*plants*

MDPI

*Article*

# Genes Associated with the Flax Plant Type (Oil or Fiber) Identified Based on Genome and Transcriptome Sequencing Data

**Liubov V. Povkhova [1,2], Nataliya V. Melnikova [1], Tatiana A. Rozhmina [3], Roman O. Novakovskiy [1], Elena N. Pushkova [1], Ekaterina M. Dvorianinova [1,2], Alexander A. Zhuchenko [3,4], Anastasia M. Kamionskaya [5], George S. Krasnov [1] and Alexey A. Dmitriev [1,*]**

1 Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 119991 Moscow, Russia; povhova.lv@phystech.edu (L.V.P.); mnv-4529264@yandex.ru (N.V.M.); 0legovich46@mail.ru (R.O.N.); pushkova18@gmail.com (E.N.P.); dvorianinova.em@phystech.edu (E.M.D.); gskrasnov@mail.ru (G.S.K.)
2 Moscow Institute of Physics and Technology, 141701 Moscow, Russia
3 Federal Research Center for Bast Fiber Crops, 172002 Torzhok, Russia; tatyana_rozhmina@mail.ru (T.A.R.); ecovilar@mail.ru (A.A.Z.)
4 All-Russian Horticultural Institute for Breeding, Agrotechnology and Nursery, 115598 Moscow, Russia
5 Institute of Bioengineering, Research Center of Biotechnology of the Russian Academy of Sciences, 119071 Moscow, Russia; rifampicin@yandex.ru
* Correspondence: Alex_245@mail.ru

**Abstract:** As a result of the breeding process, there are two main types of flax (*Linum usitatissimum* L.) plants. Linseed is used for obtaining seeds, while fiber flax is used for fiber production. We aimed to identify the genes associated with the flax plant type, which could be important for the formation of agronomically valuable traits. A search for polymorphisms was performed in genes involved in the biosynthesis of cell wall components, lignans, fatty acids, and ion transport based on genome sequencing data for 191 flax varieties. For 143 of the 424 studied genes (*4CL*, *C3′H*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *HCT*, *PAL*, *CTL*, *BGAL*, *ABC*, *HMA*, *DIR*, *PLR*, *UGT*, *TUB*, *CESA*, *RGL*, *FAD*, *SAD*, and *ACT* families), one or more polymorphisms had a strong correlation with the flax type. Based on the transcriptome sequencing data, we evaluated the expression levels for each flax type-associated gene in a wide range of tissues and suggested genes that are important for the formation of linseed or fiber flax traits. Such genes were probably subjected to the selection press and can determine not only the traits of seeds and stems but also the characteristics of the root system or resistance to stresses at a particular stage of development, which indirectly affects the ability of flax plants to produce seeds or fiber.

**Keywords:** flax; *Linum usitatissimum* L.; fiber; seeds; genome and transcriptome sequencing; polymorphisms; gene expression



**Citation:** Povkhova, L.V.; Melnikova, N.V.; Rozhmina, T.A.; Novakovskiy, R.O.; Pushkova, E.N.; Dvorianinova, E.M.; Zhuchenko, A.A.; Kamionskaya, A.M.; Krasnov, G.S.; Dmitriev, A.A. Genes Associated with the Flax Plant Type (Oil or Fiber) Identified Based on Genome and Transcriptome Sequencing Data. *Plants* **2021**, *10*, 2616. https://doi.org/10.3390/plants10122616

Academic Editors: Irina N. Anisimova, Svetlana Goryunova and Eugene Radchenko

Received: 1 November 2021
Accepted: 26 November 2021
Published: 28 November 2021



## 1. Introduction

Flax (*Linum usitatissimum* L.) is traditionally grown for obtaining fiber from stems and oil from seeds [1,2]. The differences in the use of flax products resulted in the appearance of two main varieties: fiber flax and oil flax (linseed). Compared to linseed, fiber flax is taller, has branches only in the upper part of the plant, and produces fewer seeds with lower weight [3]. In addition, fiber flax varieties predominantly have a higher fiber content than oil flax [4]. For linseed, agronomically important traits are associated with seed characteristics, such as size, yield, and biochemical composition, while for fiber flax, fiber properties that attract attention include yield, tensile strength, density, flexibility, and biochemical composition (including the cellulose and lignin contents). Linseed is rich in unsaturated fatty acids (primarily omega-3) and lignans, which are beneficial for health and reduce the risk of cancer and cardiac diseases and is, therefore, used in pharmaceutical

and food products as well as animal feed; flax oil is also a component of paints and varnishes [1,5–14]. Flax bast fiber is rich in cellulose and low in lignin, which makes fiber flax a good source for the production of high-quality textiles, medicine, and promising composite materials for the automobile, aerospace, and packaging industries [12,15–23].

In recent decades, a significant number of studies have aimed to understand the molecular–genetic determination of flax traits. The flax genome of cultivar CDC Bethune was sequenced and assembled to chromosomes (about 316 Mb for 15 chromosomes) [24,25]. Chromosome-level and scaffold-level de novo genome assemblies were also obtained for several flax varieties [26–28]. The identification of genes/genetic markers associated with agronomically important characteristics was performed using SSR (simple sequence repeats) markers [29], reduced-representation genome sequencing [30–35], whole-genome resequencing [36], combined genome-wide association analysis and transcriptome sequencing [37], and a functional approach based on the analysis of gene expression and quality parameters [38]. Transcriptomic studies of flax, including the analysis of gene and microRNA expression in particular tissues, were also carried out for the identification of genes and microRNAs that are important for flax development [37–45] and responses to stresses [46–55]. These studies have made a substantial contribution to our knowledge of the genetic basis of flax traits and have also created significant datasets that can be used in further research.

In addition, studies have identified and characterized particular gene families that likely participate in flax plant processes, including those involved in the biosynthesis of lignin (*4CLs* encode 4-coumarate:CoA ligases; *C3′Hs* encode p-coumarate 3-hydroxylases; *C4Hs* encode cinammate 4-hydroxylases; *CADs* encode cinnamyl alcohol dehydrogenases; *CCRs* encode cinnamoyl CoA reductases; *CCoAOMTs* encode affeoyl CoA 3-O-methyltransferases; *COMTs* encode caffeate/5-hydroxyferulate O-methyl-transferases; *F5Hs* encode 5-hydroxylases; *HCTs* encode hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferases; and *PALs* encode phenylalanine ammonia-lyases) [53,56] and other cell wall components (*CTLs* encode chitinase-like proteins; *BGALs* encode β-galactosidase-like proteins; *TUBs* encode tubulins; *CESAs* encode cellulose synthases; *RGLs* encode rhamnogalacturonan lyases; and *ACTs* encode actins) [42,57–64], the biosynthesis of lignans (*DIRs* encode dirigent proteins; *PLRs* encode pinoresinol–lariciresinol reductases; and *UGTs* encode UDP-glycosyltransferases) [65–68], biosynthesis of fatty acids (*FADs* encode fatty acid desaturases and *SADs* encode stearoyl-ACP desaturases) [69–72], and the transport of ions, lipids, and carbohydrates (*ABCs* encode ATP binding cassette transporters and *HMAs* encode heavy metal-associated proteins) [73]. However, it is still not clear which genes from particular families play key roles in the determination of important flax plant characteristics, which polymorphisms have the greatest impact on agronomic traits, and how gene expression affects the manifestation of a trait. A complex approach should be used to solve these issues, and novel and previously obtained data should be included in the analysis. The present study aimed to identify genes that are likely essential for the formation of a particular type of flax plant (fiber flax or linseed) based on genome and transcriptome sequencing data.

## 2. Results

### *2.1. Gene Polymorphisms*

In the present work, we focused our attention on genes that are important for cell wall formation, because these genes are crucial for the determination of flax stem traits; genes involved in the biosynthesis of lignans and fatty acids, the content of which is an important characteristic of seeds; and ABC transporter and heavy metal–associated genes, which participate in numerous processes in flax plants. We performed a search for polymorphisms among 424 genes (66 lignin biosynthesis genes (*4CL*, *C3′H*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *HCT*, and *PAL*), 35 *CTL* genes, 40 *BGAL* genes, 206 *ABC* and *HMA* genes, 9 lignan biosynthesis genes (*DIR*, *PLR*, and *UGT*), 21 *TUB* genes, 16 *CESA* genes, 10 *RGL* genes, 6 fatty acid biosynthesis genes (*FAD* and *SAD*), and 15 *ACT* genes) based

on genome sequencing data for 191 flax varieties (79 fiber flax and 112 linseed varieties) from the NCBI database (PRJNA590636 and PRJNA478805). The results are presented in Supplementary Table S1.

For lignin-related genes, we identified 2703 polymorphisms; for *CTL* genes, 1053; for *BGAL* genes, 2531; for *ABC* and *HMA* genes, 12,598; for lignan-related genes, 559; for *TUB* genes, 585; for *CESA* genes, 671; for *RGL* genes, 546; for fatty acid-related genes, 304; and for *ACT* genes, 208. Data on the number of polymorphisms for individual genes taking into account the length of the analyzed region (gene length + 1000 bp, 500 bp upstream and 500 bp downstream) are presented in Figure 1 and Supplementary Table S2.

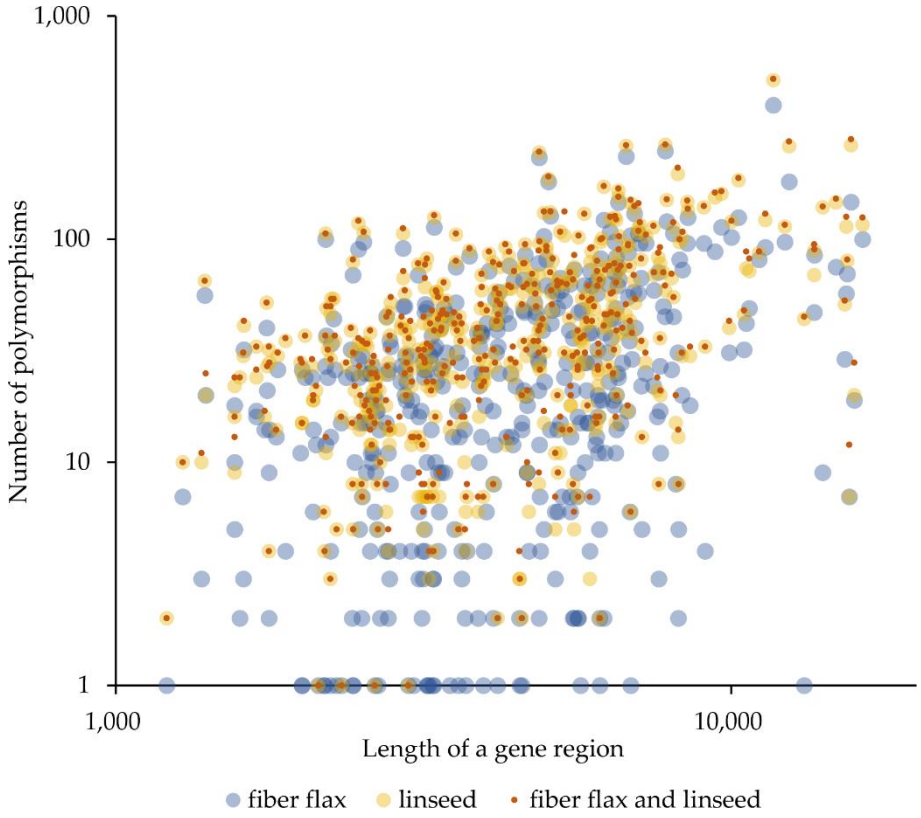


**Figure 1.** Data on the number of polymorphisms for individual genes of the *4CL*, *C3′H*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *HCT*, *PAL*, *CTL*, *BGAL*, *ABC*, *HMA*, *DIR*, *PLR*, *UGT*, *TUB*, *CESA*, *RGL*, *FAD*, *SAD*, and *ACT* families taking into account the length of the analyzed region (gene length + 1000 bp, 500 bp upstream and 500 bp downstream).

The most polymorphic genes were *ABCG11*, *F5H8*, *ABCH10*, *CTL9*, *HMA6*, *F5H7*, *ABCG68*, *4CL5*, *ABCG47*, and *ABCG16* (number of polymorphisms per 1 kb ranged from 37.8 to 50.6, Supplementary Table S2). As can be seen from Figure 1 and Supplementary Table S2, there were fewer polymorphisms in the analyzed fiber flax varieties than in the linseed varieties. We evaluated the genetic similarity in fiber flax and linseed groups based on polymorphisms in the studied 424 genes and confirmed that the linseed group was more polymorphic (Figure 2 and Supplementary Table S3).

Then, clustering of flax varieties was performed based on the identified polymorphisms for the studied gene families; the results are presented in Figure 3 and Supplementary Figure S1.

For genes involved in lignin synthesis, three clusters were revealed: the first included predominantly linseed varieties; the second, predominantly fiber flax varieties; and the third, both linseed and fiber flax varieties (Figure 3a and Supplementary Figure S1a). For *CTL* genes, we also observed a cluster with the predominance of linseed, a cluster with the prevalence of fiber flax, and a mixed cluster (Figure 3b and Supplementary Figure S1b). For

*BGAL* genes, sufficiently clear differentiation of flax varieties for a linseed group and a fiber flax group was revealed (Figure 3c and Supplementary Figure S1c); the same was observed for *ABC* and *HMA* genes (Figure 3d and Supplementary Figure S1d). For lignan-related genes, a cluster of predominantly linseed varieties emerged; the second cluster included mostly fiber flax varieties, while in the third cluster, differentiation according to flax type was not that clear (Figure 3e and Supplementary Figure S1e). The same was observed for *TUB* genes (Figure 3f and Supplementary Figure S1f). For *CESA* genes and *RGL* genes, association with flax type was revealed for some subclusters, but the majority of varieties formed mixed clusters (Figure 3g,h and Supplementary Figure S1g,h). For fatty acid–associated genes, a cluster of mostly linseed varieties and a cluster of mostly fiber flax ones were formed; however, each cluster included a significant number of varieties of the other type (Figure 3i and Supplementary Figure S1i). For actin genes, there was not clear enough differentiation of fiber flax and linseed varieties (Figure 3j and Supplementary Figure S1j). Thus, with some exceptions, clustering of fiber flax and linseed varieties was observed for most groups of studied genes, and the differentiation based on polymorphisms of *BGAL* genes and *ABC* plus *HMA* genes was considered the best (Figure 3).

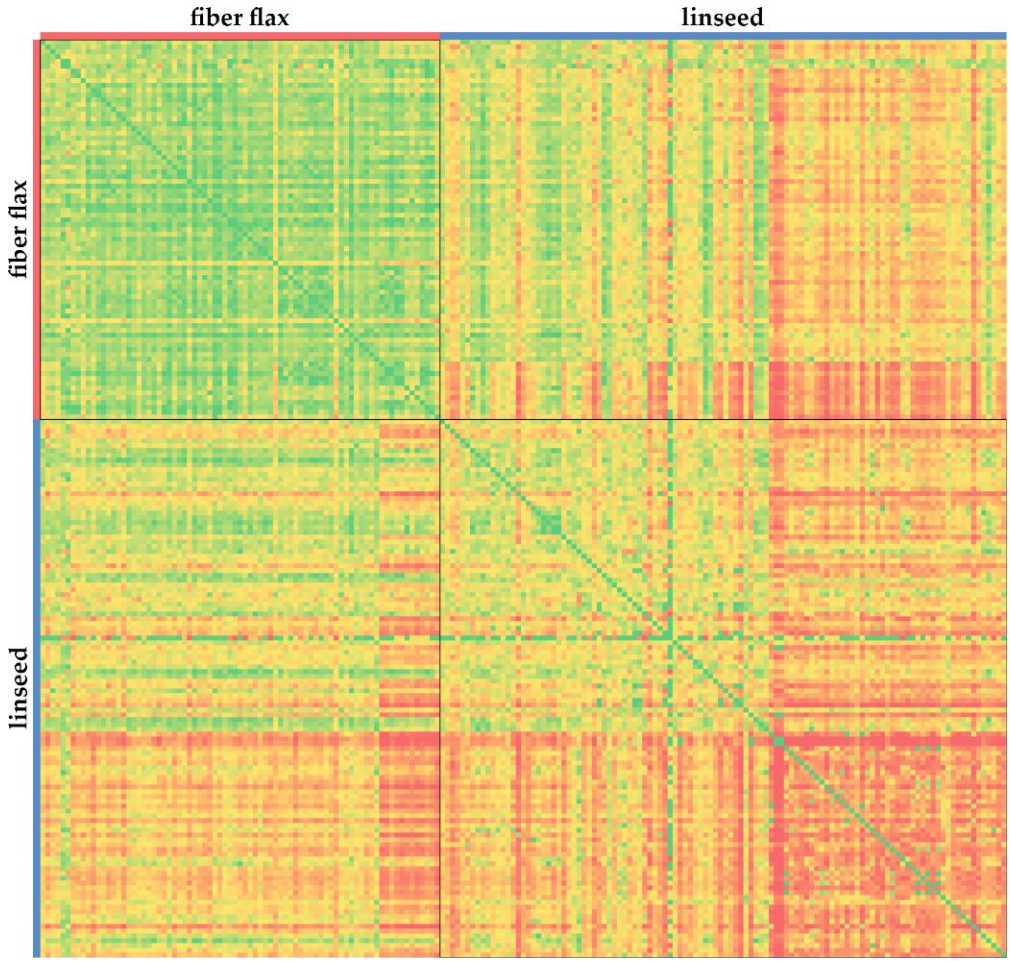


**Figure 2.** Genetic similarity between fiber flax and linseed varieties based on polymorphisms in the studied genes of the *4CL*, *C3′H*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *HCT*, *PAL*, *CTL*, *BGAL*, *ABC*, *HMA*, *DIR*, *PLR*, *UGT*, *TUB*, *CESA*, *RGL*, *FAD*, *SAD*, and *ACT* families. The green–yellow–red color scale indicates the level of genetic similarity between varieties, from the highest (green) to the lowest (red).

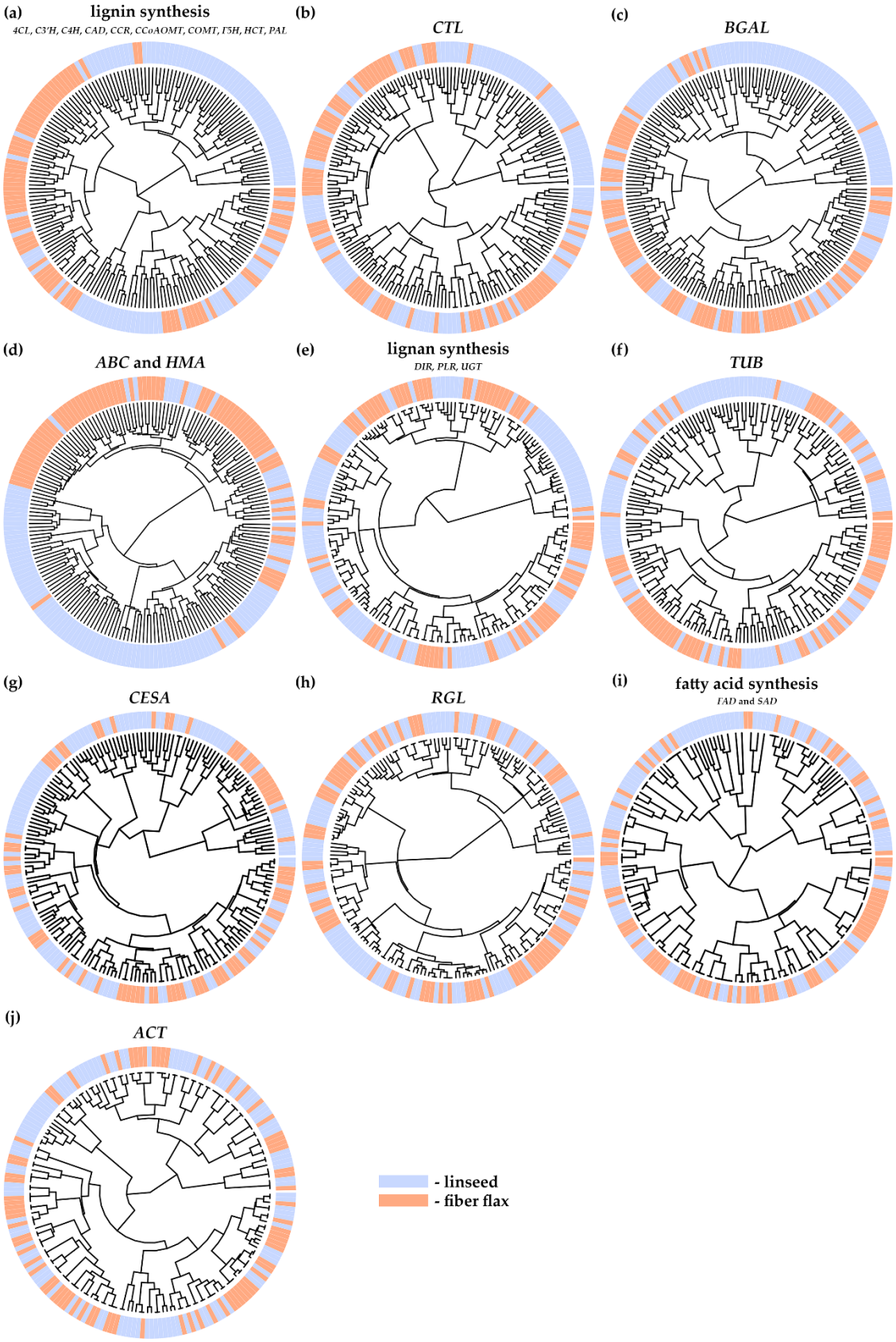


**Figure 3.** Clustering of linseed and fiber flax varieties based on polymorphisms in lignin-associated (*4CL*, *C3′H*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *HCT*, and *PAL*) (**a**), *CTL* (**b**), *BGAL* (**c**), *ABC* and *HMA* (**d**), lignan-associated (*DIR*, *PLR*, and *UGT*) (**e**), *TUB* (**f**), *CESA* (**g**), *RGL* (**h**), fatty acid–associated (*FAD* and *SAD*) (**i**), and *ACT* (**j**) gene families.

### *2.2. Analysis of Genes Associated with Flax Type*

#### 2.2.1. Genes with Flax Type–Associated Polymorphisms

For the identification of genes potentially associated with flax type, correlation analysis between variant allele frequency (VAF) values and belonging of a variety to a linseed or fiber flax group was performed. For 143 of the 424 studied genes, one or more polymorphisms had a Spearman's correlation coefficient ($r_s$) value $\geq 0.4$ or $\leq -0.4$ with flax type (Supplementary Table S4). Data on the number of flax type-associated (FTA) polymorphisms for a particular gene are presented in Table 1. The number of polymorphisms with a high correlation coefficient significantly varied among the studied genes, ranging from 1 to 39. Several genes had a very strong correlation with flax type ($r_s \geq 0.6$ or $\leq -0.6$): *PAL1*, *PAL3*, *BGAL30*, *ABCB40*, *ABCB42*, *ABCB45*, *ABCB47*, *ABCC4*, *ABCG8*, *ABCG79*, *ABCH1*, *HMA12*, *PLR1*, and *CESA1-B*. These genes could play an important role in the formation of traits specific to linseed or fiber flax and are of particular interest for further analysis.

**Table 1.** Number of flax type-associated polymorphisms ($r_s \geq 0.4$ or $\leq -0.4$) for genes of the *4CL*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *PAL*, *CTL*, *BGAL*, *ABC*, *HMA*, *PLR*, *TUB*, *CESA*, *RGL*, and *FAD* families.

| Gene | Number of FTA Polymorphisms | Gene | Number of FTA Polymorphisms | Gene | Number of FTA Polymorphisms | Gene | Number of FTA Polymorphisms |
|---|---|---|---|---|---|---|---|
| **Lignin synthesis** | | *BGAL30* | 21 | *ABCB45* | 19 | *ABCG68* | 2 |
| *4CL1* | 32 | *BGAL31* | 1 | *ABCB46* | 7 | *ABCG69* | 4 |
| *4CL4* | 2 | *BGAL32* | 4 | *ABCB47* | 30 | *ABCG71* | 30 |
| *4CL5* | 3 | *BGAL33* | 6 | *ABCB48* | 5 | *ABCG72* | 2 |
| *C4H3* | 2 | *BGAL35* | 1 | *ABCB7* | 6 | *ABCG73* | 15 |
| *C4H4* | 9 | *BGAL37* | 7 | *ABCC10* | 19 | *ABCG75* | 1 |
| *CAD1A* | 1 | *BGAL40* | 25 | *ABCC16* | 1 | *ABCG79* | 15 |
| *CAD1B* | 2 | *BGAL41* | 1 | *ABCC18* | 2 | *ABCG8* | 19 |
| *CAD4A* | 1 | *BGAL6* | 12 | *ABCC4* | 5 | *ABCG80* | 39 |
| *CAD4B* | 4 | *BGAL7* | 4 | *ABCC6* | 1 | *ABCG83* | 10 |
| *CAD7* | 2 | *BGAL9* | 4 | *ABCC5* | 1 | *ABCH1* | 3 |
| *CCR11* | 1 | ***ABC* and *HMA*** | | *ABCF8* | 2 | *ABCH10* | 15 |
| *CCR4* | 3 | *ABCA1* | 32 | *ABCG1* | 2 | *ABCH11* | 4 |
| *CCoAOMT5* | 2 | *ABCA2* | 16 | *ABCG11* | 3 | *ABCH12* | 1 |
| *COMT2* | 5 | *ABCA3* | 1 | *ABCG12* | 2 | *ABCH8* | 1 |
| *COMT3* | 2 | *ABCA4* | 1 | *ABCG13* | 3 | *HMA12* | 18 |
| *F5H1* | 2 | *ABCA5* | 11 | *ABCG14* | 8 | *HMA2* | 2 |
| *F5H7* | 1 | *ABCA6* | 2 | *ABCG16* | 15 | *HMA3* | 2 |
| *PAL1* | 2 | *ABCA7* | 27 | *ABCG22* | 6 | *HMA4* | 9 |
| *PAL3* | 3 | *ABCA8* | 12 | *ABCG24* | 1 | *HMA6* | 13 |
| ***CTL*** | | *ABCB1* | 3 | *ABCG25* | 7 | **Lignan synthesis** | |
| *CTL1* | 16 | *ABCB12* | 1 | *ABCG33* | 4 | *PLR1* | 50 |
| *CTL10* | 6 | *ABCB13* | 10 | *ABCG35* | 15 | ***TUB*** | |
| *CTL13* | 8 | *ABCB16* | 2 | *ABCG36* | 4 | *Alfa_TUB2* | 1 |
| *CTL18* | 11 | *ABCB2* | 1 | *ABCG37* | 1 | *Beta_TUB13* | 1 |
| *CTL2* | 7 | *ABCB22* | 2 | *ABCG4* | 3 | *Beta_TUB3* | 8 |
| *CTL22* | *1* | *ABCB23* | 1 | *ABCG40* | 1 | *Beta_TUB6* | 6 |
| *CTL23* | *7* | *ABCB25* | 1 | *ABCG47* | 8 | *Beta_TUB7* | 2 |
| *CTL24* | *3* | *ABCB26* | 1 | *ABCG52* | 2 | ***CESA*** | |
| *CTL26* | 2 | *ABCB29* | 23 | *ABCG56* | 13 | *CESA1-B* | 1 |
| *CTL35* | 2 | *ABCB3* | 2 | *ABCG57* | 2 | *CESA3-A* | 1 |
| *CTL4* | 1 | *ABCB32* | 2 | *ABCG58* | 12 | *CESA4* | 2 |
| ***BGAL*** | | *ABCB33* | 3 | *ABCG59* | 1 | *CESA8-A* | 1 |
| *BGAL1* | 2 | *ABCB37* | 1 | *ABCG6* | 4 | ***RGL*** | |
| *BGAL10* | 1 | *ABCB39* | 1 | *ABCG60* | 2 | *RGL1_B* | 1 |
| *BGAL2* | 11 | *ABCB40* | 22 | *ABCG61* | 1 | *RGL4_B* | 1 |
| *BGAL23* | 1 | *ABCB42* | 27 | *ABCG62* | 2 | **Fatty acid synthesis** | |
| *BGAL27* | 13 | *ABCB43* | 2 | *ABCG64* | 4 | *FAD2A* | 1 |

Note: If no FTA (flax type-associated) polymorphisms were identified for a particular gene, this gene was not presented in the table.

The proportion of genes with polymorphisms with a strong correlation ($r_s \geq 0.4$ or $\leq -0.4$) with flax type varied significantly between the studied groups of genes. For groups of *BGAL* genes and *ABC* plus *HMA* genes, a significant proportion (about 40%) of the studied genes had FTA polymorphisms. In groups of *CESA*, *TUB*, *CTL*, and *RGL* genes and genes involved in lignin synthesis, about 20–31% of the studied genes had polymorphisms with a strong correlation with flax type. In groups of fatty acid and lignan synthesis genes, fewer than 20% of the studied genes had FTA polymorphisms, while in the *ACT* group, there were no such genes. In general, these results were in concordance with those observed in the cluster analysis described above; the more polymorphisms with a high correlation coefficient in a group, the clearer the flax type-associated clusters formed for this group.

We also evaluated the expression levels of the studied genes (based on our and NCBI RNA-Seq data) with the focus on genes with polymorphisms with a strong correlation with flax type. The following types of flax tissues were included in the analysis: roots and shoots of seedlings, leaves, flowers, and stems of adult plants [74], capsules (obtained in the current study, PRJNA634481), cortical parenchyma (cPAR), intrusively growing fibers (iFIB) [43], fibers depositing tertiary cell wall (tFIB) [41,44], xylem part (sXYL) [44,45], and embryo (PRJNA720521). Data are presented in Supplementary Table S5 and Supplementary Figures S2 and S3 as heatmaps. Heatmaps revealed groups of genes that were predominantly expressed in particular flax tissues, suggesting their role in the formation of key characteristics of fiber flax or linseed. For example, genes with high expression levels in stem tissues could be crucial for fiber flax traits, while genes with high expression levels in capsules could be of paramount importance for linseed characteristics.

#### 2.2.2. Genes Involved in Lignin Synthesis

Among the genes related to lignin synthesis, *4CL1*, *4CL4*, *4CL5*, *C4H3*, *C4H4*, *CAD1A*, *CAD1B*, *CAD4A*, *CAD4B*, *CAD7*, *CCR11*, *CCR4*, *CCoAOMT5*, *COMT2*, *COMT3*, *F5H1*, *F5H7*, *PAL1*, and *PAL3* had polymorphisms with a strong correlation ($r_s \geq 0.4$ or $\leq -0.4$) with flax type. For the majority of these genes, higher expression levels were observed in seedling roots, sXYL, and capsules. Within this group of genes, *PAL1* and *PAL3* had polymorphisms with a very strong correlation ($r_s \geq 0.6$ or $\leq -0.6$) with flax type (1 and 2 polymorphisms respectively). *PAL1* was predominantly expressed in sXYL and could be associated with lignin synthesis in flax stems, while *PAL3* was predominantly expressed in seedling roots and could be associated with the root traits important for the formation of a particular flax plant type. We performed clustering based on the VAF of *PAL1* and *PAL3*; however, we did not reveal a clear enough association of clusters with flax type (Supplementary Figure S4a,b). This could be due to the fact that there are a significant number of various allelic variants containing these polymorphisms with a very strong correlation with flax type. *PAL* genes encode phenylalanine ammonia-lyases, which deaminate phenylalanine (the initial substrate in the lignin biosynthesis pathway) and result in cinnamic acid formation [75]. *PAL* is one of the key players in lignin synthesis [76], and downregulation or disruption of *PAL* results in a reduction in the lignin content and a change in the lignin composition in *Arabidopsis thaliana* [77], *Medicago sativa* [78] *Populus trichocarpa* [79], and *Nicotiana tabacum* [80,81].

The highest number (32) of polymorphisms with a strong correlation ($r_s \geq 0.4$ or $\leq -0.4$) with flax type among the lignin synthesis genes was identified for *4CL1*. This gene was expressed more highly in seedling roots and sXYL than in other analyzed tissues. A cluster of predominantly linseed varieties and a cluster of predominantly fiber flax varieties were observed in the dendrogram based on the VAF of *4CL1* (Supplementary Figure S4c). *C4H4* also contained a significant number (9) of FTA polymorphisms. This gene was mostly expressed in seedling roots and sXYL. A cluster of predominantly linseed varieties, a cluster of predominantly fiber flax varieties, and a mixed cluster were observed based on the VAF of *C4H4* (Supplementary Figure S4d). A large number of FTA polymorphisms made a significant contribution to the separation of samples according to the flax type; however, the majority of these polymorphisms, most likely, do not affect the trait associated with the

flax type but are inherited linked to a key polymorphism(s). It is known that *C4H* genes encode cinnamate 4-hydroxylases, which catalyze the second step in lignin biosynthesis (hydroxylation of cinnamic acid), while *4CL* genes, which encode 4-coumarate:CoA ligases, are responsible for the formation of *p*-coumaroyl CoA from *p*–coumaric acid [75]. *C4Hs* and *4CLs* along with *PALs* play an important role in lignin formation; the downregulation of *C4Hs* and *4CLs* was reported to decrease the lignin content [75,76,82].

Five *CAD* genes (*CAD1A*, *CAD1B*, *CAD4A*, *CAD4B*, and *CAD7*) of the thirteen studied had polymorphisms with a strong correlation ($r_s \geq 0.4$ or $\leq -0.4$) with flax type. These five genes had increased expression levels in seedling roots; in addition, *CAD1A* and *CAD1B* were also highly expressed in sXYL, while *CAD4A* was also highly expressed in stems; these genes are of the most interest because they could be involved in the determination of the lignin content in flax stems. Among these three genes, clustering based on the VAF of *CAD1B* had the greatest concordance with flax type (Supplementary Figure S4e). *CAD* genes encode cinnamyl alcohol dehydrogenases, which catalyze the reduction of hydroxycinnamyl aldehydes into monolignols [56,75]. The downregulation of *CADs* resulted in changes in lignin composition in poplar [83] and cotton [84]. The brown-midrib phenotype of stems was revealed in flax *CAD* mutants [85]. In addition, the role of *CAD* genes in responses to stresses in plants, including flax, was also revealed [53,86–89].

The association of the polymorphisms of genes involved in the synthesis of lignin with flax type was likely due to the low content of lignin in the stem, which is an important characteristic of fiber flax. In addition, as alterations in lignin biosynthesis are implicated in the regulation of plant growth and defense [90], genes with FTA polymorphisms could be essential for stages of flax plant development that have different levels of importance for the formation of fiber flax and linseed yield and for responses to stresses that have a diverse impact on the products of linseed and fiber flax. Thus, expression analysis of genes with FTA polymorphisms could improve our understanding of which processes these genes are involved in.

#### 2.2.3. Cellulose Synthases

For genes encoding cellulose synthases, a strong correlation ($r_s \geq 0.4$ or $\leq -0.4$) of polymorphisms with flax type was observed for *CESA1-B*, *CESA3-A*, *CESA4*, and *CESA8-A* (from one to two polymorphisms). These genes had increased expression levels in sXYL, and *CESA1-B* was also highly expressed in seedling roots. For *CESA1-B*, one polymorphism had a very strong correlation ($r_s = -0.63$) with flax type. Clustering based on the VAF of *CESA1-B* revealed a cluster of predominantly linseed varieties and a mixed cluster (Supplementary Figure S4f). *CESA1* and *CESA3* are involved in the cellulose synthesis of the primary cell wall, while *CESA4* and *CESA8* participate in the cellulose synthesis of the secondary cell wall [91]. Numerous studies on *CESA* genes indicated their role in flax stem formation [38,42,57–61]. Thus, the presence of FTA polymorphisms in *CESA1-B*, *CESA3-A*, *CESA4*, and *CESA8-A* is likely explained by their role in the determination of flax stem properties, which are especially important for fiber flax.

#### 2.2.4. Chitinase-Like Proteins

Among the genes encoding chitinase-like proteins, polymorphisms with a strong correlation ($r_s \geq 0.4$ or $\leq -0.4$) with flax type were identified for *CTL1*, *CTL10*, *CTL13*, *CTL18*, *CTL2*, *CTL22*, *CTL23*, *CTL24*, *CTL26*, *CTL35*, and *CTL4.* The expression profiles of these genes in different tissues were quite dissimilar: *CTL1*, *CTL2*, and *CTL23* had increased expression levels in sXYL; *CTL23*, *CTL24*, *CTL26*, *CTL22*, and *CTL4* in capsules; *CTL13* and *CTL4* in seedling roots; and *CTL10* in leaves and flowers. Meanwhile, the expression of *CTL35* and *CTL18* was low in all analyzed tissues (CPM < 10 in each sample). The greatest number of FTA polymorphisms was revealed for *CTL1* and *CTL18*. In clustering based on the VAF of *CTL1*, a small cluster of predominantly linseed varieties and a large mixed cluster were observed (Supplementary Figure S4g). For *CTL18*, two clusters were revealed: in the first cluster, the majority of linseed varieties were included, and in the

second one, fiber flax varieties prevailed. However, each cluster included a significant number of varieties of the other type (Supplementary Figure S4h). Considering that *CTL18* expression was low in all of the studied tissues, its role in the determination of traits associated with flax type is not clear, but it could be linked to a gene important for the formation of these traits. Chitinases are involved in plant stress response, development, and cell wall synthesis [92–94]. It was revealed that *CTL1* expression in flax was significantly higher in stem tissues in which thickening of the cell walls occurred, and coexpression of this gene and *CESA4*, *CESA7*, and *CESA8*, which are involved in the formation of the secondary cell wall, was observed, which could indicate the role of *CTL1* in flax cell wall synthesis [60]. Thus, *CTL1* is probably important for the formation of flax stem and, for this reason, polymorphisms of *CTL1* were associated with flax type.

#### 2.2.5. Tubulins

For tubulin-encoding genes, a strong correlation ($r_s \geq 0.4$ or $\leq -0.4$) of polymorphisms with flax type was revealed for *Alfa_TUB2*, *Beta_TUB13*, *Beta_TUB3*, *Beta_TUB6*, and *Beta_TUB7*. All of these genes had increased expression levels in seedling roots and stem tissues. Several of these genes were also highly expressed in other tissue types, but there was a significant variation in such tissue types between genes. The greatest number (8) of FTA polymorphisms was identified for *Beta_TUB3*, and clustering based on the VAF of this gene had a relatively high concordance with flax type (Supplementary Figure S4i). Tubulins are involved in numerous processes in flax plants, including cell wall formation [59,64], and thus could influence traits that play an important role in the formation of stems. However, given the involvement of tubulins in numerous processes in plants, they also could participate in the formation of other traits that are important for fiber flax or linseed plants.

#### 2.2.6. β-Galactosidases

Among the 40 studied genes encoding β-galactosidase-like proteins, 16 had polymorphisms with a strong correlation ($r_s \geq 0.4$ or $\leq -0.4$) with flax type (the number of such polymorphisms varied from 1 to 25). The expression profiles of these genes also varied; however, most genes had the highest expression levels in leaves, flowers, iFIB, and tFIB. *BGAL40* had the highest number (25) of FTA polymorphisms, and clustering of varieties based on the VAF of this gene was in high concordance with the division of varieties into fiber flax and linseed (Supplementary Figure S4j). This gene was predominantly expressed in leaves. For *BGAL30*, polymorphisms with a very strong correlation ($r_s \geq 0.6$ or $\leq -0.6$) with flax type were revealed, and the highest expression level of this gene was observed in flowers. However, clustering based on the VAF of *BGAL30* did not show clear enough differentiation of clusters according to flax type (Supplementary Figure S4k); the influence of other polymorphisms was probably strong. It was shown that β-galactosidases are involved in the modification of cell wall polysaccharides [95,96]. In flax, β-galactosidases are important for the development of the secondary cell wall [58,97]; therefore, genes encoding β-galactosidase-like proteins are likely implicated in the formation of traits of flax stems, which are especially important for fiber flax, and, for this reason, we identified FTA polymorphisms in a large number of *BGALs*.

#### 2.2.7. Rhamnogalacturonan Lyases

Among rhamnogalacturonan lyases-encoding genes, *RGL1_B* and *RGL4_B* each had one polymorphism associated with flax type ($r_s \geq 0.4$ or $\leq -0.4$). *RGL1_B* was predominantly expressed in sXYL, while the expression of *RGL4_B* was higher in leaves, flowers, and capsules; however, the expression level of this gene was also significant in all other studied tissues. It is known that rhamnogalacturonan lyases are involved in the modification of cell wall polysaccharides (RGLs degrade the rhamnogalacturonan I backbone) [98]. Transcriptomic studies of flax showed that rhamnogalacturonan lyases are probably involved in tertiary cell wall formation [41,57]. Based on gene expression data, it was suggested that

*RGL1_B* plays role in the modification of cell wall polysaccharides in xylem tissues [57]. The identification of FTA polymorphism in *RGL1_B* could also indicate a role for this gene in flax cell wall formation and its importance for stem traits of fiber flax.

#### 2.2.8. Genes Involved in Lignan Synthesis

For genes related to lignan synthesis, *PLR1* had many FTA polymorphisms. This gene was predominantly expressed in capsules, and clustering based on the VAF of *PLR1* revealed clear enough differentiation of clusters according to flax type (Supplementary Figure S4l). In addition, one polymorphism in this gene had a very strong correlation ($r_s$ = 0.64) with flax type. It was shown that *PLRs* are important for the synthesis of the major lignan in flax seeds, secoisolariciresinol diglucoside (SDG) [67,99]. SDG has antioxidative and anti-inflammatory capacities and reduces the risk of cancer and cardiac diseases [8]. PLR1 contributes to the synthesis of (+)-secoisolariciresinol by catalyzing the conversion of (−)-pinoresinol into (−)-lariciresinol and is necessary for the accumulation of SDG in flax seeds [67,100,101]. In addition, the implication of *PLR1* in the flax plant stress response was revealed [102]. Thus, a large number of FTA polymorphisms of *PLR1* could be due to its importance for the formation of a linseed trait, the lignan content. However, the difference in linseed and fiber flax based on *PLR1* polymorphisms could also be associated with the role of this gene in the formation of other characteristics, which are important for a particular flax type, namely, stress response.

#### 2.2.9. Genes Involved in Fatty Acid Synthesis

Among the studied genes related to fatty acid synthesis, *FAD2A*, encoding fatty acid desaturases 2, had one polymorphism associated with flax type ($r_s = -0.44$). The expression of *FAD2A* was higher in capsules and embryo than in other analyzed tissues, and clustering had some concordance with flax type, but it was poor enough (Supplementary Figure S4m). It is known that FAD2 catalyzes the desaturation of oleic acid into linoleic acid, while FAD3 catalyzes linoleic acid into linolenic acid, and polymorphisms of *FAD* genes are associated with the content of these fatty acids, which was especially clearly shown for *FAD3A* and *FAD3B* genes [34,71,103–105]. However, the present study did not reveal an association between flax type and polymorphisms in *FAD3A* and *FAD3B* genes. This may be due to the fact that most linseed varieties similar to fiber ones did not carry the key polymorphisms in *FAD3* genes that determine a lower content of linolenic acid [106], and the used sample set did not include a significant number of low-linolenic varieties that could form a separate cluster.

#### 2.2.10. ABC Transporter and Heavy Metal-Associated Genes

Within the studied ABC transporter and heavy metal-associated genes, 83 genes had polymorphisms with a strong correlation ($r_s \geq 0.4$ or $\leq -0.4$) with flax type, and the number of such polymorphisms per gene varied from 1 to 39. The expression profiles of these genes in the studied tissues were very different, but in general, most genes had the highest expression levels in seedling roots. For *ABCA1*, *ABCA7*, *ABCB42*, *ABCB47*, *ABCG71*, and *ABCG80* genes with a high number (27–39) of FTA polymorphisms, VAF clustering revealed a linseed cluster, whose size varied from gene to gene, while the rest clusters were not clear enough differentiated according to flax type (Supplementary Figure S4n,o–q,s), except for *ABCG71*, for which two clusters of predominantly linseed or fiber flax varieties were revealed (Supplementary Figure S4r). For *ABCG79*, *ABCB42*, *ABCH1*, *ABCB40*, *ABCB45*, *ABCB47*, *HMA12*, *ABCG8*, and *ABCC4* genes, polymorphisms with a very strong correlation ($r_s \geq 0.6$ or $\leq -0.6$) with flax type were identified, and the expression profiles of these genes in the studied flax tissues varied greatly. The highest expression level of *ABCG79* was revealed in leaves and seedling shoots; *ABCB42* in seedling roots; *ABCH1* in leaves, capsules, seedling shoots, iFIB, and tFIB; *ABCB40* in flowers; *ABCB45* in seedling roots and sXYL; *HMA12* in sXYL; *ABCG8* in seedling shoots; and *ABCC4* in iFIB. ABC transporters are essential components of plant cell membranes and are involved

in numerous processes, including growth and development, nutrition, and responses to abiotic and biotic stresses [107]. Using a general linear model (GLM), it was shown that the *Lus10016125* gene (*ABCG4* according to the classification by Khan et al. [73]) is probably involved in the determination of flax plant height [31]. In the present study, three polymorphisms of *ABCG4* had a strong association with flax type, but other *ABC* and *HMA* genes had more FTA polymorphisms or polymorphisms with a very strong correlation ($r_s \geq 0.6$ or $\leq -0.6$) with flax type. Among our analyzed groups of genes, the group of *ABC* and *HMA* genes had one of the highest proportions of genes with FTA polymorphisms (40%). As the functions of *ABC* and *HMA* genes vary and their expression profiles were diverse between flax tissues, they could be involved in the determination of traits of different organs of flax plants, which ultimately define features that are important for linseed or fiber flax.

## 3. Discussion

Linseed and fiber flax differ not only in the characteristics of plant parts that are key for agricultural use, namely, the seeds and stems, but also in the anatomy and morphology of the roots (in linseed varieties, in comparison with fiber flax ones, roots penetrate to a greater depth, lateral roots are thicker and longer, the root system has a large absorption area, more distribution to a greater depth, and more developed conducting system), as well as the consumption of nutrients in different periods of plant development (for fiber flax, the critical periods of nutrition correspond to the initial periods of vegetation development, while for linseed, the critical periods are dissimilar and later stages of development are also important), the requirements for the content of nutrients in soil (linseed has higher requirements for nitrogen and phosphorus, whereas fiber flax has a higher requirement for potassium), and the ability to grow under unfavorable conditions (unlike fiber flax, it is more significant for linseed to grow in soils with a higher salt content than in acidic soils) [108]. Therefore, for linseed and fiber flax, different stages of development are critical for the formation of seeds and fiber, respectively, and diverse stresses affect these flax plants in dissimilar ways. In this regard, an important role in the formation of fiber and oil flax plants capable of producing high yields of fiber and seeds, respectively, is probably played by different genes, some of which ensure the optimal development of linseed, while others, fiber flax, due to the distinctive features that are typical for two types of flax plants. Thus, some genes apparently play a role in the formation of stems and seeds, and some genes contribute to the formation of other plant traits, for example, the root system or resistance to specific unfavorable environments at particular stages of development, which indirectly affect the ability of flax plants to produce high levels of high-quality fiber or seeds. Likely, when linseed breeding took place, there was a selection for some alleles of these two groups of genes, and during the breeding of fiber flax, other alleles were preferred since these genes had a different significance for the formation of a flax type with favorable economically valuable traits. Thus, an approach based on the search for genes whose polymorphisms are associated with the flax plant type can provide new knowledge about genes that are important for the development of linseed or fiber flax and, therefore, affect the quality of flax production. Such genes can be used in genomic breeding and marker-assisted selection of flax.

Various approaches are used for the identification of genetic markers or genes associated with important flax plant characteristics. Linkage mapping and association mapping/genome-wide association studies (GWAS) have allowed scientists to identify quantitative trait loci (QTL) and quantitative trait nucleotides (QTN) associated with seed yield and quality traits, fiber traits, agronomic traits, disease resistance, and abiotic stress response in flax [29,31–36,109–116]. In addition to GWAS, transcriptome analysis can be useful to avoid the false-positive results in candidate gene searches; using such a combined approach, candidate genes related to fatty acid synthesis in flax seeds were identified [37]. Another approach that was successfully used for the identification of fiber quality-associated genes was based on the analysis of gene expression in different tissues

of flax to reveal genes with tissue-specific expression in developing fibers and their further analysis [38]. Such an analysis is valuable if there are transcriptomic data for the specific tissues at certain stages of plant development, and it would be preferable if there were data for several genotypes, which are more difficult to obtain. In addition, the appearance of new data (additional tissues or genotypes) may change the results of the analysis. In the present work, we focused on the study of particular gene families, whose role in plant growth, development, and stress response is known and which were identified in the flax genome. First, we searched for genes with polymorphisms associated with the flax plant type based on the data of whole-genome sequencing of a large number of varieties and then analyzed the expression profiles of these genes in various flax tissues. Such an approach allowed us to use the currently available data as efficiently as possible and narrow the range of analysis to families of sufficiently characterized functional genes in order to conduct a more detailed analysis of their possible role in the determination of traits that have different levels of significance for the formation of fiber flax or linseed plants. Genes, the polymorphisms of which had a strong association with the flax plant type, were probably subjected to selection pressure; the alleles that were more common in oil flax differed from those that were characteristic of fiber flax. Moreover, such genes can determine not only the traits of flax seeds and stems but also, for example, the characteristics of the root system or resistance of flax plants to specific stresses at a particular stage of development, which indirectly affect the ability of plants to produce seed or fiber.

## 4. Materials and Methods

### *4.1. Variant Calling and Associative Analysis of Genome Sequencing Data*

For association analysis between allele variants and flax type (fiber flax or linseed), we used Illumina whole-genome sequencing (WGS) data from two NCBI BioProjects: PRJNA590636 (12× average coverage; [36]) and PRJNA478805 (25× average coverage). In total, 191 flax varieties (79 fiber flax and 112 linseed cultivars/lines) were analyzed. The list of varieties can be found in Supplementary Tables S1, S3 and S4. Downloaded reads were trimmed and adapters were removed with Trimmomatic 0.38 [117]. Then, reads were mapped to the reference *L. usitatissimum* genome assembly GCA_000224295.2 (ASM22429v2) by BWA-MEM 0.7.17 [118] with lowered minimum seed length ('-k' argument). Subsequently, sorted BAM files were processed with FixMateInformation (picard-tools 2.21.3) (http://broadinstitute.github.io/picard/, accessed on 27 October 2021). Next, duplicated reads were marked with MarkDuplicatesWithMateCigar (picard-tools). Finally, variant calling was performed by freeBayes 1.3.2 [119] for the joint set of 191 BAM files with the following thresholds: mapping quality 10, base calling quality 15, and minimal alternative allele coverage 4 (maximal value across all samples). The following genes were analyzed: 66 genes related to lignin synthesis—*4CL1*, *4CL2*, *4CL3*, *4CL4*, *4CL5*, *4CL6*, *4CL7*, *4CL8*, *4CL9*, *C3′H1*, *C3′H2*, *C3′H3*, *C4H1*, *C4H2*, *C4H3*, *C4H4*, *C4H5*, *CAD1A*, *CAD1B*, *CAD2A*, *CAD2B*, *CAD3A*, *CAD3B*, *CAD4A*, *CAD4B*, *CAD5A*, *CAD5B*, *CAD6*, *CAD7*, *CAD8*, *CCoAOMT1*, *CCoAOMT2*, *CCoAOMT3*, *CCoAOMT4*, *CCoAOMT5*, *CCR1*, *CCR10*, *CCR11*, *CCR12*, *CCR2*, *CCR3*, *CCR4*, *CCR5*, *CCR6*, *CCR7*, *CCR8*, *CCR9*, *COMT1*, *COMT2*, *COMT3*, *F5H1*, *F5H2*, *F5H3*, *F5H4*, *F5H5*, *F5H6*, *F5H7*, *F5H8*, *HCT1*, *HCT2*, *HCT3*, *HCT4*, *HCT5*, *PAL1*, *PAL2*, and *PAL3*; 35 genes encoding chitinase-like proteins—*CTL1*, *CTL10*, *CTL11*, *CTL12*, *CTL13*, *CTL14*, *CTL16*, *CTL17*, *CTL18*, *CTL19*, *CTL2*, *CTL20*, *CTL21*, *CTL22*, *CTL23*, *CTL24*, *CTL25*, *CTL26*, *CTL27*, *CTL28*, *CTL29*, *CTL3*, *CTL30*, *CTL31*, *CTL32*, *CTL33*, *CTL35*, *CTL36*, *CTL37*, *CTL4*, *CTL5*, *CTL6*, *CTL7*, *CTL8*, and *CTL9*; 40 genes encoding β-galactosidases—*BGAL1*, *BGAL10*, *BGAL11*, *BGAL12*, *BGAL13*, *BGAL14*, *BGAL16*, *BGAL18*, *BGAL19*, *BGAL2*, *BGAL20*, *BGAL21*, *BGAL22*, *BGAL23*, *BGAL24*, *BGAL25*, *BGAL26*, *BGAL27*, *BGAL28*, *BGAL29*, *BGAL3*, *BGAL30*, *BGAL31*, *BGAL32*, *BGAL33*, *BGAL34*, *BGAL35*, *BGAL36*, *BGAL37*, *BGAL38*, *BGAL39*, *BGAL4*, *BGAL40*, *BGAL41*, *BGAL42*, *BGAL43*, *BGAL6*, *BGAL7*, *BGAL8*, and *BGAL9*; 206 genes encoding ABC transporters and heavy metal–associated genes—*ABCA1*, *ABCA2*, *ABCA3*, *ABCA4*, *ABCA5*, *ABCA6*, *ABCA7*, *ABCA8*, *ABCB1*, *ABCB10*, *ABCB11*, *ABCB12*, *ABCB13*, *ABCB14*, *ABCB15*, *ABCB16*, *ABCB17*, *ABCB18*,

*ABCB19*, *ABCB2*, *ABCB20*, *ABCB21*, *ABCB22*, *ABCB23*, *ABCB24*, *ABCB25*, *ABCB26*, *ABCB27*, *ABCB28*, *ABCB29*, *ABCB3*, *ABCB30*, *ABCB31*, *ABCB32*, *ABCB33*, *ABCB34*, *ABCB35*, *ABCB36*, *ABCB37*, *ABCB38*, *ABCB39*, *ABCB4*, *ABCB40*, *ABCB41*, *ABCB42*, *ABCB43*, *ABCB44*, *ABCB45*, *ABCB46*, *ABCB47*, *ABCB48*, *ABCB5*, *ABCB6*, *ABCB7*, *ABCB8*, *ABCB9*, *ABCC1*, *ABCC10*, *ABCC11*, *ABCC12*, *ABCC13*, *ABCC14*, *ABCC15*, *ABCC16*, *ABCC17*, *ABCC18*, *ABCC19*, *ABCC2*, *ABCC3*, *ABCC4*, *ABCC5*, *ABCC6*, *ABCC7*, *ABCC8*, *ABCC9*, *ABCD1*, *ABCD2*, *ABCD3*, *ABCD4*, *ABCD5*, *ABCE1*, *ABCE2*, *ABCF1*, *ABCF2*, *ABCF3*, *ABCF4*, *ABCF5*, *ABCF6*, *ABCF7*, *ABCF8*, *ABCF9*, *ABCG1*, *ABCG11*, *ABCG12*, *ABCG13*, *ABCG14*, *ABCG16*, *ABCG17*, *ABCG18*, *ABCG19*, *ABCG2*, *ABCG20*, *ABCG21*, *ABCG22*, *ABCG23*, *ABCG24*, *ABCG25*, *ABCG26*, *ABCG27*, *ABCG28*, *ABCG29*, *ABCG3*, *ABCG30*, *ABCG31*, *ABCG32*, *ABCG33*, *ABCG34*, *ABCG35*, *ABCG36*, *ABCG37*, *ABCG38*, *ABCG39*, *ABCG4*, *ABCG40*, *ABCG41*, *ABCG42*, *ABCG43*, *ABCG44*, *ABCG45*, *ABCG46*, *ABCG47*, *ABCG48*, *ABCG49*, *ABCG5*, *ABCG50*, *ABCG51*, *ABCG52*, *ABCG53*, *ABCG54*, *ABCG55*, *ABCG56*, *ABCG57*, *ABCG58*, *ABCG59*, *ABCG6*, *ABCG60*, *ABCG61*, *ABCG62*, *ABCG63*, *ABCG64*, *ABCG65*, *ABCG66*, *ABCG67*, *ABCG68*, *ABCG69*, *ABCG7*, *ABCG70*, *ABCG71*, *ABCG72*, *ABCG73*, *ABCG74*, *ABCG75*, *ABCG76*, *ABCG77*, *ABCG78*, *ABCG79*, *ABCG8*, *ABCG80*, *ABCG81*, *ABCG82*, *ABCG83*, *ABCG84*, *ABCG85*, *ABCG9*, *ABCH1*, *ABCH10*, *ABCH11*, *ABCH12*, *ABCH13*, *ABCH14*, *ABCH15*, *ABCH16*, *ABCH17*, *ABCH18*, *ABCH19*, *ABCH2*, *ABCH20*, *ABCH21*, *ABCH22*, *ABCH3*, *ABCH4*, *ABCH5*, *ABCH6*, *ABCH7*, *ABCH8*, *HMA1*, *HMA10*, *HMA11*, *HMA12*, *HMA2*, *HMA3*, *HMA4*, *HMA6*, *HMA7*, *HMA8*, and *HMA9*; 9 genes related to lignan synthesis—*DIR1*, *DIR2*, *DIR3*, *DIR4*, *DIR5*, *DIR6*, *PLR1*, *PLR2*, and *UGT74S1*; 21 genes encoding tubulins—*Alfa_TUB1*, *Alfa_TUB2*, *Alfa_TUB3*, *Alfa_TUB4*, *Alfa_TUB5*, *Alfa_TUB6*, *Beta_TUB1*, *Beta_TUB10*, *Beta_TUB11*, *Beta_TUB12*, *Beta_TUB13*, *Beta_TUB2*, *Beta_TUB3*, *Beta_TUB4*, *Beta_TUB5*, *Beta_TUB6*, *Beta_TUB7*, *Beta_TUB8*, *Beta_TUB9*, *Gamma_Tub1*, and *Gamma_Tub2*; 16 genes encoding cellulose synthases—*CESA1-A*, *CESA1-B*, *CESA3-A*, *CESA3-B*, *CESA3-C*, *CESA4*, *CESA6-A*, *CESA6-B*, *CESA6-C*, *CESA6-D*, *CESA6-E*, *CESA6-F*, *CESA7-A*, *CESA7-B*, *CESA8-A*, and *CESA8-B*; 10 genes encoding rhamnogalacturonate lyases—*RGL1_B*, *RGL2*, *RGL3_A*, *RGL3_B*, *RGL4_A*, *RGL4_B*, *RGL6_A*, *RGL6_B*, *RGL7_A*, and *RGL7_B*; 6 genes related to fatty acid synthesis—*FAD2A*, *FAD2B*, *FAD3A*, *FAD3B*, *SAD1*, and *SAD2*; and 15 genes encoding actins—*Act1*, *Act10*, *Act11*, *Act12*, *Act13*, *Act14*, *Act15*, *Act2*, *Act3*, *Act4*, *Act5*, *Act6*, *Act7*, *Act8*, and *Act9*. Data for gene sequences/genome coordinates were obtained from the literature [38,53,56–58,60,64,65,73,105,120]. Analyzed genome regions included the gene body, 500 bp upstream, and 500 bp downstream. To evaluate the number of polymorphisms per gene, we counted point substitutions and indels that were supported with at least four reads (regardless of variant allele frequency, VAF) and were observed in at least one sample. Such thresholds were chosen because of the low overall genome coverage (only 7× after marking duplicates) for the analyzed samples from NCBI BioProjects PRJNA590636 and PRJNA478805.

Next, we performed correlation analysis between the VAF values and whether a variety belonged to linseed or fiber flax. Spearman's, Pearson's, and Kendall's correlation coefficients and *p*-values were calculated. Additionally, VAF values were compared between these two groups using the nonparametric Mann–Whitney U-test. The derived *p*-values were adjusted for multiple testing with the Benjamini–Hochberg approach. Moreover, based on the similarity of the VAF profiles across variants (Euclidean distance), we performed hierarchical clustering of varieties using Ward's minimum variance method ('ward.D2' in R 3.6.2).

### *4.2. Analysis of Transcriptome Sequencing Data*

For gene expression analysis, we used NCBI BioProjects PRJNA475325 (RNA-Seq analysis of intrusively growing fibers from the flax stem) [43], PRJNA631357 (RNA-Seq analysis of phloem fibers during gravitropic behavior of flax plants) [45], PRJNA720521 (RNA-Seq flax data for the embryo and endosperm), PRJNA663265 (RNA-Seq dataset of multiple organs from flax), and PRJNA634481 (roots and shoots of seedlings and leaves, flowers, and stems of adult plants of six flax varieties; our data published in the data report

format) [74]. In addition, our novel data for flax capsules were also used in the present study. In brief, seven flax varieties (AGT 409/10, AGT 427/10, AGT 981/05, AGT 1568/07, Atlant, LM 98, and Lola) were grown in a greenhouse, and capsules were harvested in the stage of yellow–green ripening. RNA for each variety was extracted from a pool of five capsules using a Quick-RNA Miniprep Kit (Zymo Research, Irvine, CA, USA), cDNA library preparation was carried out with NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs, Hertfordshire, UK), and sequencing was performed on NextSeq 500 (Illumina, San Diego, CA, USA) with a read length of 86 bp, as described in our previous study [74].

In total, 69 samples were included in the analysis. To eliminate mapping bias, reads from all samples were cropped to the minimal length across all samples (86 nt) with Trimmomatic. When paired-end read data were available, only forward read data were used. Next, reads were mapped to the reference *L. usitatissimum* genome assembly GCA_000224295.2 (ASM22429v2) with splice-aware STAR 2.7.2b mapper [121]. Since this reference assembly was not annotated, we relied only on de novo splice junction discovery and launched STAR in 2-pass mode. After the first pass, we collected all found splice junctions from all the samples and supplied this list for the second STAR pass.

The derived BAM files were sorted with samtools 1.10 and reordered, and read groups were assigned with picard-tools. Then, reads that contained Ns in their CIGAR field (e.g., spanning splicing junctions) were split with SplitNCigarReads from GATK 4.2.2.0 [122]. Next, read counts "per gene" were evaluated with bedtools multcov 2.26.0. We used a BED file containing the list of 424 regions containing genes, which were also applied in the search for FTA polymorphisms. We used preliminary splitting reads by splice junctions (with SplitNCigarReads tool) to exclude from counting those reads that, without mapping themselves to the region of interest (from the BED file), cross this region "as an intron", with their first half mapping upstream of the region and the second half mapping downstream of the region. There were many such reads. The derived read counts were analyzed in edgeR [123] and normalized by the total number of mapped reads per sample. Heatmaps of expression levels of the studied genes in different flax tissues were created in the R environment using the pheatmap package. Expression values were log2-transformed and then normalized to the average value across all the samples (per each gene). Hierarchical clustering of genes and tissues was done using Ward's minimum variance method ('ward.D2').

**Supplementary Materials:** The following are available online at https://www.mdpi.com/article/10.3390/plants10122616/s1, Figure S1: Clustering of linseed and fiber flax varieties based on polymorphisms in lignin-associated (*4CL*, *C3′H*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *HCT*, and *PAL*) (a), *CTL* (b), *BGAL* (c), *ABC* and *HMA* (d), lignan-associated (*DIR*, *PLR*, and *UGT)* (e), *TUB* (f), *CESA* (g), *RGL* (h), fatty acid-associated (*FAD* and *SAD*) (i), and *ACT* (j) gene families; Figure S2: Heatmap of expression levels of genes of *4CL*, *C3′H*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *HCT*, *PAL*, *CTL*, *BGAL*, *ABC*, *HMA*, *DIR*, *PLR*, *UGT*, *TUB*, *CESA*, *RGL*, *FAD*, *SAD*, and *ACT* families in different flax tissues (averaged by tissue type); Figure S3: Heatmap of expression levels of genes of *4CL*, *C3′H*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *HCT*, *PAL*, *CTL*, *BGAL*, *ABC*, *HMA*, *DIR*, *PLR*, *UGT*, *TUB*, *CESA*, *RGL*, *FAD*, *SAD*, and *ACT* families in different flax tissues (individual samples); Figure S4: Clustering of linseed and fiber flax varieties based on polymorphisms in *PAL1* (a), *PAL3* (b), *4CL1* (c), *C4H4* (d), *CAD1B* (e), *CESA1-B* (f), *CTL1* (g), *CTL18* (h), *Beta_TUB3* (i), *BGAL40* (j), *BGAL30* (k), *PLR1* (l), *FAD2A* (m), *ABCA1* (n), *ABCA7* (o), *ABCB42* (p), *ABCB47* (q), *ABCG71* (r), and *ABCG80* (s) genes; Table S1: Polymorphisms in the studied genes of *4CL*, *C3′H*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *HCT*, *PAL*, *CTL*, *BGAL*, *ABC*, *HMA*, *DIR*, *PLR*, *UGT*, *TUB*, *CESA*, *RGL*, *FAD*, *SAD*, and *ACT* families in 191 flax varieties; Table S2: Data on the number of polymorphisms for individual genes of *4CL*, *C3′H*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *HCT*, *PAL*, *CTL*, *BGAL*, *ABC*, *HMA*, *DIR*, *PLR*, *UGT*, *TUB*, *CESA*, *RGL*, *FAD*, *SAD*, and *ACT* families taking into account the length of the analyzed region (gene length + 1000 bp, 500 bp upstream and 500 bp downstream); Table S3: Genetic similarity between fiber flax and linseed varieties based on polymorphisms in the studied genes of *4CL*, *C3′H*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *HCT*, *PAL*, *CTL*, *BGAL*, *ABC*, *HMA*, *DIR*, *PLR*,

*UGT*, *TUB*, *CESA*, *RGL*, *FAD*, *SAD*, and *ACT* families; Table S4: Polymorphisms in the genes of *4CL*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *PAL*, *CTL*, *BGAL*, *ABC*, *HMA*, *PLR*, *TUB*, *CESA*, *RGL*, and *FAD* families with a strong correlation ($r_s \geq 0.4$ or $\leq -0.4$) with flax type; Table S5: Expression levels of the genes of *4CL*, *C3′H*, *C4H*, *CAD*, *CCR*, *CCoAOMT*, *COMT*, *F5H*, *HCT*, *PAL*, *CTL*, *BGAL*, *ABC*, *HMA*, *DIR*, *PLR*, *UGT*, *TUB*, *CESA*, *RGL*, *FAD*, *SAD*, and *ACT* families in different flax tissues.

**Author Contributions:** Conceptualization, N.V.M. and A.A.D.; performing experiments, L.V.P., T.A.R., R.O.N. and E.N.P.; data analysis, L.V.P., N.V.M., E.M.D., A.A.Z., A.M.K., G.S.K. and A.A.D.; writing, N.V.M., G.S.K. and A.A.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Ministry of Science and Higher Education of the Russian Federation in accordance with agreement number 075-15-2020-907, 16 November 2020, on providing a grant in the form of subsidies from the Federal budget of the Russian Federation. The grant was provided for state support for the creation and development of a World-class Scientific Center "Agrotechnologies for the Future".

**Data Availability Statement:** The raw sequencing data have been deposited in the NCBI Sequence Read Archive (SRA) under the BioProject accession number PRJNA634481.

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

## References

1. Muir, A.D.; Westcott, N.D. *Flax: The Genus Linum*; CRC Press: Boca Raton, FL, USA, 2003.
2. Weiss, E.; Zohary, D.; Hopf, M. *Domestication of Plants in the Old World—The Origin and Spread of Domesticated Plants in South-West Asia, Europe, and the Mediterranean Basin*; Oxford Scholarship: Oxford, UK, 2012.
3. Diederichsen, A.; Richards, K. *Cultivated Flax and the Genus linum L.: Taxonomy and Germplasm Conservation*; CRC Press: Boca Raton, FL, USA, 2003; pp. 22–54.
4. Diederichsen, A.; Ulrich, A. Variability in stem fibre content and its association with other characteristics in 1177 flax (*Linum usitatissimum* L.) genebank accessions. *Ind. Crop. Prod.* **2009**, *30*, 33–39. [CrossRef]
5. Goyal, A.; Sharma, V.; Upadhyay, N.; Gill, S.; Sihag, M. Flax and flaxseed oil: An ancient medicine & modern functional food. *J. Food Sci. Technol.* **2014**, *51*, 1633–1653. [CrossRef] [PubMed]
6. Imran, M.; Ahmad, N.; Anjum, F.M.; Khan, M.K.; Mushtaq, Z.; Nadeem, M.; Hussain, S. Potential protective properties of flax lignan secoisolariciresinol diglucoside. *Nutr. J.* **2015**, *14*, 71. [CrossRef]
7. Parikh, M.; Netticadan, T.; Pierce, G.N. Flaxseed: Its bioactive components and their cardiovascular benefits. *Am. J. Physiol. Heart Circ. Physiol.* **2018**, *314*, H146–H159. [CrossRef] [PubMed]
8. Kezimana, P.; Dmitriev, A.A.; Kudryavtseva, A.V.; Romanova, E.V.; Melnikova, N.V. Secoisolariciresinol Diglucoside of Flaxseed and Its Metabolites: Biosynthesis and Potential for Nutraceuticals. *Front. Genet.* **2018**, *9*, 641. [CrossRef] [PubMed]
9. Mali, A.V.; Padhye, S.B.; Anant, S.; Hegde, M.V.; Kadam, S.S. Anticancer and antimetastatic potential of enterolactone: Clinical, preclinical and mechanistic perspectives. *Eur. J. Pharmacol.* **2019**, *852*, 107–124. [CrossRef]
10. Cullis, C.A. *Genetics and Genomics of Linum*; Springer International Publishing: Cham, Switzerland, 2019.
11. Campos, J.R.; Severino, P.; Ferreira, C.S.; Zielinska, A.; Santini, A.; Souto, S.B.; Souto, E.B. Linseed Essential Oil—Source of Lipids as Active Ingredients for Pharmaceuticals and Nutraceuticals. *Curr. Med. Chem.* **2019**, *26*, 4537–4558. [CrossRef]
12. Fombuena, V.; Petrucci, R.; Dominici, F.; Jorda-Vilaplana, A.; Montanes, N.; Torre, L. Maleinized Linseed Oil as Epoxy Resin Hardener for Composites with High Bio Content Obtained from Linen Byproducts. *Polymers* **2019**, *11*, 301. [CrossRef]
13. Corino, C.; Rossi, R.; Cannata, S.; Ratti, S. Effect of dietary linseed on the nutritional value and quality of pork and pork products: Systematic review and meta-analysis. *Meat Sci.* **2014**, *98*, 679–688. [CrossRef]
14. Singh, K.K.; Mridula, D.; Rehal, J.; Barnwal, P. Flaxseed: A Potential Source of Food, Feed and Fiber. *Crit. Rev. Food Sci. Nutr.* **2011**, *51*, 210–222. [CrossRef]
15. Costa, S.M.; Ferreira, D.P.; Ferreira, A.; Vaz, F.; Fangueiro, R. Multifunctional Flax Fibres Based on the Combined Effect of Silver and Zinc Oxide (Ag/ZnO) Nanostructures. *Nanomaterials* **2018**, *8*, 1069. [CrossRef] [PubMed]
16. Baley, C.; Gomina, M.; Breard, J.; Bourmaud, A.; Davies, P. Variability of mechanical properties of flax fibres for composite reinforcement. A review. *Ind. Crop. Prod.* **2019**, *145*, 111984. [CrossRef]
17. Zhu, J.; Zhu, H.; Njuguna, J.; Abhyankar, H. Recent Development of Flax Fibres and Their Reinforced Composites Based on Different Polymeric Matrices. *Materials* **2013**, *6*, 5171–5198. [CrossRef]
18. Goudenhooft, C.; Bourmaud, A.; Baley, C. Flax (*Linum usitatissimum* L.) Fibers for Composite Reinforcement: Exploring the Link Between Plant Growth, Cell Walls Development, and Fiber Properties. *Front. Plant Sci.* **2019**, *10*, 411. [CrossRef] [PubMed]
19. Zhang, H.; Liu, D.; Huang, T.; Hu, Q.; Lammer, H. Three-Dimensional Printing of Continuous Flax Fiber-Reinforced Thermoplastic Composites by Five-Axis Machine. *Materials* **2020**, *13*, 1678. [CrossRef]
20. Mokhothu, T.H.; John, M.J. Review on hygroscopic aging of cellulose fibres and their biocomposites. *Carbohydr. Polym.* **2015**, *131*, 337–354. [CrossRef]

21. Dhakal, H.N.; Sain, M. Enhancement of Mechanical Properties of Flax-Epoxy Composite with Carbon Fibre Hybridisation for Lightweight Applications. *Materials* **2019**, *13*, 109. [CrossRef]
22. Wu, C.M.; Lai, W.Y.; Wang, C.Y. Effects of Surface Modification on the Mechanical Properties of Flax/beta-Polypropylene Composites. *Materials* **2016**, *9*, 314. [CrossRef]
23. Kymäläinen, H.-R.; Sjöberg, A.-M. Flax and hemp fibres as raw materials for thermal insulations. *Build. Environ.* **2008**, *43*, 1261–1269. [CrossRef]
24. Wang, Z.; Hobson, N.; Galindo, L.; Zhu, S.; Shi, D.; McDill, J.; Yang, L.; Hawkins, S.; Neutelings, G.; Datla, R.; et al. The genome of flax (*Linum usitatissimum*) assembled de novo from short shotgun sequence reads. *Plant J. Cell Mol. Biol.* **2012**, *72*, 461–473. [CrossRef]
25. You, F.M.; Xiao, J.; Li, P.; Yao, Z.; Jia, G.; He, L.; Zhu, T.; Luo, M.C.; Wang, X.; Deyholos, M.K.; et al. Chromosome-scale pseudomolecules refined by optical, physical and genetic maps in flax. *Plant J. Cell Mol. Biol.* **2018**, *95*, 371–384. [CrossRef] [PubMed]
26. Sa, R.; Yi, L.; Siqin, B.; An, M.; Bao, H.; Song, X.; Wang, S.; Li, Z.; Zhang, Z.; Hazaisi, H.; et al. Chromosome-Level Genome Assembly and Annotation of the Fiber Flax (*Linum usitatissimum*) Genome. *Front. Genet.* **2021**, *12*, 735690. [CrossRef] [PubMed]
27. Zhang, J.; Qi, Y.; Wang, L.; Wang, L.; Yan, X.; Dang, Z.; Li, W.; Zhao, W.; Pei, X.; Li, X.; et al. Genomic Comparison and Population Diversity Analysis Provide Insights into the Domestication and Improvement of Flax. *iScience* **2020**, *23*, 100967. [CrossRef] [PubMed]
28. Dmitriev, A.A.; Pushkova, E.N.; Novakovskiy, R.O.; Beniaminov, A.D.; Rozhmina, T.A.; Zhuchenko, A.A.; Bolsheva, N.L.; Muravenko, O.V.; Povkhova, L.V.; Dvorianinova, E.M.; et al. Genome Sequencing of Fiber Flax Cultivar Atlant Using Oxford Nanopore and Illumina Platforms. *Front. Genet.* **2020**, *11*, 590282. [CrossRef] [PubMed]
29. Soto-Cerda, B.J.; Duguid, S.; Booker, H.; Rowland, G.; Diederichsen, A.; Cloutier, S. Genomic regions underlying agronomic traits in linseed (*Linum usitatissimum* L.) as revealed by association mapping. *J. Integr. Plant Biol.* **2014**, *56*, 75–87. [CrossRef]
30. Xie, D.; Dai, Z.; Yang, Z.; Sun, J.; Zhao, D.; Yang, X.; Zhang, L.; Tang, Q.; Su, J. Genome-Wide Association Study Identifying Candidate Genes Influencing Important Agronomic Traits of Flax (*Linum usitatissimum* L.) Using SLAF-seq. *Front. Plant Sci.* **2017**, *8*, 2232. [CrossRef] [PubMed]
31. Xie, D.; Dai, Z.; Yang, Z.; Tang, Q.; Sun, J.; Yang, X.; Song, X.; Lu, Y.; Zhao, D.; Zhang, L.; et al. Genomic variations and association study of agronomic traits in flax. *BMC Genom.* **2018**, *19*, 512. [CrossRef] [PubMed]
32. Soto-Cerda, B.J.; Cloutier, S.; Quian, R.; Gajardo, H.A.; Olivos, M.; You, F.M. Genome-Wide Association Analysis of Mucilage and Hull Content in Flax (*Linum usitatissimum* L.) Seeds. *Int. J. Mol. Sci.* **2018**, *19*, 2870. [CrossRef]
33. Zhang, J.; Long, Y.; Wang, L.; Dang, Z.; Zhang, T.; Song, X.; Dang, Z.; Pei, X. Consensus genetic linkage map construction and QTL mapping for plant height-related traits in linseed flax (*Linum usitatissimum* L.). *BMC Plant Biol.* **2018**, *18*, 160. [CrossRef]
34. You, F.M.; Xiao, J.; Li, P.; Yao, Z.; Jia, G.; He, L.; Kumar, S.; Soto-Cerda, B.; Duguid, S.D.; Booker, H.M.; et al. Genome-Wide Association Study and Selection Signatures Detect Genomic Regions Associated with Seed Yield and Oil Quality in Flax. *Int. J. Mol. Sci.* **2018**, *19*, 2303. [CrossRef]
35. Wu, J.; Zhao, Q.; Zhang, L.; Li, S.; Ma, Y.; Pan, L.; Lin, H.; Wu, G.; Yuan, H.; Yu, Y.; et al. QTL Mapping of Fiber-Related Traits Based on a High-Density Genetic Map in Flax (*Linum usitatissimum* L.). *Front. Plant Sci.* **2018**, *9*, 885. [CrossRef] [PubMed]
36. Guo, D.; Jiang, H.; Yan, W.; Yang, L.; Ye, J.; Wang, Y.; Yan, Q.; Chen, J.; Gao, Y.; Duan, L.; et al. Resequencing 200 Flax Cultivated Accessions Identifies Candidate Genes Related to Seed Size and Weight and Reveals Signatures of Artificial Selection. *Front. Plant Sci.* **2019**, *10*, 1682. [CrossRef] [PubMed]
37. Xie, D.; Dai, Z.; Yang, Z.; Tang, Q.; Deng, C.; Xu, Y.; Wang, J.; Chen, J.; Zhao, D.; Zhang, S.; et al. Combined genome-wide association analysis and transcriptome sequencing to identify candidate genes for flax seed fatty acid metabolism. *Plant Sci. Int. J. Exp. Plant Biol.* **2019**, *286*, 98–107. [CrossRef]
38. Galinousky, D.; Mokshina, N.; Padvitski, T.; Ageeva, M.; Bogdan, V.; Kilchevsky, A.; Gorshkova, T. The Toolbox for Fiber Flax Breeding: A Pipeline From Gene Expression to Fiber Quality. *Front. Genet.* **2020**, *11*, 589881. [CrossRef] [PubMed]
39. Long, S.H.; Deng, X.; Wang, Y.F.; Li, X.; Qiao, R.Q.; Qiu, C.S.; Guo, Y.; Hao, D.M.; Jia, W.Q.; Chen, X.B. Analysis of 2297 expressed sequence tags (ESTs) from a cDNA library of flax (*Linum ustitatissimum* L.) bark tissue. *Mol. Biol. Rep.* **2012**, *39*, 6289–6296. [CrossRef]
40. Guo, Y.; Qiu, C.; Long, S.; Chen, P.; Hao, D.; Preisner, M.; Wang, H.; Wang, Y. Digital gene expression profiling of flax (*Linum usitatissimum* L.) stem peel identifies genes enriched in fiber-bearing phloem tissue. *Gene* **2017**, *626*, 32–40. [CrossRef]
41. Gorshkov, O.; Mokshina, N.; Gorshkov, V.; Chemikosova, S.; Gogolev, Y.; Gorshkova, T. Transcriptome portrait of cellulose-enriched flax fibres at advanced stage of specialization. *Plant Mol. Biol.* **2017**, *93*, 431–449. [CrossRef]
42. Mokshina, N.; Gorshkov, O.; Ibragimova, N.; Chernova, T.; Gorshkova, T. Cellulosic fibres of flax recruit both primary and secondary cell wall cellulose synthases during deposition of thick tertiary cell walls and in the course of graviresponse. *Funct. Plant Biol. FPB* **2017**, *44*, 820–831. [CrossRef]
43. Gorshkova, T.; Chernova, T.; Mokshina, N.; Gorshkov, V.; Kozlova, L.; Gorshkov, O. Transcriptome Analysis of Intrusively Growing Flax Fibers Isolated by Laser Microdissection. *Sci. Rep.* **2018**, *8*, 14570. [CrossRef]
44. Gorshkov, O.; Chernova, T.; Mokshina, N.; Gogoleva, N.; Suslov, D.; Tkachenko, A.; Gorshkova, T. Intrusive Growth of Phloem Fibers in Flax Stem: Integrated Analysis of miRNA and mRNA Expression Profiles. *Plants* **2019**, *8*, 47. [CrossRef]

45. Mokshina, N.; Gorshkov, O.; Galinousky, D.; Gorshkova, T. Genes with bast fiber-specific expression in flax plants—Molecular keys for targeted fiber crop improvement. *Ind. Crop. Prod.* **2020**, *152*, 112549. [CrossRef]
46. Galindo-Gonzalez, L.; Deyholos, M.K. RNA-seq Transcriptome Response of Flax (*Linum usitatissimum* L.) to the Pathogenic Fungus Fusarium oxysporum f. sp. lini. *Front. Plant Sci.* **2016**, *7*, 1766. [CrossRef]
47. Yu, Y.; Huang, W.; Chen, H.; Wu, G.; Yuan, H.; Song, X.; Kang, Q.; Zhao, D.; Jiang, W.; Liu, Y.; et al. Identification of differentially expressed genes in flax (*Linum usitatissimum* L.) under saline-alkaline stress by digital gene expression. *Gene* **2014**, *549*, 113–122. [CrossRef] [PubMed]
48. Dmitriev, A.A.; Krasnov, G.S.; Rozhmina, T.A.; Kishlyan, N.V.; Zyablitsin, A.V.; Sadritdinova, A.F.; Snezhkina, A.V.; Fedorova, M.S.; Yurkevich, O.Y.; Muravenko, O.V.; et al. Glutathione S-transferases and UDP-glycosyltransferases Are Involved in Response to Aluminum Stress in Flax. *Front. Plant Sci.* **2016**, *7*, 1920. [CrossRef] [PubMed]
49. Dmitriev, A.A.; Kudryavtseva, A.V.; Krasnov, G.S.; Koroban, N.V.; Speranskaya, A.S.; Krinitsina, A.A.; Belenikin, M.S.; Snezhkina, A.V.; Sadritdinova, A.F.; Kishlyan, N.V.; et al. Gene expression profiling of flax (*Linum usitatissimum* L.) under edaphic stress. *BMC Plant Biol.* **2016**, *16*, 237. [CrossRef] [PubMed]
50. Dash, P.K.; Cao, Y.; Jailani, A.K.; Gupta, P.; Venglat, P.; Xiang, D.; Rai, R.; Sharma, R.; Thirunavukkarasu, N.; Abdin, M.Z.; et al. Genome-wide analysis of drought induced gene expression changes in flax (*Linum usitatissimum*). *GM Crop. Food* **2014**, *5*, 106–119. [CrossRef]
51. Dash, P.K.; Rai, R.; Mahato, A.K.; Gaikwad, K.; Singh, N.K. Transcriptome Landscape at Different Developmental Stages of a Drought Tolerant Cultivar of Flax (*Linum usitatissimum*). *Front. Chem.* **2017**, *5*, 82. [CrossRef]
52. Dmitriev, A.A.; Krasnov, G.S.; Rozhmina, T.A.; Novakovskiy, R.O.; Snezhkina, A.V.; Fedorova, M.S.; Yurkevich, O.Y.; Muravenko, O.V.; Bolsheva, N.L.; Kudryavtseva, A.V.; et al. Differential gene expression in response to Fusarium oxysporum infection in resistant and susceptible genotypes of flax (*Linum usitatissimum* L.). *BMC Plant Biol.* **2017**, *17*, 253. [CrossRef]
53. Preisner, M.; Wojtasik, W.; Kostyn, K.; Boba, A.; Czuj, T.; Szopa, J.; Kulma, A. The cinnamyl alcohol dehydrogenase family in flax: Differentiation during plant growth and under stress conditions. *J. Plant Physiol.* **2018**, *221*, 132–143. [CrossRef]
54. Dmitriev, A.A.; Krasnov, G.S.; Rozhmina, T.A.; Zyablitsin, A.V.; Snezhkina, A.V.; Fedorova, M.S.; Pushkova, E.N.; Kezimana, P.; Novakovskiy, R.O.; Povkhova, L.V.; et al. Flax (*Linum usitatissimum* L.) response to non-optimal soil acidity and zinc deficiency. *BMC Plant Biol.* **2019**, *19*, 54. [CrossRef]
55. Wu, J.; Zhao, Q.; Wu, G.; Yuan, H.; Ma, Y.; Lin, H.; Pan, L.; Li, S.; Sun, D. Comprehensive Analysis of Differentially Expressed Unigenes under NaCl Stress in Flax (*Linum usitatissimum* L.) Using RNA-Seq. *Int. J. Mol. Sci.* **2019**, *20*, 369. [CrossRef]
56. Le Roy, J.; Blervacq, A.S.; Creach, A.; Huss, B.; Hawkins, S.; Neutelings, G. Spatial regulation of monolignol biosynthesis and laccase genes control developmental and stress-related lignin in flax. *BMC Plant Biol.* **2017**, *17*, 124. [CrossRef] [PubMed]
57. Mokshina, N.; Makshakova, O.; Nazipova, A.; Gorshkov, O.; Gorshkova, T. Flax rhamnogalacturonan lyases: Phylogeny, differential expression and modeling of protein structure. *Physiol. Plant* **2019**, *167*, 173–187. [CrossRef] [PubMed]
58. Hobson, N.; Deyholos, M.K. Genomic and expression analysis of the flax (*Linum usitatissimum*) family of glycosyl hydrolase 35 genes. *BMC Genom.* **2013**, *14*, 344. [CrossRef] [PubMed]
59. Morello, L.; Pydiura, N.; Galinousky, D.; Blume, Y.; Breviario, D. Flax tubulin and CesA superfamilies represent attractive and challenging targets for a variety of genome- and base-editing applications. *Funct. Integr. Genom.* **2020**, *20*, 163–176. [CrossRef]
60. Mokshina, N.; Gorshkova, T.; Deyholos, M.K. Chitinase-like (CTL) and cellulose synthase (CESA) gene expression in gelatinous-type cellulosic walls of flax (*Linum usitatissimum* L.) bast fibers. *PLoS ONE* **2014**, *9*, e97949. [CrossRef]
61. Chantreau, M.; Chabbert, B.; Billiard, S.; Hawkins, S.; Neutelings, G. Functional analyses of cellulose synthase genes in flax (*Linum usitatissimum*) by virus-induced gene silencing. *Plant Biotechnol. J.* **2015**, *13*, 1312–1324. [CrossRef]
62. Galinousky, D.; Padvitski, T.; Bayer, G.; Pirko, Y.; Pydiura, N.; Anisimova, N.; Nikitinskaya, T.; Khotyleva, L.; Yemets, A.; Kilchevsky, A.; et al. Expression analysis of cellulose synthase and main cytoskeletal protein genes in flax (*Linum usitatissimum* L.). *Cell Biol. Int.* **2019**, *43*, 1065–1071. [CrossRef]
63. Yurkevich, O.Y.; Kirov, I.V.; Bolsheva, N.L.; Rachinskaya, O.A.; Grushetskaya, Z.E.; Zoschuk, S.A.; Samatadze, T.E.; Bogdanova, M.V.; Lemesh, V.A.; Amosova, A.V.; et al. Integration of Physical, Genetic, and Cytogenetic Mapping Data for Cellulose Synthase (CesA) Genes in Flax (*Linum usitatissimum* L.). *Front. Plant Sci.* **2017**, *8*, 1467. [CrossRef]
64. Pydiura, N.; Pirko, Y.; Galinousky, D.; Postovoitova, A.; Yemets, A.; Kilchevsky, A.; Blume, Y. Genome-wide identification, phylogenetic classification, and exon-intron structure characterization of the tubulin and actin genes in flax (*Linum usitatissimum*). *Cell Biol. Int.* **2019**, *43*, 1010–1019. [CrossRef]
65. Dalisay, D.S.; Kim, K.W.; Lee, C.; Yang, H.; Rubel, O.; Bowen, B.P.; Davin, L.B.; Lewis, N.G. Dirigent Protein-Mediated Lignan and Cyanogenic Glucoside Formation in Flax Seed: Integrated Omics and MALDI Mass Spectrometry Imaging. *J. Nat. Prod.* **2015**, *78*, 1231–1242. [CrossRef] [PubMed]
66. Ghose, K.; Selvaraj, K.; McCallum, J.; Kirby, C.W.; Sweeney-Nixon, M.; Cloutier, S.J.; Deyholos, M.; Datla, R.; Fofana, B. Identification and functional characterization of a flax UDP-glycosyltransferase glucosylating secoisolariciresinol (SECO) into secoisolariciresinol monoglucoside (SMG) and diglucoside (SDG). *BMC Plant Biol.* **2014**, *14*, 82. [CrossRef] [PubMed]
67. Hemmati, S.; von Heimendahl, C.B.; Klaes, M.; Alfermann, A.W.; Schmidt, T.J.; Fuss, E. Pinoresinol-lariciresinol reductases with opposite enantiospecificity determine the enantiomeric composition of lignans in the different organs of *Linum usitatissimum* L. *Planta Med.* **2010**, *76*, 928–934. [CrossRef] [PubMed]

68. Corbin, C.; Drouet, S.; Markulin, L.; Auguin, D.; Laine, E.; Davin, L.B.; Cort, J.R.; Lewis, N.G.; Hano, C. A genome-wide analysis of the flax (*Linum usitatissimum* L.) dirigent protein family: From gene identification and evolution to differential regulation. *Plant Mol. Biol.* **2018**, *97*, 73–101. [CrossRef] [PubMed]
69. Fofana, B.; Duguid, S.; Cloutier, S. Cloning of fatty acid biosynthetic genes β-ketoacyl CoA synthase, fatty acid elongase, stearoyl-ACP desaturase, and fatty acid desaturase and analysis of expression in the early developmental stages of flax (*Linum usitatissimum* L.) seeds. *Plant Sci.* **2004**, *166*, 1487–1496. [CrossRef]
70. Vrinten, P.; Hu, Z.; Munchinsky, M.A.; Rowland, G.; Qiu, X. Two FAD3 desaturase genes control the level of linolenic acid in flax seed. *Plant Physiol.* **2005**, *139*, 79–87. [CrossRef] [PubMed]
71. Thambugala, D.; Duguid, S.; Loewen, E.; Rowland, G.; Booker, H.; You, F.M.; Cloutier, S. Genetic variation of six desaturase genes in flax and their impact on fatty acid composition. *Theor. Appl. Genet.* **2013**, *126*, 2627–2641. [CrossRef]
72. Fofana, B.; Cloutier, S.; Duguid, S.; Ching, J.; Rampitsch, C. Gene expression of stearoyl-ACP desaturase and delta12 fatty acid desaturase 2 is modulated during seed development of flax (*Linum usitatissimum*). *Lipids* **2006**, *41*, 705–712. [CrossRef]
73. Khan, N.; You, F.M.; Datla, R.; Ravichandran, S.; Jia, B.; Cloutier, S. Genome-wide identification of ATP binding cassette (ABC) transporter and heavy metal associated (HMA) gene families in flax (*Linum usitatissimum* L.). *BMC Genom.* **2020**, *21*, 722. [CrossRef]
74. Dmitriev, A.A.; Novakovskiy, R.O.; Pushkova, E.N.; Rozhmina, T.A.; Zhuchenko, A.A.; Bolsheva, N.L.; Beniaminov, A.D.; Mitkevich, V.A.; Povkhova, L.V.; Dvorianinova, E.M.; et al. Transcriptomes of Different Tissues of Flax (*Linum usitatissimum* L.) Cultivars With Diverse Characteristics. *Front. Genet.* **2020**, *11*, 565146. [CrossRef]
75. Boerjan, W.; Ralph, J.; Baucher, M. Lignin biosynthesis. *Annu. Rev. Plant Biol.* **2003**, *54*, 519–546. [CrossRef] [PubMed]
76. Vanholme, R.; Morreel, K.; Ralph, J.; Boerjan, W. Lignin engineering. *Curr. Opin. Plant Biol.* **2008**, *11*, 278–285. [CrossRef] [PubMed]
77. Rohde, A.; Morreel, K.; Ralph, J.; Goeminne, G.; Hostyn, V.; De Rycke, R.; Kushnir, S.; Van Doorsselaere, J.; Joseleau, J.P.; Vuylsteke, M.; et al. Molecular phenotyping of the pal1 and pal2 mutants of Arabidopsis thaliana reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. *Plant Cell* **2004**, *16*, 2749–2771. [CrossRef]
78. Chen, F.; Srinivasa Reddy, M.S.; Temple, S.; Jackson, L.; Shadle, G.; Dixon, R.A. Multi-site genetic modulation of monolignol biosynthesis suggests new routes for formation of syringyl lignin and wall-bound ferulic acid in alfalfa (*Medicago sativa* L.). *Plant J. Cell Mol. Biol.* **2006**, *48*, 113–124. [CrossRef] [PubMed]
79. Wang, J.P.; Matthews, M.L.; Williams, C.M.; Shi, R.; Yang, C.; Tunlaya-Anukit, S.; Chen, H.C.; Li, Q.; Liu, J.; Lin, C.Y.; et al. Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. *Nat. Commun.* **2018**, *9*, 1579. [CrossRef]
80. Bate, N.J.; Orr, J.; Ni, W.; Meromi, A.; Nadler-Hassar, T.; Doerner, P.W.; Dixon, R.A.; Lamb, C.J.; Elkind, Y. Quantitative relationship between phenylalanine ammonia-lyase levels and phenylpropanoid accumulation in transgenic tobacco identifies a rate-determining step in natural product synthesis. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 7608–7612. [CrossRef]
81. Sewalt, V.; Ni, W.; Blount, J.W.; Jung, H.G.; Masoud, S.A.; Howles, P.A.; Lamb, C.; Dixon, R.A. Reduced Lignin Content and Altered Lignin Composition in Transgenic Tobacco Down-Regulated in Expression of L-Phenylalanine Ammonia-Lyase or Cinnamate 4-Hydroxylase. *Plant Physiol.* **1997**, *115*, 41–50. [CrossRef]
82. Cao, S.; Huang, C.; Luo, L.; Zheng, S.; Zhong, Y.; Sun, J.; Gui, J.; Li, L. Cell-Specific Suppression of 4-Coumarate-CoA Ligase Gene Reveals Differential Effect of Lignin on Cell Physiological Function in Populus. *Front. Plant Sci.* **2020**, *11*, 589729. [CrossRef]
83. Ozparpucu, M.; Gierlinger, N.; Burgert, I.; Van Acker, R.; Vanholme, R.; Boerjan, W.; Pilate, G.; Dejardin, A.; Ruggeberg, M. The effect of altered lignin composition on mechanical properties of CINNAMYL ALCOHOL DEHYDROGENASE (CAD) deficient poplars. *Planta* **2018**, *247*, 887–897. [CrossRef]
84. Zhang, S.; Jia, T.; Zhang, Z.; Zou, X.; Fan, S.; Lei, K.; Jiang, X.; Niu, D.; Yuan, Y.; Shang, H. Insight into the relationship between S-lignin and fiber quality based on multiple research methods. *Plant Physiol. Biochem. PPB* **2020**, *147*, 251–261. [CrossRef]
85. Chantreau, M.; Grec, S.; Gutierrez, L.; Dalmais, M.; Pineau, C.; Demailly, H.; Paysant-Leroux, C.; Tavernier, R.; Trouve, J.P.; Chatterjee, M.; et al. PT-Flax (phenotyping and TILLinG of flax): Development of a flax (*Linum usitatissimum* L.) mutant population and TILLinG platform for forward and reverse genetics. *BMC Plant Biol.* **2013**, *13*, 159. [CrossRef]
86. Liu, W.; Jiang, Y.; Wang, C.; Zhao, L.; Jin, Y.; Xing, Q.; Li, M.; Lv, T.; Qi, H. Lignin synthesized by CmCAD2 and CmCAD3 in oriental melon (Cucumis melo L.) seedlings contributes to drought tolerance. *Plant Mol. Biol.* **2020**, *103*, 689–704. [CrossRef] [PubMed]
87. Liu, W.; Jiang, Y.; Jin, Y.; Wang, C.; Yang, J.; Qi, H. Drought-induced ABA, H2O2 and JA positively regulate CmCAD genes and lignin synthesis in melon stems. *BMC Plant Biol.* **2021**, *21*, 83. [CrossRef] [PubMed]
88. Novakovskiy, R.O.; Povkhova, L.V.; Krasnov, G.S.; Rozhmina, T.A.; Zhuchenko, A.A.; Kudryavtseva, L.P.; Pushkova, E.N.; Kezimana, P.; Kudryavtseva, A.V.; Dmitriev, A.A.; et al. The cinnamyl alcohol dehydrogenase gene family is involved in the response to Fusarium oxysporum in resistant and susceptible flax genotypes. *Vavilov J. Genet. Breed.* **2019**, *23*, 896–901. [CrossRef]
89. Bagniewska-Zadworna, A.; Barakat, A.; Lakomy, P.; Smolinski, D.J.; Zadworny, M. Lignin and lignans in plant defence: Insight from expression profiling of cinnamyl alcohol dehydrogenase genes during development and following fungal infection in Populus. *Plant Sci. Int. J. Exp. Plant Biol.* **2014**, *229*, 111–121. [CrossRef] [PubMed]
90. Xie, M.; Zhang, J.; Tschaplinski, T.J.; Tuskan, G.A.; Chen, J.G.; Muchero, W. Regulation of Lignin Biosynthesis and Its Role in Growth-Defense Tradeoffs. *Front. Plant Sci.* **2018**, *9*, 1427. [CrossRef] [PubMed]

91. Robert, S.; Mouille, G.; Höfte, H. The mechanism and regulation of cellulose synthesis in primary walls: Lessons from cellulose-deficient Arabidopsis mutants. *Cellulose* **2004**, *11*, 351–364. [CrossRef]
92. Grover, A. Plant Chitinases: Genetic Diversity and Physiological Roles. *Crit. Rev. Plant Sci.* **2012**, *31*, 57–73. [CrossRef]
93. Kombrink, E.; Schroder, M.; Hahlbrock, K. Several "pathogenesis-related" proteins in potato are 1,3-beta-glucanases and chitinases. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 782–786. [CrossRef]
94. Hermans, C.; Porco, S.; Verbruggen, N.; Bush, D.R. Chitinase-like protein CTL1 plays a role in altering root system architecture in response to multiple environmental conditions. *Plant Physiol.* **2010**, *152*, 904–917. [CrossRef]
95. Levy, I.; Shani, Z.; Shoseyov, O. Modification of polysaccharides and plant cell wall by endo-1,4-beta-glucanase and cellulose-binding domains. *Biomol. Eng.* **2002**, *19*, 17–30. [CrossRef]
96. Chandrasekar, B.; van der Hoorn, R.A. Beta galactosidases in Arabidopsis and tomato—A mini review. *Biochem. Soc. Trans.* **2016**, *44*, 150–158. [CrossRef] [PubMed]
97. Roach, M.J.; Mokshina, N.Y.; Badhan, A.; Snegireva, A.V.; Hobson, N.; Deyholos, M.K.; Gorshkova, T.A. Development of cellulosic secondary walls in flax fibers requires beta-galactosidase. *Plant Physiol.* **2011**, *156*, 1351–1363. [CrossRef] [PubMed]
98. Silva, I.R.; Jers, C.; Meyer, A.S.; Mikkelsen, J.D. Rhamnogalacturonan I modifying enzymes: An update. *Nat. Biotechnol.* **2016**, *33*, 41–54. [CrossRef]
99. Hano, C.; Martin, I.; Fliniaux, O.; Legrand, B.; Gutierrez, L.; Arroo, R.R.; Mesnard, F.; Lamblin, F.; Laine, E. Pinoresinol-lariciresinol reductase gene expression and secoisolariciresinol diglucoside accumulation in developing flax (*Linum usitatissimum*) seeds. *Planta* **2006**, *224*, 1291–1301. [CrossRef]
100. Corbin, C.; Drouet, S.; Mateljak, I.; Markulin, L.; Decourtil, C.; Renouard, S.; Lopez, T.; Doussot, J.; Lamblin, F.; Auguin, D.; et al. Functional characterization of the pinoresinol-lariciresinol reductase-2 gene reveals its roles in yatein biosynthesis and flax defense response. *Planta* **2017**, *246*, 405–420. [CrossRef]
101. Renouard, S.; Tribalatc, M.A.; Lamblin, F.; Mongelard, G.; Fliniaux, O.; Corbin, C.; Marosevic, D.; Pilard, S.; Demailly, H.; Gutierrez, L.; et al. RNAi-mediated pinoresinol lariciresinol reductase gene silencing in flax (*Linum usitatissimum* L.) seed coat: Consequences on lignans and neolignans accumulation. *J. Plant Physiol.* **2014**, *171*, 1372–1377. [CrossRef]
102. Hamade, K.; Fliniaux, O.; Fontaine, J.X.; Molinie, R.; Otogo Nnang, E.; Bassard, S.; Guenin, S.; Gutierrez, L.; Laine, E.; Hano, C.; et al. NMR and LC-MS-Based Metabolomics to Study Osmotic Stress in Lignan-Deficient Flax. *Molecules* **2021**, *26*, 767. [CrossRef]
103. Rajwade, A.V.; Kadoo, N.Y.; Borikar, S.P.; Harsulkar, A.M.; Ghorpade, P.B.; Gupta, V.S. Differential transcriptional activity of SAD, FAD2 and FAD3 desaturase genes in developing seeds of linseed contributes to varietal variation in alpha-linolenic acid content. *Phytochemistry* **2014**, *98*, 41–53. [CrossRef]
104. Zhang, Y.; Maximova, S.N.; Guiltinan, M.J. Characterization of a stearoyl-acyl carrier protein desaturase gene family from chocolate tree, *Theobroma cacao* L. *Front. Plant Sci.* **2015**, *6*, 239. [CrossRef]
105. Dmitriev, A.A.; Kezimana, P.; Rozhmina, T.A.; Zhuchenko, A.A.; Povkhova, L.V.; Pushkova, E.N.; Novakovskiy, R.O.; Pavelek, M.; Vladimirov, G.N.; Nikolaev, E.N.; et al. Genetic diversity of SAD and FAD genes responsible for the fatty acid composition in flax cultivars and lines. *BMC Plant Biol.* **2020**, *20*, 301. [CrossRef]
106. Kezimana, P.; Rozhmina, T.A.; Krasnov, G.S.; Povkhova, L.V.; Novakovskiy, R.O.; Pushkova, E.N.; Zhuchenko, A.A.; Bjelková, M.; Pavelek, M.; Dmitriev, A.A.; et al. Evaluation of polymorphism of SAD and FAD genes in flax (*Linum usitatissimum* L.) cultivars and lines using deep sequencing. In Proceedings of the Theory and Practice of Adaptive Plant Breeding (Zhuchenkov's Readings VI), Krasnodar, Russia, 25 September 2020; pp. 49–52.
107. Kang, J.; Park, J.; Choi, H.; Burla, B.; Kretzschmar, T.; Lee, Y.; Martinoia, E. Plant ABC Transporters. *Arab. Book* **2011**, *9*, e0153. [CrossRef]
108. D'yakov, A.B. *Flax Physiology and Ecology*; LAP LAMBERT Academic Publishing: Sunnyvale, CA, USA, 2006.
109. Kumar, S.; You, F.M.; Duguid, S.; Booker, H.; Rowland, G.; Cloutier, S. QTL for fatty acid composition and yield in linseed (*Linum usitatissimum* L.). *Theor. Appl. Genet.* **2015**, *128*, 965–984. [CrossRef] [PubMed]
110. Cloutier, S.; Ragupathy, R.; Niu, Z.; Duguid, S. SSR-based linkage map of flax (*Linum usitatissimum* L.) and mapping of QTLs underlying fatty acid composition traits. *Mol. Breed.* **2011**, *28*, 437–451. [CrossRef]
111. Soto-Cerda, B.J.; Duguid, S.; Booker, H.; Rowland, G.; Diederichsen, A.; Cloutier, S. Association mapping of seed quality traits using the Canadian flax (*Linum usitatissimum* L.) core collection. *Theor. Appl. Genet.* **2014**, *127*, 881–896. [CrossRef]
112. Asgarinia, P.; Cloutier, S.; Duguid, S.; Rashid, K.; Mirlohi, A.; Banik, M.; Saeidi, G. Mapping Quantitative Trait Loci for Powdery Mildew Resistance in Flax (*Linum usitatissimum* L.). *Crop Sci.* **2013**, *53*, 2462–2472. [CrossRef]
113. He, L.; Xiao, J.; Rashid, K.Y.; Yao, Z.; Li, P.; Jia, G.; Wang, X.; Cloutier, S.; You, F.M. Genome-Wide Association Studies for Pasmo Resistance in Flax (*Linum usitatissimum* L.). *Front. Plant Sci.* **2018**, *9*, 1982. [CrossRef]
114. Lan, S.; Zheng, C.; Hauck, K.; McCausland, M.; Duguid, S.D.; Booker, H.M.; Cloutier, S.; You, F.M. Genomic Prediction Accuracy of Seven Breeding Selection Traits Improved by QTL Identification in Flax. *Int. J. Mol. Sci.* **2020**, *21*, 1577. [CrossRef]
115. Sertse, D.; You, F.M.; Ravichandran, S.; Soto-Cerda, B.J.; Duguid, S.; Cloutier, S. Loci harboring genes with important role in drought and related abiotic stress responses in flax revealed by multiple GWAS models. *Theor. Appl. Genet.* **2021**, *134*, 191–212. [CrossRef] [PubMed]
116. Soto-Cerda, B.J.; Aravena, G.; Cloutier, S. Genetic dissection of flowering time in flax (*Linum usitatissimum* L.) through single- and multi-locus genome-wide association studies. *Mol. Genet. Genom. MGG* **2021**, *296*, 877–891. [CrossRef]

117. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef] [PubMed]
118. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997v2.
119. Garrison, E.; Marth, G. Haplotype-based variant detection from short-read sequencing. *arXiv* **2012**, arXiv:1207.3907v2.
120. You, F.M.; Cloutier, S. Mapping Quantitative Trait Loci onto Chromosome-Scale Pseudomolecules in Flax. *Methods Protoc.* **2020**, *3*, 28. [CrossRef]
121. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21. [CrossRef] [PubMed]
122. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [CrossRef]
123. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [CrossRef]