* had the greatest LOD and PVE values.

Overall, a number of major QTLs associated with protein, oil, and nutrient translocation efficiency were detected. Chromosomes 6, 10, and 20 were represented across multiple traits, and there appear to be interesting regions of colocalization. The set of loci detected herein provides numerous targets for potential future study, particularly in regions of colocalization.

3.5. Loci of Colocalized Seed Quality and Nutrient Translocation Efficiency Traits in Soybean

In this study, two major QTL regions for seed protein were identified and named *qPro6* and *qPro20.* In addition, QTLs for N and P translocation efficiency were identified near *qPro6* and *qPro20* (Figure 4). *qPro6* is located between 17,296,433 bp and 20,707,785 bp, in a genetic interval of 6.202 cm where *qNTE6* and *qPTE6* were also identified (Figure 4, Table 3). Positive Add values mean that favorable alleles were contributed by JD17, and there appeared to be some positive synergistic interactions between allele effects on traits, as suggested by the observed positive correlation between NTE, PTE, and protein content (Figure 1C). Furthermore, the seed P and N concentration loci *qPcS6* and *qNcS6*, and the leaf P concentration locus *qPcL6* were all also detected within the recurring region of significance on chromosome 6. Here, the LOD value was higher for *qPcS6* than for *qNcS6* (Table 3), which suggests that P concentration effects might have been more impactful on protein production than N concentration effects for this locus under the given experimental conditions. In addition, correlation analysis also revealed clear positive correlations of protein with PcL, PcS, and NcS (Figure S1). Similarly, the locus *qOil6* was also located in the region on chromosome 6 that appears to hold loci of interest for many of the traits studied herein. Chromosome 20 appears to harbor an important region, with detailed analysis in this report showing that *qPro20* and *qOil20* localize near 31,540,803 bp on chromosome 20, which is close to *Glyma.20G085100* (31,774,769–31,779,804 bp), a newly reported potential region of candidate genes behind *cqSeed protein-003* QTL effects [47]. This is the same region in which *qPTE20, qNcS20, qPcS20, qNcD20, and qPcD20* were also detected in this study, with *qPcD20* returning the highest LOD value of 11.91 and a PVE value that was 28% greater than the PVE for *qNcD20*. Taken together, these results imply that concentration differences for P were more impactful on seed quality than those of N, and *Glyma.20G085100* is a strong candidate gene for imparting these effects in regulating nutrient translocation efficiency in soybean.

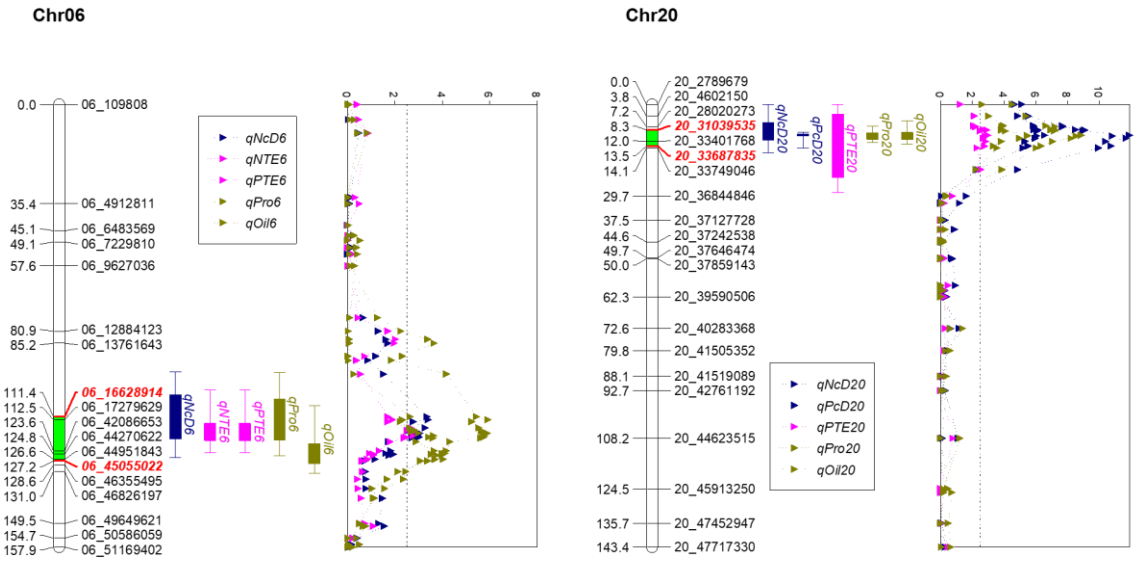

Figure 4. Colocalization for nutrient translocation efficiency and quality traits in RIL population. Genetic maps constructed using MapChart 2.32. The LOD graphs and linkage groups are shown together with a constant line indicating the LOD threshold of 2.5, while 1- and 2-LOD intervals are also shown.

3.6. Effects of Glyma.20G085100 on Nutrient Translocation Efficiency

The QTLs on chromosome 20 associated with seed protein and oils identified in this study co-localize with a newly reported major locus for seed protein and oil, with variation of the gene caused by a transposable element (TE) InDel in the candidate gene *Glyma.20G085100* [13,47]. In order to evaluate whether the QTL identified in this study is also caused by the TE Indel, fragments near the TE insertion site within the gene in each of the two parents were re-sequenced after PCR using the reported primers [47]. Here, the PCR fragment length was about 900 bp for JD17 and about 600 bp for HC6 (Figure 5A). In addition, re-sequencing results showed that HC6 appeared to experience a 321 bp deletion in comparisons with JD17 and Williams82 (Figure 5B). These results demonstrate that the candidate gene of *qPro20* and *qOil20* is likely the same gene *Glyma.20G085100* as previously reported.

To further investigate the effect of *Glyma.20G085100* on N and P translocation efficiencies, the RIL population was grouped into two groups according to the TE genotype. The flank markers of the gene were 20_31540803 and 20_31966484 in the present RIL population. The fragment from 20_30257371 to 20_31540803 was then considered as the upstream fragment of the gene, while fragment from 20_31966484 to 20_33401768 was the downstream fragment of the gene. When the upstream and downstream fragments of tested lines both came from JD17, the line was defined as +TE genotype, while those lacking both JD17 fragments were defined as −TE genotypes. This led to the identification of 77 +TE lines and 78 −TE lines in the RIL population (Figure 5C).

NTE, PTE, NcD, and PcD served as indexes of N and P translocation efficiencies. The higher the NTE and PTE, the more N and P were allocated to seeds. The higher NcD and PcD were, the more N and P were translocated from leaves to seeds. According to the genotype categories described above, 77 +TE progeny and 78 −TE progeny were selected to study the effect of this TE on N and P translocation efficiency (Figure 6). The results showed that −TE lines had significantly higher NTE, PTE, NcD, and PcD values than the +TE lines. In comparison with the translocation abilities of −TE lines with +TE lines, we found that NcD and PcD were greater than NTE and PTE, which indicates that *Glyma.20G085100* may be most impactful on NcD and PcD, especially PcD (Figure 6A–D). These results are consistent with the observation that *qPcD20* had the largest LOD and PVE values of all QTLs detected in the study. Meanwhile, −TE lines had higher protein contents than +TE

lines (Figure 6E). Taken together, the results above imply that higher N and P translocation efficiencies facilitate the accumulation of protein in seeds.

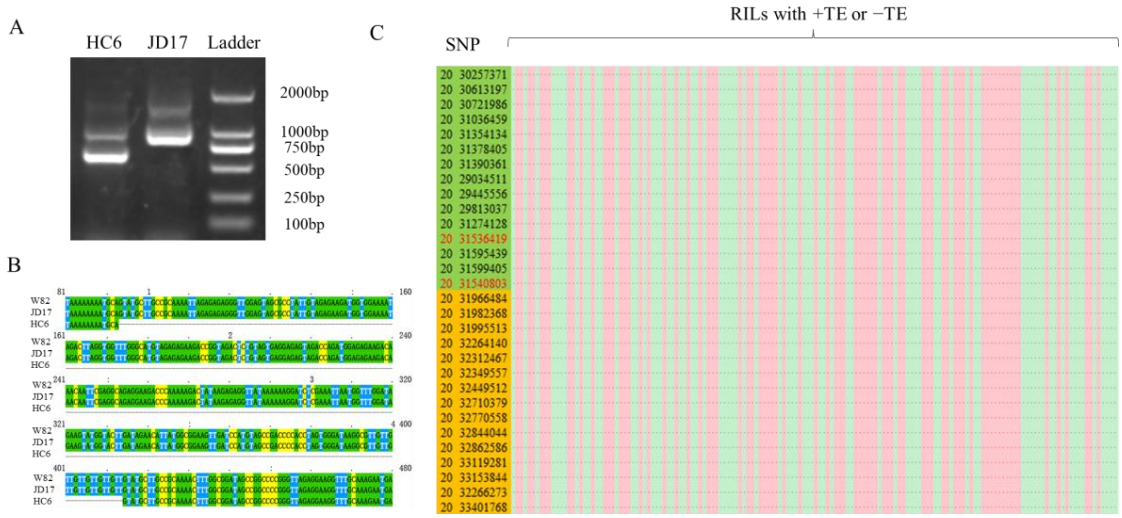

Figure 5. Validation of transposition in causative gene. PCR genotyping assay for TE insertion (**A**) was conducted in the parental lines of the studied RIL populations. PCR amplicons of about 900 bp and 600 bp represent the presence and absence of the TE insertion, respectively. Alignments of *Glyma.20G085100* sequence fragments from Williams82, JD17, and HC6 (**B**) showing the 321-bp InDel (dashed line), with the whole alignment also provided in Figure S3. Genotypes of RILs (**C**) with the upstream (green highlighted) and downstream (orange highlighted) fragments of *Glyma.20G085100* are shown. The SNP highlighted in red had the greatest LOD value for *qPro20* and *qOil20*. Light green filled columns represents +TE lines and light red filled columns represents −TE lines.

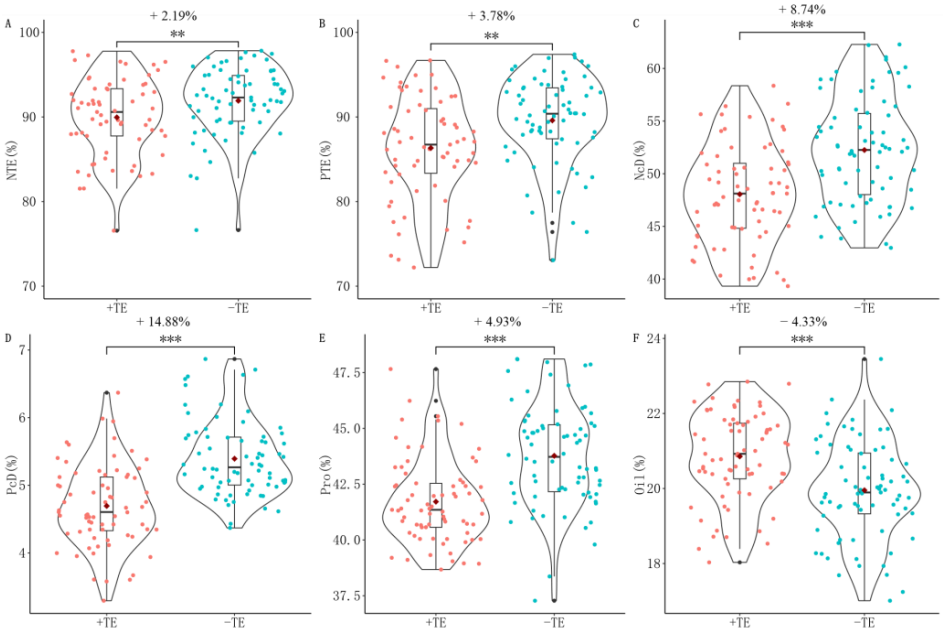

Figure 6. Comparison of nitrogen and phosphorus translocation efficiencies and quality traits of RIL lines carrying +TE or −TE. Of 155 graphed genotypes, 77 were +TE and 78 were −TE. Asterisks denote significance of differences (threshold $p = 0.05$) according to Student's t tests; ** $p < 0.01$ and *** $p < 0.001$. Trait values are shown with violin and box plots. The percentage values above asterisks denote increase or decrease in −TE lines compared with +TE lines. The square red dot represents the mean value. Black round dots denote outliers. (**A, B**) Villon plot of N and P translocation efficiency (N/PTE); (**C, D**) Villon plot of N and P concentration difference (N/PcD); (**E, F**) Villon plot of protein and oil content.

4. Discussion

Soybean is a main source of protein for humans due to its high seed protein content. To improve seed quality, both seed protein and oil contents have been studied in numerous populations, with the result that a few interesting candidate genes have been identified [13,48–50].

Acquisition and utilization of nutrients, especially N and P, is critical for soybean quality [14]. A number of studies on NUE have been reported in soybean [51,52] and rice [53–55], though potential genetic connections between efficiency of nutrient translocation and seed quality have received little attention in soybean. In this study, two parents contrasting in N and P translocation efficiencies, as well as seed quality traits, were used to construct a population of recombinant inbred lines (RILs) comprising 179 individuals (Figure 1). Then, a high-density genetic map was constructed with 5250 markers (Figure 3). QTLs related to N and P translocation efficiencies and quality were detected and analyzed for possible genetic relationships.

Through genetic analysis, it was found that the observed nutrient and quality traits were quantitative (Table 1), which is consistent with previous reports [24,56], suggesting they are controlled by multiple genes. The heritabilities of NTE and PTE observed herein were about 0.7, while those of seed protein and oil were about 0.90 and 0.88, respectively, all of which is consistent with previous reports [4,57–59]. The heritability of PcS in this study was higher than the heritability of PcS reported in a previous study, which may result from the testing of different populations or experimenting in different environments [24].

Based on a constructed high-resolution genetic map, the identified quality loci and some N translocation efficiency loci are consistent with those reported in previous studies. For example, *qOil5* overlaps with *GmST05* [60], while *qOil9* lies near *qOC9(21)* [4], *Seed oil 42–26* [61] and *Seed oil 43–22* [4,59,61]. N and P translocation efficiency markers colocalize with *qPro6* and *qOil6* in a heterochromatic region of chromosome 6 near *qQ6.1* and *qQ6.2* [4]. Plus, N and P translocation efficiency markers colocalizing with qPro*20* and *qOil20* also overlap with the gene *Glyma.20G085100* [13,47]. In addition, the *qNcS18* locus was found near Satt 570, which is known to control seed N concentration at the R7 stage [56].

The production of seeds with high protein contents requires high levels of P and N [62–64]. This seems to be affected by N and P translocation efficiencies, as evident in the present results through the observed positive correlations of seed quality with N and P translocation efficiencies (Figure 1 and Figure S1), as well as the colocalization of numerous relevant loci in chromosomes 6 and 20 (Figure 4). The region on chromosome 6 holding multiple significant QTLs was mined for candidate genes related to N or P translocation efficiency. The interval from 06_18864382 to 06_19585884 contains *Glyma.06G204300*, encoding a TCP-type transcription factor with roles in plant development at the cellular, organ, and tissue levels [65], along with *Glyma.06g196400*, with functions in phospholipid transport according to the gene ontology annotation in SoyBase (http://soybase.org, accessed on 5 December 2022). In addition, the region from 06_17275501 to 06_17280964 is known to harbor *Glyma.06G194500*, which, according to the annotation on SoyBase (http://soybase.org, accessed on 5 December 2022), is a GDSL-like Lipase/Acyl hydrolase superfamily protein, a family with functions that are highly correlated with seed quality and size in other plants [66–69]. Finally, the *qNTE06* and *qPTE06* loci were found near *Glyma.06G200200*, a homologue to *SWEET17* in Arabidopsis, which codes a facilitative transporter that mediates fructose transport across the tonoplast of Arabidopsis root and leaf cells [70].

Another region of interest was identified on chromosome 20, which appears to contain genes regulating soybean seed quality. This region is known to hold a CCT gene, *Glyma.20G085100*, that underlies a large-effect protein/oil QTL, along with a TE insertion in *Glyma.20G085100* that is associated with seed oil and protein content, all of which have been suggested to describe a gene involved in nutrient transport, though direct genetic evidence remains lacking [13]. In our work, re-sequencing showed that HC6 had a 321 bp deletion compared with JD17 and Williams82 (Figure 5), and confirmed that the candidate

gene for *qPro20* and *qOil20* was *Glyma.20G085100*, with −TE lines displaying significantly higher NTE, PTE, NcD, PcD and seed protein than +TE lines (Figure 6). Taken together, these colocalization results suggest that high nutrient translocation efficiency facilitates protein accumulation. In addition, the results herein also indicate that *Glyma.20G085100* may have a bigger effect on NcD and PcD, especially PcD, than on NTE and PTE (Figure 6), which suggests a preference for maintaining high concentration differences over nutrient allocation, especially for P. Most directly, this study provides evidence that a TE increases N and P translocation efficiencies and leads to increasing protein content in soybean seeds.

In short, the present study identified and localized QTLs for soybean N and P translocation efficiencies, along with quality traits in a high-resolution genetic map. Loci for N and P translocation efficiencies and seed protein content were mapped to colocalized regions on chromosome 6 and chromosome 20. The gene *Glyma.20G085100* could be the likely candidate gene for *qPro20.* Our study also provides direct proof that a TE in *Glyma.20G085100* is behind increasing N and P translocation efficiencies that have led to increasing protein contents in soybean seeds. On the whole, this study provides a theoretical basis for breeding soybean lines with high seed protein concentrations and high N and P translocation efficiencies, with a TE InDel identified as a likely genetic marker.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agronomy13020598/s1, Figure S1: Correlation analysis of nutrient and quality traits in seeds. Figure S2: Plots of genetic distance vs physical distance for soybean chromosomes. Figure S3: BLAST of resequenced JD17 and HC6 sequences with William82. Table S1: Genetic correlation analysis of nutrient efficiency traits and quality traits in seeds. Table S2: ANOVA analysis for nutrient efficiency traits and quality traits in seeds.

Author Contributions: Conceptualization, Q.Z. and H.L.; Data curation, N.M., R.L. and Y.Z.; Funding acquisition, H.L.; Investigation, N.M. and R.L.; Methodology, Q.Z., Y.Z., X.L. and H.L.; Resources, H.L.; Validation, N.M.; Visualization, Q.Z.; Writing—original draft, Q.Z. and N.M.; Writing—Review and editing, Y.Z., X.L. and H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was support by the National Natural Science Foundation of China (31830083) and the National Key Research and Development Program of China (2021YFF1000500).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to thank Thomas Walk of Golden Fidelity LLC for English assistance.

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

References

1. Li, X.H.; Shao, Z.Q.; Tian, R.; Zhang, H.; Du, H.; Kong, Y.B.; Li, W.L.; Zhang, C.Y. Mining QTLs and candidate genes for seed protein and oil contents across multiple environments and backgrounds in soybean. *Mol. Breed.* **2019**, *39*, 139. [CrossRef]
2. Wilson, R.F. Soybean: Market Driven Research Needs. In *Genetics and Genomics of Soybean*; Springer: New York, NY, USA, 2008.
3. Ray, D.K.; Mueller, N.D.; West, P.C.; Foley, J.A. Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS ONE* **2013**, *8*, e66428. [CrossRef]
4. Zhao, Q.; Qin, J.; Li, X.; Liu, B.; Liu, Y.; Yang, Q.; Liu, S.; Zhao, X.; Ma, N.; Yan, L.; et al. Coordinate inheritance of seed isoflavone and protein in soybean. *Agriculture* **2022**, *12*, 1178. [CrossRef]
5. Association, A.S. SoyStats. 2022. Available online: http://soystats.com/ (accessed on 5 December 2022).
6. David, G.; Nelson, R.T.; Cannon, S.B.; Shoemaker, R.C. SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res.* **2010**, *38* (Suppl. 1), D843–D846.
7. 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]

8. 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.* **2021**, *296*, 91–102. [CrossRef] [PubMed]
9. Hwang, E.-Y.; Song, Q.; Jia, G.; Specht, J.E.; Hyten, D.L.; Costa, J.; Cregan, P.B. A genome-wide association study of seed protein and oil content in soybean. *BMC Genom.* **2014**, *15*, 1. [CrossRef] [PubMed]
10. Sonah, H.; O'Donoughue, L.; Cober, E.; Rajcan, I.; Belzile, F. Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. *Plant Biotechnol. J.* **2015**, *13*, 211–221. [CrossRef]
11. Lee, S.; Van, K.; Sung, M.; Nelson, R.; LaMantia, J.; McHale, L.K.; Mian, M.A.R. 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]
12. 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]
13. Goettel, W.; Zhang, H.; Li, Y.; Qiao, Z.; Jiang, H.; Hou, D.; Song, Q.; Pantalone, V.R.; Song, B.-H.; Yu, D.; et al. POWR1 is a domestication gene pleiotropically regulating seed quality and yield in soybean. *Nat. Commun.* **2022**, *13*, 3051. [CrossRef]
14. Marschner, H. *Mineral Nutrition of Higher Plants*; Academic Press: San Diego, CA, USA, 2012.
15. Elser, J.J.; Bracken, M.E.; Cleland, E.E.; Gruner, D.S.; Harpole, W.S.; Hillebrand, H.; Ngai, J.T.; Seabloom, E.W.; Shurin, J.B.; Smith, J.E. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecol. Lett.* **2007**, *10*, 1135–1142. [CrossRef] [PubMed]
16. Maathuis, F.J. Physiological functions of mineral macronutrients. *Curr. Opin. Plant Biol.* **2009**, *12*, 250–258. [CrossRef]
17. Hirel, B.; Le Gouis, J.; Ney, B.; Gallais, A. The challenge of improving nitrogen use efficiency in crop plants: Towards a more central role for genetic variability and quantitative genetics within integrated approaches. *J. Exp. Bot.* **2007**, *58*, 2369–2387. [CrossRef] [PubMed]
18. Gaude, N.; Nakamura, Y.; Scheible, W.R.; Ohta, H.; Dormann, P. Phospholipase C5 (NPC5) is involved in galactolipid accumulation during phosphate limitation in leaves of Arabidopsis. *Plant J.* **2008**, *56*, 28–39. [CrossRef]
19. Stigter, K.A.; Plaxton, W.C. Molecular Mechanisms of Phosphorus Metabolism and Transport during Leaf Senescence. *Plants* **2015**, *4*, 773–798. [CrossRef] [PubMed]
20. Mitsuhashi, N.; Ohnishi, M.; Sekiguchi, Y.; Kwon, Y.U.; Chang, Y.T.; Chung, S.K.; Inoue, Y.; Reid, R.J.; Yagisawa, H.; Mimura, T. Phytic acid synthesis and vacuolar accumulation in suspension-cultured cells of Catharanthus roseus induced by high concentration of inorganic phosphate and cations. *Plant Physiol.* **2005**, *138*, 1607–1614. [CrossRef]
21. Rochester, I.J. Using seed nitrogen concentration to estimate crop N use-efficiency in high-yielding irrigated cotton. *Field Crops Res.* **2012**, *127*, 140–145. [CrossRef]
22. Rose, T.J.; Pariasca-Tanaka, J.; Rose, M.T.; Fukuta, Y.; Wissuwa, M. Genotypic variation in grain phosphorus concentration, and opportunities to improve P-use efficiency in rice. *Field Crops Res.* **2010**, *119*, 154–160. [CrossRef]
23. Egli, D.B.; Bruening, W.P. Accumulation of Nitrogen and Dry Matter by Soybean Seeds with Genetic Differences in Protein Concentration. *Crop Sci.* **2007**, *47*, 359–366. [CrossRef]
24. King, K.E.; Lauter, N.; Lin, S.F.; Scott, M.P.; Shoemaker, R.C. Evaluation and QTL mapping of phosphorus concentration in soybean seed. *Euphytica* **2012**, *189*, 261–269. [CrossRef]
25. Zhao, X.; Zheng, S.H.; Suzuki, A.; Arima, S. Varietal Difference in Nitrogen Redistribution from Leaves and Its Contribution to Seed Yield in Soybean. *Plant Prod. Sci.* **2014**, *17*, 103–108. [CrossRef]
26. Chen, L.; Liao, H. Engineering crop nutrient efficiency for sustainable agriculture. *J. Integr. Plant Biol.* **2017**, *59*, 710–735. [CrossRef] [PubMed]
27. Yang, Y.; Tong, Y.; Li, X.; He, Y.; Xu, R.; Liu, D.; Yang, Q.; Lv, H.; Liao, H. Genetic analysis and fine mapping of phosphorus efficiency locus 1 (PE1) in soybean. *Theor. Appl. Genet.* **2019**, *132*, 2847–2858. [CrossRef] [PubMed]
28. Kichey, T.; Hirel, B.; Heumez, E.; Dubois, F.; Le Gouis, J. In winter wheat (*Triticum aestivum* L.), post-anthesis nitrogen uptake and remobilisation to the grain correlates with agronomic traits and nitrogen physiological markers. *Field Crops Res.* **2007**, *102*, 22–32. [CrossRef]
29. He, M.; Yan, Z.; Cui, X.; Gong, Y.; Li, K.; Han, W. Scaling the leaf nutrient resorption efficiency: Nitrogen vs. phosphorus in global plants. *Sci. Total Environ.* **2020**, *729*, 138920. [CrossRef] [PubMed]
30. Feller, S.H.U. Nitrogen metabolism and remobilization during senescence. *J. Exp. Bot.* **2002**, *53*, 927–937.
31. Masclaux-Daubresse, C.; Daniel-Vedele, F.; Dechorgnat, J.; Chardon, F.; Gaufichon, L.; Suzuki, A. Nitrogen uptake, assimilation and remobilization in plants: Challenges for sustainable and productive agriculture. *Ann. Bot.* **2010**, *105*, 1141–1157. [CrossRef]
32. Hammond, J.P.; Bennett, M.J.; Bowen, H.C.; Broadley, M.R.; Eastwood, D.C.; May, S.T.; Rahn, C.; Swarup, R.; Woolaway, K.E.; White, P.J. Changes in gene expression in Arabidopsis shoots during phosphate starvation and the potential for developing smart plants. *Plant Physiol.* **2003**, *132*, 578–596. [CrossRef]
33. Veneklaas, E.J.; Lambers, H.; Bragg, J.; Finnegan, P.M.; Lovelock, C.E.; Plaxton, W.C.; Price, C.A.; Scheible, W.R.; Shane, M.W.; White, P.J.; et al. Opportunities for improving phosphorus-use efficiency in crop plants. *New Phytol.* **2012**, *195*, 306–320. [CrossRef]
34. Zhang, B.; Chen, P.; Shi, A.; Hou, A.; Ishibashi, T.; Wang, D. Putative quantitative trait loci associated with calcium content in soybean seed. *J. Hered.* **2009**, *100*, 263–269. [CrossRef]
35. Ramamurthy, R.K.; Jedlicka, J.; Graef, G.L.; Waters, B.M. Identification of new QTLs for seed mineral, cysteine, and methionine concentrations in soybean [*Glycine max* (L.) Merr.]. *Mol. Breed.* **2014**, *34*, 431–445. [CrossRef]

36. Watson, A.; Ghosh, S.; Williams, M.J.; Cuddy, W.S.; Simmonds, J.; Rey, M.D.; Md Hatta, M.A.; 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]
37. Cui, B.; Chen, L.; Yang, Y.; Liao, H.; Bohra, A. Genetic analysis and map-based delimitation of a major locus qSS3 for seed size in soybean. *Plant Breed.* **2020**, *139*, 1145–1157. [CrossRef]
38. Bao, S. *Soil and Agricultural Chemistry Analysis*; Chinese Agriculture Publishing House: Beijing, China, 2000.
39. Yan, L.; Jiang, C.; Yu, X.; Yang, C.; Zhang, M. Development and reliability of near infrared spectroscopy (NIS) models of protein and oil content in soybean. *Soybean Sci.* **2008**, *27*, 5.
40. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **2011**, *6*, e19379. [CrossRef] [PubMed]
41. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef] [PubMed]
42. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **2010**, *38*, e164. [CrossRef]
43. Voorrips, R. MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.* **2002**, *93*, 77–78. [CrossRef] [PubMed]
44. VanOoijen, J.W. *MapQTL 6, Software for the Mapping of Quantitative Trait Loci in experimental Populations of Diploid Species*; Kyazma B.V.: Wageningen, The Netherlands, 2004.
45. Ihaka, R.; Gentleman, R. R: A language for data analysis and graphics. *J. Comput. Graph. Stat.* **1996**, *5*, 299–314.
46. Takahashi, R.; Dubouzet, J.G.; Matsumura, H.; Asuda, K.Y.; Iwashina, T. A new allele of flower color gene W1 encoding flavonoid 3′5′-hydroxylase is responsible for light purple flowers in wild soybean Glycine soja. *BMC Plant Biol.* **2010**, *10*, 155. [CrossRef]
47. Fliege, C.E.; Ward, R.A.; Vogel, P.; Nguyen, H.; Quach, T.; Guo, M.; Viana, J.P.G.; Santos, L.B.; Specht, J.E.; Clemente, T.E.; et al. Fine mapping and cloning of the major seed protein quantitative trait loci on soybean chromosome 20. *Plant J.* **2022**, *110*, 114–128. [CrossRef]
48. Miao, L.; Yang, S.N.; Zhang, K.; He, J.B.; Wu, C.H.; Ren, Y.H.; Gai, J.Y.; Li, Y. Natural variation and selection in *GmSWEET39* affect soybean seed oil content. *New Phytol.* **2020**, *225*, 1651–1666. [CrossRef] [PubMed]
49. Zhang, T.F.; Wu, T.T.; Wang, L.W.; Jiang, B.J.; Zhen, C.X.; Yuan, S.; Hou, W.S.; Wu, C.X.; Han, T.F.; Sun, S. A combined linkage and GWAS analysis identifies QTLs linked to soybean seed protein and oil content. *Int. J. Mol. Sci.* **2019**, *20*, 5915. [CrossRef] [PubMed]
50. Qin, J.; Wang, F.; Zhao, Q.; Shi, A.; Zhao, T.; Song, Q.; Ravelombola, W.; An, H.; Yan, L.; Yang, C.; et al. Identification of candidate genes and genomic selection for seed protein in soybean breeding pipeline. *Front. Plant Sci.* **2022**, *13*, 882732. [CrossRef] [PubMed]
51. Chen, Z.; Li, L.; Halford, N.G.; Xu, H.; Huang, L.; Gao, R.; Lu, R.; Liu, C. Advances in barley breeding for improving nitrogen use efficiency. *Agronomy* **2022**, *12*, 1682. [CrossRef]
52. Rosolem, C.A.; Batista, T.B.; Dias, P.P.; Neto, L.V.D.M.; Calonego, J.C. The joint application of phosphorus and ammonium enhances soybean root growth and P uptake. *Agriculture* **2022**, *12*, 880. [CrossRef]
53. Yamaji, N.; Takemoto, Y.; Miyaji, T.; Mitani-Ueno, N.; Yoshida, K.T.; Ma, J.F. Reducing phosphorus accumulation in rice grains with an impaired transporter in the node. *Nature* **2017**, *541*, 92–95. [CrossRef]
54. Zhang, Z.; Gao, S.; Chu, C. Improvement of nutrient use efficiency in rice: Current toolbox and future perspectives. *Theor. Appl. Genet.* **2020**, *133*, 1365–1384. [CrossRef]
55. Yang, X.; Nian, J.; Xie, Q.; Feng, J.; Zhang, F.; Jing, H.; Zhang, J.; Dong, G.; Liang, Y.; Peng, J. Rice ferredoxin-dependent glutamate synthase regulates nitrogen–carbon metabolomes and is genetically differentiated between japonica and indica subspecies. *Mol. Plant* **2016**, *9*, 1520–1534. [CrossRef]
56. Panthee, D.; Pantalone, V.; Sams, C.; Saxton, A.; West, D.; Rayford, W. Genomic regions governing soybean seed nitrogen accumulation. *J. Am. Oil Chem. Soc.* **2004**, *81*, 77–81. [CrossRef]
57. Gu, R.; Chen, F.; Long, L.; Cai, H.; Liu, Z.; Yang, J.; Wang, L.; Li, H.; Li, J.; Liu, W.; et al. Enhancing phosphorus uptake efficiency through QTL-based selection for root system architecture in maize. *J. Genet. Genom.* **2016**, *43*, 663–672. [CrossRef] [PubMed]
58. Shen, C.; Chen, K.; Cui, Y.; Chen, J.; Mi, X.; Zhu, S.; Zhu, Y.; Ali, J.; Ye, G.; Li, Z.; et al. QTL mapping and favorable allele mining of nitrogen deficiency tolerance using an interconnected breeding population in rice. *Front. Genet.* **2021**, *12*, 616428. [CrossRef]
59. Mao, T.T.; Jiang, Z.F.; Han, Y.P.; Teng, W.L.; Zhao, X.; Li, W.B. Identification of quantitative trait loci underlying seed protein and oil contents of soybean across multi-genetic backgrounds and environments. *Plant Breed.* **2013**, *132*, 630–641. [CrossRef]
60. Duan, Z.; Zhang, M.; Zhang, Z.; Liang, S.; Fan, L.; Yang, X.; Yuan, Y.; Pan, Y.; Zhou, G.; Liu, S.; et al. Natural allelic variation of GmST05 controlling seed size and quality in soybean. *Plant Biotechnol. J.* **2022**, *20*, 1807–1818. [CrossRef]
61. Han, Y.P.; Teng, W.L.; Wang, Y.; Zhao, X.; Wu, L.; Li, D.M.; Li, W.B. Unconditional and conditional QTL underlying the genetic interrelationships between soybean seed isoflavone, and protein or oil contents. *Plant Breed.* **2015**, *134*, 300–309. [CrossRef]
62. Fabre, F.; Planchon, C. Nitrogen nutrition, yield and protein content in soybean. *Plant Sci.* **2000**, *152*, 51–58. [CrossRef]
63. Yang, Y. Fine Mapping and Candidate Gene Identification of a Soybean Seed Protein and Oil QTL from a Wild Soybean Accession and Linkage Analysis for Whole Plant Biomass, Carbon, Nitrogen, and Seed Composition Using a RIL Mapping Population. Ph.D. Thesis, University of Missouri Columbia, Columbia, MO, USA, 2021.

64. Yin, X.; Bellaloui, N.; Mcclure, A.M.; Tyler, D.D.; Mengistu, A. Phosphorus Fertilization Differentially Influences Fatty Acids, Protein, and Oil in Soybean. *Am. J. Plant Sci.* **2016**, *7*, 1975–1992. [CrossRef]
65. Manassero, N.G.; Viola, I.L.; Welchen, E.; Gonzalez, D.H. TCP transcription factors: Architectures of plant form. *Biomol. Concepts* **2013**, *4*, 111–127. [CrossRef]
66. Ursache, R.; De Jesus Vieira Teixeira, C.; Tendon, V.D.; Gully, K.; De Bellis, D.; Schmid-Siegert, E.; Andersen, T.G.; Shekhar, V.; Calderon, S.; Pradervand, S.; et al. GDSL-domain proteins have key roles in suberin polymerization and degradation. *Nat. Plants* **2021**, *7*, 353–364. [CrossRef]
67. Clauss, K.; von Roepenack-Lahaye, E.; Bottcher, C.; Roth, M.R.; Welti, R.; Erban, A.; Kopka, J.; Scheel, D.; Milkowski, C.; Strack, D. Overexpression of sinapine esterase BnSCE3 in oilseed rape seeds triggers global changes in seed metabolism. *Plant Physiol.* **2011**, *155*, 1127–1145. [CrossRef] [PubMed]
68. Ma, R.; Yuan, H.; An, J.; Hao, X.; Li, H. A Gossypium hirsutum GDSL lipase/hydrolase gene (GhGLIP) appears to be involved in promoting seed growth in Arabidopsis. *PLoS ONE* **2018**, *13*, e0195556. [CrossRef] [PubMed]
69. Kondou, Y.; Nakazawa, M.; Kawashima, M.; Ichikawa, T.; Yoshizumi, T.; Suzuki, K.; Ishikawa, A.; Koshi, T.; Matsui, R.; Muto, S.; et al. RETARDED GROWTH OF EMBRYO1, a new basic helix-loop-helix protein, expresses in endosperm to control embryo growth. *Plant Physiol.* **2008**, *147*, 1924–1935. [CrossRef] [PubMed]
70. Guo, W.J.; Nagy, R.; Chen, H.Y.; Pfrunder, S.; Yu, Y.C.; Santelia, D.; Frommer, W.B.; Martinoia, E. SWEET17, a facilitative transporter, mediates fructose transport across the tonoplast of Arabidopsis roots and leaves. *Plant Physiol.* **2014**, *164*, 777–789. [CrossRef] [PubMed]