

Article

# On the Inverse Correlation of Protein and Oil: Examining the Effects of Altered Central Carbon Metabolism on Seed Composition Using Soybean Fast Neutron Mutants

Shrikaar Kambhampati <sup>1</sup>, Jose A. Aznar-Moreno <sup>2,†</sup>, Cooper Hostetler <sup>1,‡</sup>, Tara Caso <sup>1,‡</sup>, Sally R. Bailey <sup>3,‡</sup>, Allen H. Hubbard <sup>1</sup>, Timothy P. Durrett <sup>2,\*</sup> and Doug K. Allen <sup>1,3,\*</sup>

<sup>1</sup> Donald Danforth Plant Science Center, St. Louis, MO 63132, USA; skambhampati@danforthcenter.org (S.K.); cooperhostetler@gmail.com (C.H.); tsc152@humboldt.edu (T.C.); ahubbard@danforthcenter.org (A.H.H.)

<sup>2</sup> Department of Biochemistry and Molecular Biophysics, Kansas State University, Manhattan, KS 66506, USA; ja.aznar@upm.es

<sup>3</sup> United States Department of Agriculture, Agricultural Research Service, St. Louis, MO 63132, USA; sally.bailey@usda.gov

\* Correspondence: tdurrett@ksu.edu (T.P.D.); doug.allen@usda.gov (D.K.A.)

† Current address: Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid, Pozuelo de Alarcón, 28223 Madrid, Spain.

‡ These authors contributed equally to this work.

Received: 6 December 2019; Accepted: 21 December 2019; Published: 28 December 2019



**Abstract:** Protein and oil levels measured at maturity are inversely correlated across soybean lines; however, carbon is in limited supply during maturation resulting in tradeoffs for the production of other reserves including oligosaccharides. During the late stages of seed development, the allocation of carbon for storage reserves changes. Lipid and protein levels decline while concentrations of indigestible raffinose family oligosaccharides (RFOs) increase, leading to a decreased crop value. Since the maternal source of carbon is diminished during seed maturation stages of development, carbon supplied to RFO synthesis likely comes from an internal, turned-over source and may contribute to the reduction in protein and lipid content in mature seeds. In this study, fast neutron (FN) mutagenized soybean populations with deletions in central carbon metabolic genes were examined for trends in oil, protein, sugar, and RFO accumulation leading to an altered final composition. Two lines with concurrent increases in oil and protein, by combined 10%, were identified. A delayed switch in carbon allocation towards RFO biosynthesis resulted in extended lipid accumulation and without compromising protein. Strategies for future soybean improvement using FN resources are described.

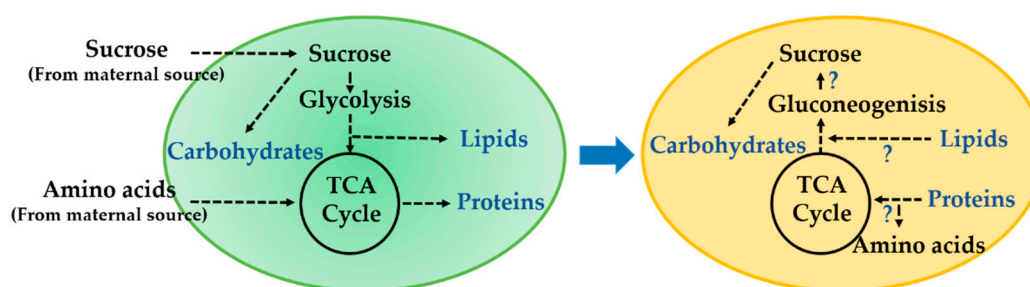
**Keywords:** soybeans; fast neutron mutants; protein; oil; raffinose family oligosaccharides; correlations; linear models; central carbon metabolism; seed development

## 1. Introduction

Soybean (*Glycine max*) is a versatile and important crop with many uses. In 2017, 80 million acres of soybeans were planted in the United States (United States Department of Agriculture (USDA), National Agricultural Statistics Service (NASS)). The value of soybean is mostly due to high protein content with balanced amino acid composition that is used to produce feed for livestock and high seed oil content that is a primary source of vegetable oil and the feed stock for biofuel production. Soybean protein represents approximately 35–40% of the total biomass in soybeans and is the component of greatest value in soybean meal. Depending on the cultivar and planting zone [1–3], an additional

3–12% of biomass is present as sugars and oligosaccharides [4]. Sugars, such as sucrose, represent a metabolizable energy source in animal feed, but raffinose and stachyose that are members of the raffinose family oligosaccharides (RFOs) are indigestible and undesirable for livestock production [5,6]. RFOs have hypothesized roles in seed desiccation tolerance [7–9], stability of liposomes during dehydration, and seed germination [10–12], although efforts in breeding and genetic engineering have demonstrated that reduced levels do not impact seed viability [13–15]. The RFO and sucrose levels are the result of central carbon metabolism and could adversely impact the levels of oil and protein, which are known to be inversely correlated in mature seeds [16–21].

Protein and lipid accumulation over the course of seed development is dependent on the supply of amino acids and sugars from the maternal sources (i.e., organic carbon assimilated in the leaf) as well as the metabolism within the developing seed (i.e., cotyledons) [22–25]. The levels of intermediates of central carbon (C) and nitrogen (N) metabolism that are precursors for storage reserve production, vary depending on the stage of reproductive development and may indicate temporal changes in metabolism [26–28]. During maturation, 10–15% of lipids are degraded [28–30], coinciding with RFO accumulation, and occurring at the time when there are little to no exogenous resources supplied by the maternal plant. The expression of genes involved in gluconeogenesis and glyoxylate cycle suggest carbon remobilization from lipid may occur. Along with existing sucrose present late in development, carbon from lipid could contribute to central metabolism and enable RFO production during the maturation phase. Protein may also contribute carbon through turnover induced by proteases expressed late in development [28,31]. Figure 1 summarizes the hypothetical shift in metabolism from seed fill to maturity.



**Figure 1.** Drawing showing hypothetical shift in carbon partitioning from seed filling to maturation phase. During seed filling (R5–R7), central carbon (C) and nitrogen (N) sources (i.e., sucrose and amino acids) from maternal tissue are distributed to lipids, carbohydrates, and proteins through glycolysis, tricarboxylic acid (TCA) cycle, and amino acid metabolic pathways. As the seed transitions to maturation phase (R7 to R8), some accumulated proteins are degraded by proteolysis, while lipids are degraded via  $\beta$ -oxidation and potentially used for raffinose family oligosaccharide (RFO) biosynthesis through gluconeogenesis.

The relationship between the accumulation of protein and oil has been examined. Through *in vivo* embryo culture, changes in the supply of C and N sources resulted in inverse correlations between the two biomass components [22,32]. Based on the changes to maternal supply and seed metabolism that occur over the course of seed development, we hypothesized that in late seed development, carbon derived from the turnover of lipids and proteins contributes to the synthesis of RFOs. This concept was tested by inspecting correlations between biomass components in developing soybean seeds. Fast neutron (FN) lines, which had differences in genes involved in C and N partitioning, were evaluated.

Fast neutron (FN) mutagenesis-based forward genetic screens have been successful at identifying the genetic basis of valuable phenotypes in soybean [33–36]. Advances in the use of whole genome sequencing and comparative genome hybridization (CGH) techniques for mapping the deletions in mutant populations [37–39] also provide a means to identify novel mutations in genomic studies [40,41]

that can have applications in plant breeding programs for improved agronomic traits. Studying the role of altered C and N partitioning as a function of development may provide clues to overcome the inverse correlation between oil and protein and to develop soybean lines with improved seed composition.

## 2. Materials and Methods

### 2.1. Plant Material and Growth Conditions

Soybean (*Glycine max*) seeds selected for gene mutations in central carbon metabolism were obtained from two fast neutron (FN) mutagenized populations developed at the University of Missouri (MO) and the University of Minnesota (MN) with the USDA (Table 1). Seeds were grown in greenhouses at Kansas State University (KSU) in 2018 for initial lipid and protein evaluation with fatty acid methyl ester (FAME) (described in Section 2.3) and NIR analysis, respectively. One-gallon pots were filled with 4:2:1 mixture of Metromix Professional Growing Mix/Vermiculite/Perlite supplemented with Osmocote 14-14-14 slow release fertilizer. Plants were watered three times per week and grown with day/night temperatures ranging from 25 °C to 27 °C/21 °C to 23 °C and humidity maintained at 45% or greater. Supplemental light (400–1000 W m<sup>-2</sup>) was used to maintain a 14 h day/10 h night photoperiod. Select FN lines were replanted in 2019, at the Donald Danforth Plant Science Center. Two seeds from each mutant line were planted in individual one-gallon pots containing Fafard 4M soil and grown in greenhouses. Temperature, humidity, and light conditions were maintained similar to the conditions at KSU. Plants were watered daily and fertilized with Jack's 15-16-17 fertilizer (JR Peters, Allentown, PA, USA) three times per week. Pods were harvested from each plant at R5, R6, R7, R7.5, and R8 stages of reproductive development [42,43] and placed directly on ice. Seeds were removed from the pods, seed coats were excised, and the cotyledons were flash frozen in liquid nitrogen. Samples were stored at –80 °C until further use.

**Table 1.** Fast neutron (FN) mutagenized population used for preliminary oil and protein analysis in 2018.

FN Descriptor	Descriptors Used Here	50-Seed Weight (g)	Genes of Interest within FN Deletion or Phenotype Used as Selection Criteria
Williams 82	WT-Wm 82	10.05	Wild type
FN300660	FN300660	11.87	$\alpha/\beta$ hydrolase, Glyma14g09181
FN300088	FN300088	10.49	$\alpha/\beta$ hydrolase, Glyma17g35990
FN300068.01.01.M4	FN300068	13.62	F1,6BPase, Glyma16g28310
FN300012.02.BC1F3	FN300012	11.15	UDP-Gluc-6DH, Glyma01g06970
FN300026.02.01.M4	FN300026	11.57	PEPCK, Glyma04g09510
FN300108	FN300108	7.66	ACX2, Glyma06g43840
FN301952.01.01.M4	FN301952	13.33	Enolase, Glyma03g34830
FN301653.01.01.M4	FN301653	12.11	F1,6BPase, Glyma18g41940
M92-220	WT-M92-220	13.65	Wild type
FN0171815	FN0171815	14.63	high oil
FN0110332.03.M3	FN0110332	11.98	Aldolase, Glyma12g24190
FN0170714.03.01.02.M5	FN0170714	16.3	F1,6BPase Glyma18g41940
P1082Daar541BMN15 <sup>1</sup>	P1082D	10.76	high protein
FN0175143	FN0175143	13.8	high oil
FN0171855.08.01.01.10.M6	FN0171855	13.32	PEPCK Glyma04g09510
5R16CO1Ddar117MN15 <sup>1</sup>	5R16CO1D	12.92	high oil
FN0170904.01.M3	FN0170904	13.93	F1,6BPase Glyma16g28310
FN0171734	FN0171734	9.93	SUS1/4 Glyma09g08550
PO559Caar541MN15 <sup>1</sup>	PO559C	12.75	high protein + oil
FN0141075.03.04.10.M5	FN0141075	12.29	ACX1 Glyma01g41600
1R37C45cbdaar254MN15 <sup>1</sup>	1R37C45	13.28	high protein + oil
FN0171466	FN0171466	14.23	high oil
FN0173708	FN0173708	13.32	high oil
FN0175116	FN0175116	13.08	high oil
FN0111996	FN0111996	13.53	PXA1 Glyma11g38160
FN0173054.M2	FN0173054	14.18	PGM Glyma08g04890

<sup>1</sup> Lines belonged to the 2015 core collection with no FN ID.

## 2.2. Extraction of Lipids, Protein, and Soluble Sugars

Following the method described in Ma et al. [44] with limited modifications, lipids, protein, and soluble sugars were obtained from the cotyledons (analyzed as single seeds). Stored samples were removed from  $-80\text{ }^{\circ}\text{C}$  and placed in liquid nitrogen, two metal beads were then added to each tube, and the samples were pulverized using a ball mill at 30 Hz for 25 s for three rounds or until completely powdered. Pulverized samples were lyophilized for  $\sim 48\text{ h}$  at 2.4 Pa and  $-48\text{ }^{\circ}\text{C}$ . After lyophilization, the metal beads were removed, and the dry weight of the sample was obtained. Samples were collected in triplicate from two individual plants to obtain a total of six replicates used for lipid and soluble sugar extraction and another six for protein hydrolysis.

To extract and quantify the lipids and soluble sugars, 1 mL 7:3 methanol/chloroform ( $-20\text{ }^{\circ}\text{C}$ ), a triheptadecanoin standard, a tripentadecanoin standard, and a norvaline/ribitol mixed standard were added to each sample. Samples were vortexed and mixed at  $4\text{ }^{\circ}\text{C}$  for  $\sim 2\text{ h}$  with all subsequent steps occurring at  $4\text{ }^{\circ}\text{C}$ . A 500  $\mu\text{L}$  portion of ddH<sub>2</sub>O was added to each sample prior to vortexing, then centrifugation at 14,000 rpm for 10 min, to generate three distinct phases. The upper aqueous phase, containing water-soluble metabolites (sugars and free amino acids), was transferred to a 1.5 mL tube. The samples containing soluble sugars were dried in a speed vacuum centrifuge (Labconco®, Kansas City, MO, USA) and stored at  $-80\text{ }^{\circ}\text{C}$  for LC/MS analysis. A 200  $\mu\text{L}$  portion of 100% methanol was added to the remaining  $\sim 300\text{ }\mu\text{L}$  organic phase (containing lipids) and protein layers. The samples were vortexed and centrifuged at 14,000 rpm for 10 min, resulting in a pellet that contained insoluble proteins and carbohydrates along with an upper layer of chloroform/methanol. This upper layer was then transferred to an 8 mL glass vial and dried in a speed vacuum centrifuge to be processed into fatty acid methyl esters (FAMES).

Protein quantification was performed by hydrolyzing the peptide bonds and quantifying the resulting individual amino acids as described in Kambhampati et al. [45]. Protein hydrolysis was performed in 100  $\mu\text{L}$  of 4 M methanesulfonic acid with 0.2% tryptamine spiked with 20  $\mu\text{L}$  of 1 mM  $^{13}\text{C}$   $^{15}\text{N}$ -labeled amino acid standard mix. Samples were incubated at  $110\text{ }^{\circ}\text{C}$  for  $\sim 22\text{ h}$  followed by neutralization with 100  $\mu\text{L}$  of 4 M NaOH. Samples were then completely dried down and 1 mL 50% MeOH was added to resuspend hydrolyzed amino acids prior to filtering with a 0.8  $\mu\text{m}$  PES membrane centrifuge filter. Chromatography and mass spectrometry conditions for separation and detection of amino acids were identical to the method described in Kambhampati et al. [45].

## 2.3. FAME Analysis for the Quantification of Lipids

Analysis of lipid content was carried out according to an adapted version of the method described in Allen and Young [22]. In brief, freshly prepared 5% sulfuric acid:methanol (*v/v*) was added to the dried product along with 25  $\mu\text{L}$  0.2% butylated hydroxytoluene (BHT) in methanol to prevent oxidation before being vortexed and heated at  $110\text{ }^{\circ}\text{C}$  for 2.5–3 h, vortexing hourly. After cooling to room temperature, 0.9% NaCl (*w/v*) was added to each sample to quench the reaction. The FAMES were then extracted using hexane and quantified by gas chromatography–flame ionization detection (GC-FID) using a DB23 column (30 m, 0.25 mm i.d., 0.25  $\mu\text{m}$  film; J&W Scientific, Folsom, CA, USA). The GC was operated in a split mode (50:1). The flame ionization detector was operated with a temperature of  $250\text{ }^{\circ}\text{C}$  with an oven program that ramped from  $180\text{ }^{\circ}\text{C}$  to  $260\text{ }^{\circ}\text{C}$  at a rate of  $20\text{ }^{\circ}\text{C min}^{-1}$  followed by a hold time of 7 min. Comparisons of peak areas to the two internal standards were used for quantification.

## 2.4. Quantification of Sugars using HILIC-MS/MS

The aqueous phase samples containing water-soluble metabolites were separated for sugars and analyzed using a Shimadzu HPLC system (UFLCXR, Columbia, MD, USA) connected to an AB Sciex 6500 triple quadrupole/trap MS instrument equipped with Turbo V™ electrospray ionization (ESI) source. Negative ion mode was used for ionization of sugars. A 3  $\mu\text{L}$  sample was injected on the Infinity

Lab Poroshell 120 Z-HILIC column (2.7  $\mu\text{m}$ , 100  $\times$  2.1 mm; Agilent Technologies, Santa Clara, CA, USA) and the metabolites were eluted with an increasing gradient of acetonitrile/10 mM ammonium formate (90:10 *v/v* ACN:H<sub>2</sub>O) (A) and 10 mM ammonium formate in water (B) with pH adjusted to 6.9 for both buffers. Both the buffers were spiked with 5  $\mu\text{M}$  medronic acid to improve compound retention on the column. The flow rate used was 0.20 mL/min. Sugars were separated using a binary gradient of 100–70% B over 5 min and then to 30% B over the next 7 min followed by a hold at 30% B for 1 min. The gradient was then returned to 100% B over the next minute before re-equilibrating the column for six minutes. Eluted compounds separated by the HPLC were introduced to the mass spectrometer by an electrospray ionization (ESI) source with the following conditions: ion spray voltage, 4.5 kV (ESI-); ion source temperature, 550 °C; source gas 1, 45 psi; source gas 2, 40 psi; curtain gas, 35 psi; and entrance potential, 10. Ions were detected and monitored using a targeted multiple reaction monitoring (MRM) approach with the parameters listed in Table S1 (Supplementary Materials). The value for entrance potential was default (–10) for all analytes. Other parameters for MRMs were optimized using direct injections of individual sugar standards. Data were analyzed using the quantitation wizard available in Analyst (v. 1.6.2) software (AB SCIEX, Concord, ON, Canada). Metabolite concentrations and recoveries were calculated based on a calibration curve and ribitol as an internal standard, respectively.

### 2.5. Linear Modeling and Statistical Analyses

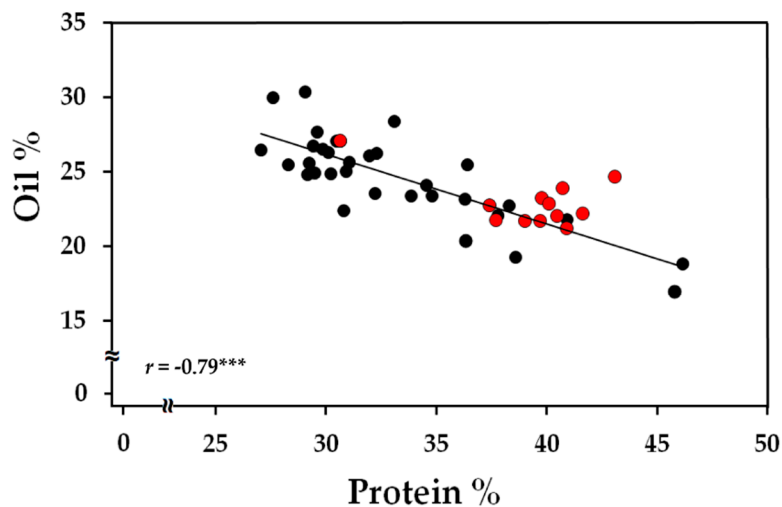
Correlations between bulk proteins, lipids, and oil measurements were calculated at the developmental time points R5, R6, R7, R7.5, and R8. Multiple measurements taken from a single plant made it possible to observe the impact of developmental time on correlations between pairs of compounds and to determine if the levels at one time point are predictive of levels at the other time points. Linear modeling, implemented in R programming language, was then used to determine if the relationship between a pair of compounds changed as a function of time during at least one developmental stage (*p*-value of interaction term <0.05). For each pair of metabolites (A,B), models of the form  $A \sim \text{time point} + B \times \text{time point}$  and  $B \sim \text{time point} + A \times \text{time point}$  were calculated.

## 3. Results

### 3.1. Identification of FN Lines with Alterations in Central Carbon Metabolic Genes

A total of 25 lines from two core collections of soybean fast neutron mutant population were selected. The first collection consisted of 8 lines from a mutant population developed at the University of Missouri, Columbia from the “Williams 82” background and the second consisted of 17 lines from a mutant population developed at the University of Minnesota and USDA using the background “M92-220” [37]. A list of the FN descriptors for the lines, 50-seed weight at maturity, and the selection criteria for central carbon metabolic gene deletions or oil phenotype are presented in Table 1. In order to assess the oil content and stability of loci leading to a high oil phenotype within these populations, total oil content at maturity was measured from four individual seeds selected from four different plants for each of the selected lines. Three lines from the Williams 82 mutagenized population and eight from the M92-220 mutagenized population showed a high oil phenotype (Figure S1, Supplementary Materials).

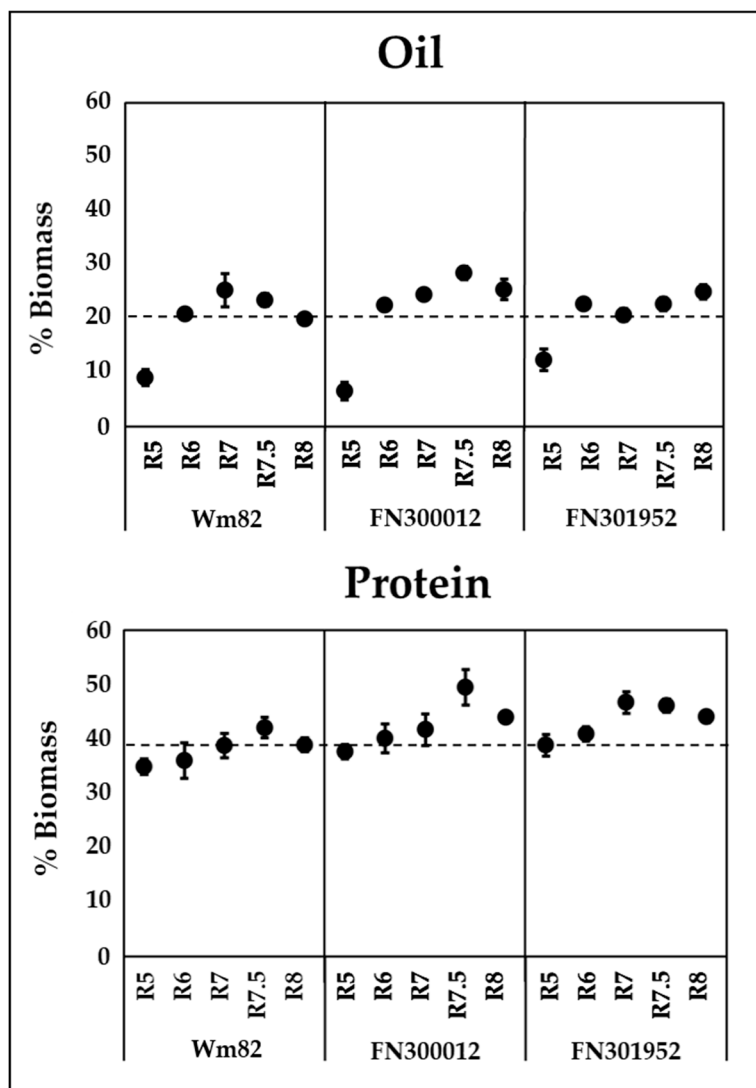
The protein content in the selected eleven lines along with the two wild types, Williams 82 and M92-220 (Table S2, Supplementary Materials), was measured to assess the relationship between oil and protein. An overall negative correlation was observed between the two biomass components with a Pearson’s correlation coefficient,  $r = -0.79$  and  $p < 0.0001$  (Figure 2). However, in three FN lines under Williams 82 background (FN300660, FN300012, and FN301952), the high oil content did not result in a significant penalty in the protein level. The seeds obtained from these lines were replanted in 2019 for biomass component description over the course of seed development in the subsequent generation.



**Figure 2.** Correlation plot for oil and protein content represented as % biomass at maturity. A total of 11 FN lines with four biological replicates were used in the analysis. Red markers represent the three lines, FN300660, FN300012, and FN301952, from the Williams 82 mutagenized population.  $r$  represents the Pearson's correlation coefficient and \*\*\* indicates a significance of  $p < 0.001$ .

### 3.2. FN300012 and FN301952 Accumulated High Oil without a Decline in Their Total Protein Content

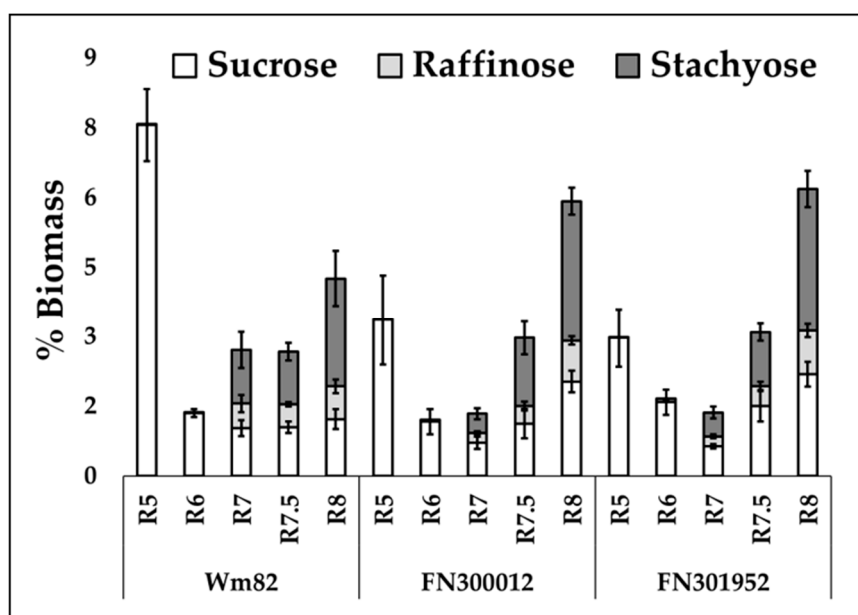
Total oil and protein contents in cotyledons were quantified in all three lines and the wild-type Williams 82 over the course of development from R5 to R8. Reproductive developmental stages were defined based on the seed weights (Table S3, Supplementary Materials) and were consistent with previous descriptions [42,43]. FN300660 displayed large variation in protein and oil accumulation based on single-seed analysis, likely due to segregation of causal genes for the metabolic phenotype and hence was not pursued further. The accumulation of oil over the course of development has been previously reported for Williams 82 [28]. A similar trend was observed in our study, where the content increased up to R7 and declined during the maturation phase (R7 and R8). This decline was delayed in line FN300012 and no decrease in oil content was observed in FN301952 (Figure 3). Protein accumulation in Williams 82 continued to R7.5 and declined towards the end. No difference in the trends of accumulation or decline were observed for protein between the two FN lines and Williams 82; however, the quantity of protein accumulated reached as high as  $49.5 \pm 3\%$  in FN300012 at R7.5 and  $46.2 \pm 1\%$  in FN301952 at R7, whereas the highest protein content in Williams 82 reached only  $42.1 \pm 2\%$  at R7.5. The overall protein content at maturity was also elevated in both the mutant lines ( $44 \pm 1\%$ ) compared to the wild type ( $38.9 \pm 1\%$ ). The increased protein content in mutants was expected to be accompanied by reduced oil levels, based on the inverse correlation between protein and oil [16–21], yet an overall increase in oil content at maturity was also observed for both the mutant lines ( $25.5 \pm 2\%$ ,  $25 \pm 1.3\%$  for FN300012 and FN301952, respectively) relative to the wild type ( $20 \pm 0.7\%$ ).



**Figure 3.** Oil (top panel) and protein (bottom panel) content in cotyledons over the course of development in two FN lines FN300012 and FN301952 compared to the wild-type Williams 82 (Wm82). Dotted lines represent the oil and protein quantities of the wild type at maturity for comparison. Error bars are the standard error of the mean (SEM),  $n = 6$ . Absolute values are presented in Table S3.

### 3.3. No Decrease in RFOs Was Observed in the Final Composition of Selected Lines

RFOs constitute 3–12% of biomass at maturity [1–3] and are a significant carbon sink during the maturation phase. Since the stages corresponding to the decline in oil coincide with the accumulation of RFOs in Williams 82 (Figure 4), we hypothesized that there was a change in resource allocation to accumulate raffinose and stachyose (the most abundant RFOs) along with sucrose in the mutant lines. The total raffinose and stachyose content was found to be ~2.5-fold lower in both the mutant lines (FN300012, FN301952) compared to Williams 82 at R7 (Figure 4), indicating a delayed RFO synthesis when lipid accumulation continued into the maturation phase (Section 2.2, Figure 3). However, the RFO content at maturity did not show a significant difference between the two mutants and the wild-type Williams 82. The sucrose levels at maturity in the two mutants, FN300012 and FN301952, were elevated by 1.7-fold ( $p = 0.021$ ) and 1.8-fold ( $p = 0.017$ ), respectively.

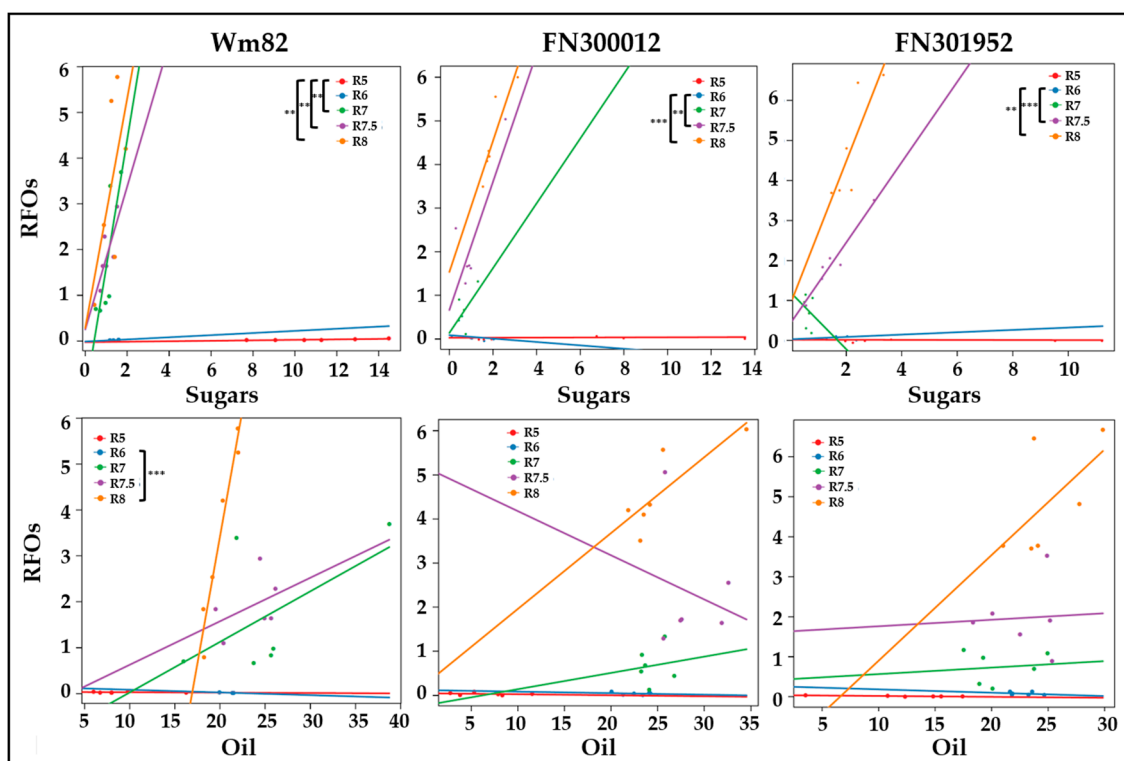


**Figure 4.** Sucrose, raffinose, and stachyose content in the two lines, FN300012 and FN301952, compared to Williams 82 over the course of development. Error bars represent standard error of the mean ( $n = 6$ ). Absolute values are presented in Table S3.

#### 3.4. A Delayed Switch towards RFO Accumulation Leads to Extended Oil and Protein Accumulation

The change in resource allocation between sugars, oil, and RFOs was investigated using a correlation-based approach followed by linear modeling [46]. Correlation matrices between biomass components including sugars (glucose + fructose + sucrose), RFOs (raffinose + stachyose), oil, and protein, at all stages were generated for Williams 82 and the two related FN lines, FN300012 and FN301952 (Figure S2, Supplementary Materials). A positive correlation between sugars and RFOs was observed in the wild type in all stages of development except maturity, indicating the channeling of free sugars towards RFOs from the beginning of storage reserve deposition. This correlation, however, was found to be negative at R6 in FN300012, and no significant positive correlations were observed until R7.5 in both FN300012 and FN301952. The results suggest that the allocation of carbon from sugars to RFOs is delayed and may contribute to enhanced lipid accumulation. Linear modeling (where the difference in slopes of the regression lines are compared over time) used to identify the relationship between sugars and RFOs revealed a significant difference between R5 and maturation phases (R7 and R8) (Figure 5). This difference in slope, indicated by a significant interaction term ( $p$ -value  $< 0.05$ ), may represent a metabolic shift. Linear models have been used previously to identify compounds that are differentially directed towards metabolic fates under regulation [46]. In the types of metabolic shifts detected in soybean, the biochemical relationship between classes of compounds changes as a function of developmental time. The difference was not apparent until R7.5 in the mutant lines FN300012 and FN301952, consistent with extended oil accumulation and delayed RFO production.





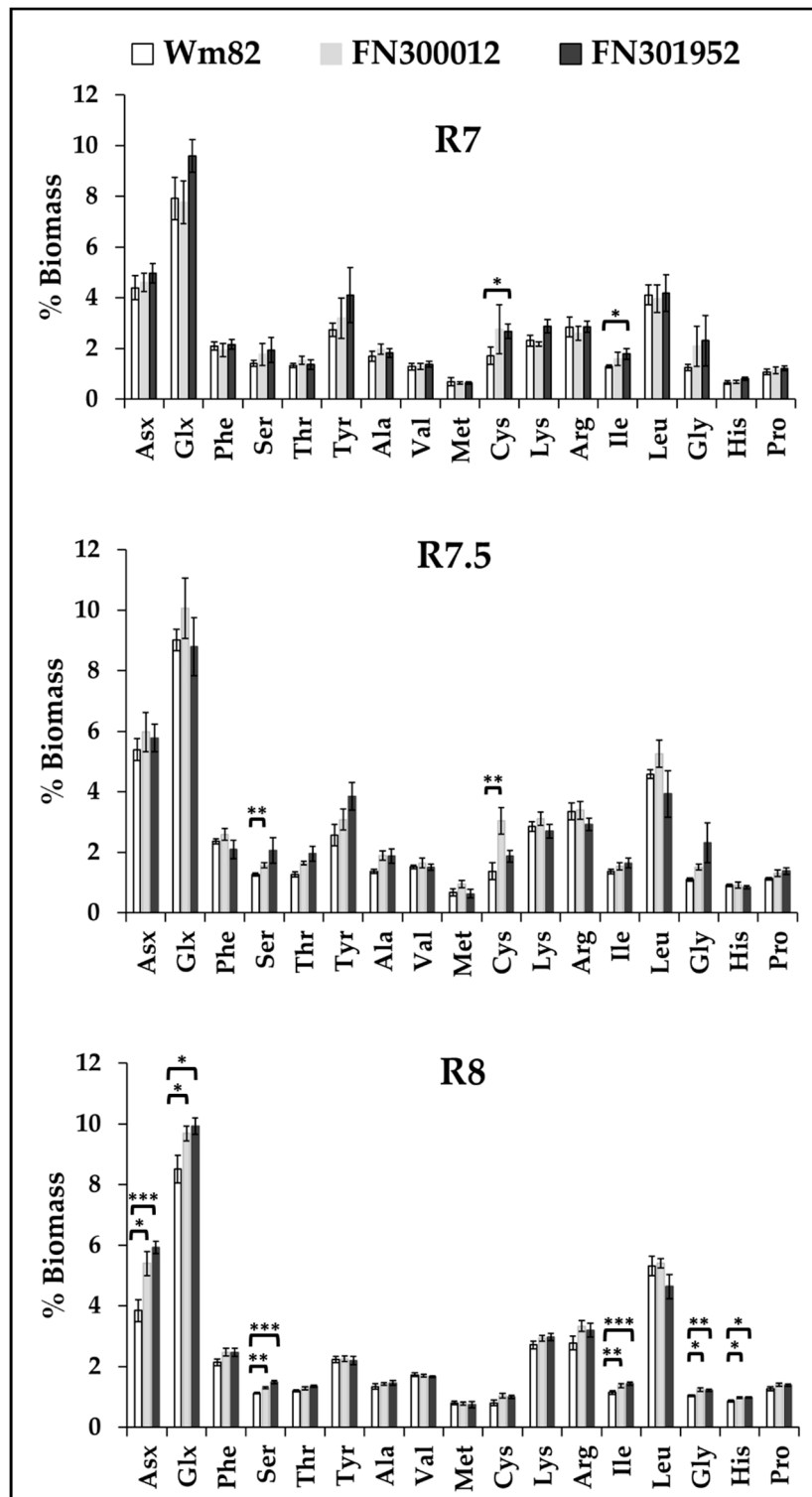
**Figure 5.** Linear models between sugars and RFOs (top panel) or oil content and RFOs (bottom panel) represented by regression lines, color coded for developmental time points (R5–R8). Statistically significant differences in slopes between time points that represent a change in the direction of carbon flow are marked for each of the three lines, namely Williams 82, FN300012, and FN301952. \*\*  $p < 0.01$  and \*\*\*  $p < 0.001$ . Units for both axes represent % biomass. Sugars represent the sum of glucose, fructose, and sucrose. RFOs represent raffinose and stachyose together.

Out of the possible paired relationships examined using the linear modeling approach (sugars, RFOs, oil, and protein) over time (R5–R8), other significant differences in slopes were identified related to oil and RFO accumulation, indicating that the carbon from the lipid may be repurposed for RFO synthesis in Williams 82 (Figure 5). Although not statistically significant, a positive relationship between oil and RFOs was observed in wild-type Williams 82 at R7, which was not observed until R8 in both the mutant lines as indicated by the slope of regression lines.

### 3.5. Cysteine Content That Contributes to Protein Quality Peaked Earlier in Development

The protein quality of the meal from soybeans, in addition to total quantity, contributes to the market value. The mutant lines, FN300012 and FN301952, had protein levels greater (Section 3.2) than the wild-type (Williams 82) and were further evaluated for protein quality through quantification of the amino acids. The sulfur-containing essential amino acids, methionine and cysteine, are generally lower in soybeans than other sources of protein; thus, improving the production of these amino acids has received considerable attention [47,48]. Although the levels of these two amino acids at maturity are not significantly different between the two FN lines compared to Williams 82, the content of cysteine is elevated in R7 in FN300012 and R7.5 in FN301952 before declining to wild-type levels (Figure 6), suggesting that the genes involved in either biosynthesis or degradation of these amino acids are affected. The level of serine, that is the precursor of cysteine biosynthesis, is also significantly higher in FN300012 at R7.5. This may indicate that precursors for cysteine biosynthesis are more available and could contribute to enhanced protein quality. The total protein composition at maturity is different in the two mutants compared to the wild type. Asx (asparagine + aspartate), Glx (glutamine + glutamate),

serine, isoleucine, glycine, and histidine are all greater in the mutants compared to the wild type (Figure 6), suggesting an altered protein profile possibly resulting from changes to protease activities as a result of the FN mutations.



**Figure 6.** Amino acid composition of the total protein between the three lines, Williams 82, FN300012, and FN301952, at R7 (beginning of maturation), R7.5 (mid-maturation), and R8 (maturity). Values represent a mean ( $n = 6$ ). Student's *t*-test was used for statistical analyses between the wild type and the two mutants. \*  $p < 0.05$ , \*\*  $p < 0.01$ , and \*\*\*  $p < 0.001$ .

#### 4. Discussion

Interconnected enzymatic steps result in the flow of carbon through central metabolic pathways to the production of valuable biomass components, including oil, protein, and carbohydrates [23,49]. Conventional breeding and transgenic approaches for improving the quality or quantity of biomass reserves often results in rebalancing between components within the seed [18,48]. Protein and lipid are co-produced during development, including during the seed maturation phase when resources from the maternal plant are scarce. During the maturation phase, the production of one storage component likely results from repartitioning of carbon amongst storage reserves. Attempts at rebalancing protein and oil have given mixed results, with strategies to enhance protein such as heterologous protein expression [50] and suppression of specific seed storage proteins [51] providing small gains in the overall protein content due to concurrent changes within the proteome whereas the expression of the *Arabidopsis thaliana* orphan gene *QUA-QUINE STARCH (QQS)* in soybean seeds resulted in 4–10% increase in protein predominantly at the expense of starch [52,53]. Conversely, natural variants of *RAFFINOSE SYNTHASE* responsible for RFO accumulation have decreased RFO content and increased sucrose without further gains in oil or protein accumulation [5]. Thus altering a strategic combination of genes to reshuffle carbon and nitrogen between biomass components can improve the quality of soybeans; but the outcomes remain difficult to predict a priori. We hypothesize that since resources (such as carbon) are finite, the reallocation of carbon to lipid and/or protein instead of RFOs could be important to consider for enhanced seed value at maturity.

FN mutagenized populations have a unique advantage in generating random genomic deletions at multiple loci creating multiple phenotypes [37], and provide a means to select lines that have improved seed composition traits. Genome mapping techniques such as array CGH for locating gene deletions and duplications have expanded the use of FN mutagenized populations as a resource for reverse genetic studies. A number of examples now indicate enhanced protein levels through gene combinations in lines with high oil ([54] and references within) and the use of FN mutagenized populations to identify lines with similar potential [40,41]. In this study, a set of FN lines with high oil, protein, and/or known genetic alterations was selected based on deletions to central carbon metabolic genes responsible for partitioning the maternal nutrient sources into biomass (Table 1). An initial analysis from a subset of these lines with high oil revealed a strong negative correlation (Figure 2) for most lines as has been previously observed [16–21]. The two FN lines, FN300012 and FN301952, however, yielded a combined protein and oil quantity roughly 10% (69.5% and 69.1%, respectively) higher compared to the wild-type level (58.9%) at maturity, under greenhouse conditions, avoiding the inverse correlation between oil and protein.

The lines FN300012 and FN301952 were investigated over the course of seed development to assess changes in biomass that result from maternal contribution and temporal changes in metabolism within the seed [22–24,55]. The delayed switch in carbon allocation towards raffinose and stachyose occurred concomitantly with prolonged lipid accumulation in FN300012 and FN301952. Furthermore, the decline of lipids generally associated with the maturation period in oilseeds [28–30] was delayed in FN300012 and not apparent in FN301952, leading to an increased final oil content. The source of excess carbon deposited into lipids, however, remains unclear since significant decreases in final RFO levels were not observed. The sucrose and protein contents at maturity were also higher in the two mutants, thus the dietary fiber or other unquantified biomass pools could offset the increases in lipids. Indeed, FN300012 was selected due to a knockout in chromosome 1 (~0.7 Mb) that contained a gene coding for a putative UDP-glucose 6-dehydrogenase activity (Glyma01g06970, E.C:1.1.1.22) identified using CGH analysis and catalogued in SoyBase [56]. This enzyme has been shown to enable cell wall polysaccharide biosynthesis as the seed matures [57]. In addition, FN301952 had a ~2.3 Mb homozygous deletion in chromosome 3 that contained three glycolytic enzymes, namely an aldolase (Glyma03g34950), a putative enolase (Glyma03g34830), and a pyruvate kinase (Glyma03g34740), all of which were predicted to be localized in cytosol using WoLF PSORT [58]. The loss of cytosolic components of glycolysis may impact the available hexose phosphates in the chloroplast and result in

increased de novo lipid synthesis, which would be distinct from recent strategies to augment lipid levels by reducing lipases in oilseeds, including *Arabidopsis* [59], rapeseed [60], *Jatropha* [61], and soybean [62].

Proteolysis of storage proteins in developing seeds has previously been described [63], and transcripts of several proteases and protein degradation systems are known to be expressed throughout seed development [28,31]. The two FN lines contained several proteases among the list of deleted genes, likely contributing to reduced or delayed proteolysis leading to loss of protein during the maturation phase observed in Williams 82. The mutant lines contained higher protein-linked cysteine content important to meal quality [48], during mid-maturation phase. A gene related to the methionine degradation pathway, S-adenosyl methionine synthetase (Glyma03g34120), that was deleted in FN301952 may affect methionine levels affecting degradation and produce cysteine that accumulates in protein. Several protease-related genes were deleted in the mutant lines and, along with others that impact sulfur-enriched proteins, remain a target for higher quality protein at maturity.

## 5. Conclusions

Fast neutron mutagenesis resulted in lines with different carbon partitioning to storage reserves and served as a valuable resource for identification and characterization of novel phenotypes and their genetic basis. We identified two FN lines with increased oil and protein by a combined 10%. Temporal changes in biomass composition revealed a delayed carbon allocation to RFO synthesis in the mutant lines compared to the wild type. Further studies on the temporal changes in metabolism during soybean development can be used to identify candidate genes for elite soybean germplasm production.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2218-1989/10/1/18/s1>, Figure S1: Total lipid content represented as percent biomass in the selected FN population derived from the wild type, M92-220 (top panel), and the ones derived from Williams 82 (bottom panel). Figure S2: Correlation matrices showing temporal relationships between the four biomass components, namely, sugars, RFOs, oil, and protein in the two FN lines, FN300012 and FN301952, along with the wild-type Williams 82. Grey shaded values represent the biologically relevant comparisons. \*  $p < 0.05$ , \*\*  $p < 0.01$ , and \*\*\*  $p < 0.001$ . Table S1: Multiple reaction monitoring (MRM) method settings for the quantification of sugars using SCIEX Triple Quad™ 6500+ system. Table S2: Oil and protein content of the eleven fast neutron mutant lines along with their wild-type backgrounds used for correlation, as described in Figure 2. Table S3: Fresh weights of seeds used for analyses at different stages of seed development along with oil, protein, sucrose, and RFO content that represent the data presented in Figure 3; Figure 4. The R scripts for correlation analysis and the linear modeling as well as the data used are available at [https://github.com/victorfrnak/allen\\_lab](https://github.com/victorfrnak/allen_lab).

**Author Contributions:** Conceptualization, project administration, and funding acquisition, D.K.A. and T.P.D.; methodology, software, and validation, S.K., J.A.A.-M., C.H., T.C., and S.R.B.; formal analysis, investigation, and data curation, S.K., A.H.H., and J.A.A.-M.; writing—original draft preparation, S.K.; writing—review and editing, S.K., J.A.A.-M., D.K.A., and T.P.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the United States Department of Agriculture, Agricultural Research Service, the United States Department of Agriculture, National Institute of Food and Agriculture through award number USDA-NIFA #2017-67013-26156, and the Multistate Research Project NC1203, the United Soybean Board (#1820-152-0134), and the National Science Foundation (NSF-DBI #1659812). Support for the acquisition of the 6500 QTRAP LC-MS/MS was also provided by the National Science Foundation (NSF-DBI #1427621).

**Acknowledgments:** The authors would like to acknowledge the Plant Growth Facility (PGF) and Proteomics and Mass Spectrometry Facility (PMSF) at the Donald Danforth Plant Science Center for accommodating large planting volumes in their growth spaces and instrumentation used for analyses, respectively, as well as Robert Stupar at the University of Minnesota, and Minviluz Stacey and Gary Stacey at the University of Missouri, Columbia for their generous donation of FN seeds.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Leprince, O.; Pellizzaro, A.; Berriri, S.; Buitink, J. Late seed maturation: Drying without dying. *J. Exp. Bot.* **2016**, *68*, 827–841. [[CrossRef](#)] [[PubMed](#)]

2. Verdier, J.; Lalanne, D.; Pelletier, S.; Torres-Jerez, I.; Righetti, K.; Bandyopadhyay, K.; Leprince, O.; Chatelain, E.; Vu, B.L.; Gouzy, J.; et al. A regulatory network-based approach dissects late maturation processes related to the acquisition of desiccation tolerance and longevity of medicago truncatula seeds. *Plant Physiol.* **2013**, *163*, 757–774. [[CrossRef](#)] [[PubMed](#)]
3. Kumar, V.; Rani, A.; Goyal, L.; Dixit, A.K.; Manjaya, J.G.; Dev, J.; Swamy, M. Sucrose and raffinose family oligosaccharides (RFOs) in soybean seeds as influenced by genotype and growing location. *J. Agric. Food Chem.* **2010**, *58*, 5081–5085. [[CrossRef](#)] [[PubMed](#)]
4. Openshaw, S.J.; Hadley, H.H. Maternal effects on sugar content in soybean seeds. *Crop Sci.* **1978**, *18*, 581–584. [[CrossRef](#)]
5. Dierking, E.C.; Bilyeu, K.D. Association of a soybean raffinose synthase gene with low raffinose and stachyose seed phenotype. *Plant Genome* **2008**, *1*, 135–145. [[CrossRef](#)]
6. Coon, C.N.; Leske, K.L.; Akavanichan, O.; Cheng, T.K. Effect of oligosaccharide-free soybean meal on true metabolizable energy and fiber digestion in adult roosters<sup>1</sup>. *Poult. Sci.* **1990**, *69*, 787–793. [[CrossRef](#)]
7. Koster, K.L.; Leopold, A.C. Sugars and desiccation tolerance in seeds. *Plant Physiol.* **1988**, *88*, 829–832. [[CrossRef](#)]
8. Horbowicz, M.; Obendorf, R.L. Seed desiccation tolerance and storability: Dependence on flatulence-producing oligosaccharides and cyclitols—Review and survey. *Seed Sci. Res.* **2008**, *4*, 385–405. [[CrossRef](#)]
9. Peterbauer, T.; Richter, A. Biochemistry and physiology of raffinose family oligosaccharides and galactosyl cyclitols in seeds. *Seed Sci. Res.* **2001**, *11*, 185–197.
10. Blöchl, A.; Peterbauer, T.; Richter, A. Inhibition of raffinose oligosaccharide breakdown delays germination of pea seeds. *J. Plant Physiol.* **2007**, *164*, 1093–1096. [[CrossRef](#)]
11. Gangl, R.; Tenhaken, R. Raffinose family oligosaccharides act as galactose stores in seeds and are required for rapid germination of arabidopsis in the dark. *Front. Plant Sci.* **2016**, *7*. [[CrossRef](#)] [[PubMed](#)]
12. Hincha, D.K.; Zuther, E.; Heyer, A.G. The preservation of liposomes by raffinose family oligosaccharides during drying is mediated by effects on fusion and lipid phase transitions. *Biochim. Biophys. Acta-Biomembr.* **2003**, *1612*, 172–177. [[CrossRef](#)]
13. Neus, J.D.; Fehr, W.R.; Schnebly, S.R. Agronomic and seed characteristics of soybean with reduced raffinose and stachyose. *Crop Sci.* **2005**, *45*, 589–592. [[CrossRef](#)]
14. Dierking, E.C.; Bilyeu, K.D. Raffinose and stachyose metabolism are not required for efficient soybean seed germination. *J. Plant. Physiol.* **2009**, *166*, 1329–1335. [[CrossRef](#)] [[PubMed](#)]
15. Valentine, M.F.; De Tar, J.R.; Mookkan, M.; Firman, J.D.; Zhang, Z.J. Silencing of soybean raffinose synthase gene reduced raffinose family oligosaccharides and increased true metabolizable energy of poultry feed. *Front. Plant Sci.* **2017**, *8*. [[CrossRef](#)] [[PubMed](#)]
16. Clemente, T.E.; Cahoon, E.B. Soybean Oil: Genetic Approaches for modification of functionality and total content. *Plant Physiol.* **2009**, *151*, 1030–1040. [[CrossRef](#)]
17. Chaudhary, J.; Patil, G.B.; Sonah, H.; Deshmukh, R.K.; Vuong, T.D.; Valliyodan, B.; Nguyen, H.T. Expanding omics resources for improvement of soybean seed composition traits. *Front. Plant. Sci.* **2015**, *6*. [[CrossRef](#)]
18. Patil, G.; Mian, R.; Vuong, T.; Pantalone, V.; Song, Q.; Chen, 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](#)]
19. Rincker, K.; Nelson, R.; Specht, J.; Sleper, D.; Cary, T.; Cianzio, S.R.; Casteel, S.; Conley, S.; Chen, P.; Davis, V.; et al. Genetic improvement of U.S. soybean in maturity groups II, III, and IV. *Crop Sci.* **2014**, *54*, 1419–1432. [[CrossRef](#)]
20. Bandillo, N.; Jarquin, D.; Song, Q.; Nelson, R.; Cregan, P.; Specht, J.; Lorenz, A. A population structure and genome-wide association analysis on the USDA soybean germplasm collection. *Plant Genome* **2015**, *8*. [[CrossRef](#)]
21. Brzostowski, L.F.; Diers, B.W. Agronomic evaluation of a high protein allele from PI407788A on chromosome 15 across two soybean backgrounds. *Crop Sci.* **2017**, *57*, 2972–2978. [[CrossRef](#)]
22. Allen, D.K.; Young, J.D. Carbon and nitrogen provisions alter the metabolic flux in developing soybean embryos. *Plant Physiol.* **2013**, *161*, 1458–1475. [[CrossRef](#)] [[PubMed](#)]
23. Allen, D.K.; Ohlrogge, J.B.; Shachar-Hill, Y. The role of light in soybean seed filling metabolism. *Plant J.* **2009**, *58*, 220–234. [[CrossRef](#)] [[PubMed](#)]

24. Truong, Q.; Koch, K.; Yoon, J.M.; Everard, J.D.; Shanks, J.V. Influence of carbon to nitrogen ratios on soybean somatic embryo (cv. Jack) growth and composition. *J. Exp. Bot.* **2013**, *64*, 2985–2995. [[CrossRef](#)] [[PubMed](#)]
25. Rotundo, J.L.; Borrás, L.; Westgate, M.E.; Orf, J.H. Relationship between assimilate supply per seed during seed filling and soybean seed composition. *Field Crop. Res.* **2009**, *112*, 90–96. [[CrossRef](#)]
26. Fait, A.; Angelovici, R.; Less, H.; Ohad, I.; Urbanczk-Wochniak, E.; Fernie, A.R.; Galili, G. Arabidopsis seed development and germination is associated with temporally distinct metabolic switches. *Plant Physiol.* **2006**, *142*, 839–854. [[CrossRef](#)]
27. Galili, G.; Avin-Wittenberg, T.; Angelovici, R.; Fernie, A.R. The role of photosynthesis and amino acid metabolism in the energy status during seed development. *Front. Plant Sci.* **2014**, *5*. [[CrossRef](#)]
28. Collakova, E.; Aghamirzaie, D.; Fang, Y.; Klumas, C.; Tabataba, F.; Kakumanu, A.; Myers, E.; Heath, L.S.; Grene, R. Metabolic and transcriptional reprogramming in developing soybean (*Glycine max*) embryos. *Metab.* **2013**, *3*, 347–372. [[CrossRef](#)]
29. Baud, S.; Boutin, J.-P.; Miquel, M.; Lepiniec, L.; Rochat, C. An integrated overview of seed development in *Arabidopsis thaliana* ecotype WS. *Plant Physiol. Biochem.* **2002**, *40*, 151–160. [[CrossRef](#)]
30. Chia, T.Y.P.; Pike, M.J.; Rawsthorne, S. Storage oil breakdown during embryo development of *Brassica napus* (L.). *J. Exp. Bot.* **2005**, *56*, 1285–1296. [[CrossRef](#)]
31. Angelovici, R.; Galili, G.; Fernie, A.R.; Fait, A. Seed desiccation: A bridge between maturation and germination. *Trends Plant Sci.* **2010**, *15*, 211–218. [[CrossRef](#)] [[PubMed](#)]
32. Pipolo, A.E.; Sinclair, T.R.; Camara, G.M.S. Protein and oil concentration of soybean seed cultured in vitro using nutrient solutions of differing glutamine concentration. *Ann. Appl. Biol.* **2004**, *144*, 223–227. [[CrossRef](#)]
33. Campbell, B.W.; Hofstad, A.N.; Sreekanta, S.; Fu, F.; Kono, T.J.Y.; O'Rourke, J.A.; Vance, C.P.; Muehlbauer, G.J.; Stupar, R.M. Fast neutron-induced structural rearrangements at a soybean NAP1 locus result in gnarled trichomes. *Theor. Appl. Genet.* **2016**, *129*, 1725–1738. [[CrossRef](#)] [[PubMed](#)]
34. Stacey, M.G.; Cahoon, R.E.; Nguyen, H.T.; Cui, Y.; Sato, S.; Nguyen, C.T.; Phoka, N.; Clark, K.M.; Liang, Y.; Forrester, J.; et al. Identification of homogentisate dioxygenase as a target for vitamin E biofortification in oilseeds. *Plant Physiol.* **2016**, *172*, 1506–1518. [[CrossRef](#)]
35. Dobbels, A.A.; Michno, J.-M.; Campbell, B.W.; Viridi, K.S.; Stec, A.O.; Muehlbauer, G.J.; Naeve, S.L.; Stupar, R.M. An induced chromosomal translocation in soybean disrupts a *KASI* ortholog and is associated with a high-sucrose and low-oil seed phenotype. *G3 Genes Genomes Genet.* **2017**, *7*, 1215–1223. [[CrossRef](#)]
36. Men, A.E.; Laniya, T.S.; Searle, I.R.; Iturbe-Ormaetxe, I.; Gresshoff, I.; Jiang, Q.; Carroll, B.J.; Gresshoff, P.M. Fast neutron mutagenesis of soybean (*Glycine soja* L.) produces a supernodulating mutant containing a large deletion in linkage group H. *Genome Lett.* **2002**, *1*, 147–155. [[CrossRef](#)]
37. Bolon, Y.-T.; Haun, W.J.; Xu, W.W.; Grant, D.; Stacey, M.G.; Nelson, R.T.; Gerhardt, D.J.; Jeddelloh, J.A.; Stacey, G.; Muehlbauer, G.J.; et al. Phenotypic and genomic analyses of a fast neutron mutant population resource in soybean. *Plant Physiol.* **2011**, *156*, 240–253. [[CrossRef](#)]
38. Bolon, Y.-T.; Stec, A.O.; Michno, J.-M.; Roessler, J.; Bhaskar, P.B.; Ries, L.; Dobbels, A.A.; Campbell, B.W.; Young, N.P.; Anderson, J.E.; et al. Genome resilience and prevalence of segmental duplications following fast neutron irradiation of soybean. *Genetics* **2014**, *198*, 967–981. [[CrossRef](#)]
39. Hwang, W.J.; Kim, M.Y.; Kang, Y.J.; Shim, S.; Stacey, M.G.; Stacey, G.; Lee, S.-H. Genome-wide analysis of mutations in a dwarf soybean mutant induced by fast neutron bombardment. *Euphytica* **2015**, *203*, 399–408. [[CrossRef](#)]
40. Islam, N.; Stupar, R.M.; Qijian, S.; Luthria, D.L.; Garrett, W.; Stec, A.O.; Roessler, J.; Natarajan, S.S. Genomic changes and biochemical alterations of seed protein and oil content in a subset of fast neutron induced soybean mutants. *BMC Plant Biol.* **2019**, *19*, 420. [[CrossRef](#)]
41. Prenger, E.M.; Ostezan, A.; Mian, M.A.R.; 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](#)]
42. Naeve, S.L. Soybean Growth Stages. Available online: <https://extension.umn.edu/growing-soybean/soybean-growth-stages> (accessed on 31 October 2019).
43. Licht, M. Soybean Growth and Development. Available online: <https://store.extension.iastate.edu/product/6451> (accessed on 31 October 2019).
44. Ma, F.; Jazmin, L.J.; Young, J.D.; Allen, D.K. Isotopically nonstationary metabolic flux analysis (INST-MFA) of photosynthesis and photorespiration in plants. In *Photorespiration: Methods and Protocols*; Fernie, A.R., Bauwe, H., Weber, A.P.M., Eds.; Springer: New York, NY, USA, 2017; pp. 167–194.

45. Kambhampati, S.; Li, J.; Evans, B.S.; Allen, D.K. Accurate and efficient amino acid analysis for protein quantification using hydrophilic interaction chromatography coupled tandem mass spectrometry. *Plant Methods* **2019**, *15*, 46. [[CrossRef](#)] [[PubMed](#)]
46. Hubbard, A.H.; Zhang, X.; Jastrebski, S.; Lamont, S.J.; Singh, A.; Schmidt, C.J. Identifying mechanisms of regulation to model carbon flux during heat stress and generate testable hypotheses. *PLoS ONE* **2018**, *13*, e0205824. [[CrossRef](#)] [[PubMed](#)]
47. Krishnan, H.B. Engineering soybean for enhanced sulfur amino acid content. *Crop Sci.* **2005**, *45*, 454–461. [[CrossRef](#)]
48. Krishnan, H.B.; Jez, J.M. Review: The promise and limits for enhancing sulfur-containing amino acid content of soybean seed. *Plant Sci.* **2018**, *272*, 14–21. [[CrossRef](#)]
49. Li, L.; Hur, M.; Lee, J.-Y.; Zhou, W.; Song, Z.; Ransom, N.; Demirkale, C.Y.; Nettleton, D.; Westgate, M.; Arendsee, Z.; et al. A systems biology approach toward understanding seed composition in soybean. *BMC Genom.* **2015**, *16*, S9. [[CrossRef](#)]
50. Kim, W.-S.; Jez, J.M.; Krishnan, H.B. Effects of proteome rebalancing and sulfur nutrition on the accumulation of methionine rich  $\delta$ -zein in transgenic soybeans. *Front. Plant Sci.* **2014**, *5*. [[CrossRef](#)]
51. Schmidt, M.A.; Barbazuk, W.B.; Sandford, M.; May, G.; Song, Z.; Zhou, W.; Nikolau, B.J.; Herman, E.M. Silencing of soybean seed storage proteins results in a rebalanced protein composition preserving seed protein content without major collateral changes in the metabolome and transcriptome. *Plant Physiol.* **2011**, *156*, 330–345. [[CrossRef](#)]
52. Li, L.; Zheng, W.; Zhu, Y.; Ye, H.; Tang, B.; Arendsee, Z.W.; Jones, D.; Li, R.; Ortiz, D.; Zhao, X.; et al. QQS orphan gene regulates carbon and nitrogen partitioning across species via NF-YC interactions. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 14734. [[CrossRef](#)]
53. Li, L.; Wurtele, E.S. The QQS orphan gene of Arabidopsis modulates carbon and nitrogen allocation in soybean. *Plant Biotechnol. J.* **2015**, *13*, 177–187. [[CrossRef](#)]
54. Wilson, R.F. Seed composition. In *Soybeans: Improvement, Production, and Uses*; Boerma, H.R., Specht, J.E., Eds.; American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America: Madison, WI, USA, 2004; pp. 621–677.
55. Rotundo, J.L. Linking assimilate supply and seed developmental processes that determine soybean seed composition. *Eur. J. Agron.* **2011**, *35*, 184–191. [[CrossRef](#)]
56. Grant, D.; Nelson, R.T.; Cannon, S.B.; Shoemaker, R.C. SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res.* **2009**, *38*, 843–846. [[CrossRef](#)] [[PubMed](#)]
57. Litterer, L.A.; Plaisance, K.L.; Schnurr, J.A.; Storey, K.K.; Jung, H.-J.G.; Gronwald, J.W.; Somers, D.A. Biosynthesis of UDP-glucuronic acid in developing soybean embryos: Possible role of UDP-sugar pyrophosphorylase. *Physiol. Plant.* **2006**, *128*, 200–211. [[CrossRef](#)]
58. Horton, P.; Park, K.-J.; Obayashi, T.; Fujita, N.; Harada, H.; Adams-Collier, C.J.; Nakai, K. WoLF PSORT: Protein localization predictor. *Nucleic Acids Res.* **2007**, *35*, 585–587. [[CrossRef](#)]
59. van Erp, H.; Kelly, A.A.; Menard, G.; Eastmond, P.J. Multigene engineering of triacylglycerol metabolism boosts seed oil content in arabidopsis. *Plant Physiol.* **2014**, *165*, 30–36. [[CrossRef](#)]
60. Kelly, A.A.; Shaw, E.; Powers, S.J.; Kurup, S.; Eastmond, P.J. Suppression of the SUGAR-DEPENDENT1 triacylglycerol lipase family during seed development enhances oil yield in oilseed rape (*Brassica napus* L.). *Plant Biotechnol. J.* **2013**, *11*, 355–361. [[CrossRef](#)]
61. Kim, M.J.; Yang, S.W.; Mao, H.-Z.; Veena, S.P.; Yin, J.-L.; Chua, N.-H. Gene silencing of sugar-dependent 1 (JcSDP1), encoding a patatin-domain triacylglycerol lipase, enhances seed oil accumulation in *Jatropha curcas*. *Biotechnol. Biofuels* **2014**, *7*, 36. [[CrossRef](#)]
62. Kanai, M.; Yamada, T.; Hayashi, M.; Mano, S.; Nishimura, M. Soybean (*Glycine max* L.) triacylglycerol lipase GmSDP1 regulates the quality and quantity of seed oil. *Sci. Rep.* **2019**, *9*, 8924. [[CrossRef](#)]
63. Müntz, K. Proteases and proteolytic cleavage of storage proteins in developing and germinating dicotyledonous seeds. *J. Exp. Bot.* **1996**, *47*, 605–622. [[CrossRef](#)]

