



Article

Genomic Regions Associated with Fusarium Wilt Resistance in Flax

Alexander Kanapin ¹, Mikhail Bankin ², Tatyana Rozhmina ³, Anastasia Samsonova ¹ and Maria Samsonova ^{2,*}

¹ Centre for Computational Biology, Peter the Great St. Petersburg Polytechnic University, 195251 St. Petersburg, Russia; a.kanapin@gmail.com (A.K.); a.a.samsonova@gmail.com (A.S.)

² Mathematical Biology & Bioinformatics Laboratory, Peter the Great St. Petersburg Polytechnic University, 195251 St. Petersburg, Russia; mikle.p.bankin@gmail.com

³ Laboratory of Breeding Technologies, Federal Research Center for Bast Fiber Crops, 172002 Torzhok, Russia; tatyana_rozhmina@mail.ru

* Correspondence: m.g.samsonova@gmail.com; Tel.: +7-812-290-9645

Abstract: Modern flax cultivars are susceptible to many diseases; arguably, the most economically damaging of these is the Fusarium wilt fungal disease. Over the past decades international flax breeding initiatives resulted in the development of resistant cultivars. However, much remains to be learned about the mechanisms of resistance to Fusarium infection in flax. As a first step to uncover the genetic factors associated with resistance to Fusarium wilt disease, we performed a genome-wide association study (GWAS) using 297 accessions from the collection of the Federal Research Centre of the Bast Fiber Crops, Torzhok, Russia. These genotypes were infected with a highly pathogenic *Fusarium oxysporum* f.sp. *lini* MI39 strain; the wilt symptoms were documented in the course of three successive years. Six different single-locus models implemented in GAPIT3 R package were applied to a selected subset of 72,526 SNPs. A total of 15 QTNs (Quantitative Trait Nucleotides) were detected during at least two years of observation, while eight QTNs were found during all three years of the experiment. Of these, ten QTNs occupied a region of 640 Kb at the start of chromosome 1, while the remaining QTNs mapped to chromosomes 8, 11 and 13. All stable QTNs demonstrate a statistically significant allelic effect across 3 years of the experiment. Importantly, several QTNs spanned regions that harbored genes involved in the pathogen recognition and plant immunity response, including the KIP1-like protein (*Lus10025717*) and NBS-LRR protein (*Lus10025852*). Our results provide novel insights into the genetic architecture of flax resistance to Fusarium wilt and pinpoint potential candidate genes for further in-depth studies.

Keywords: fusarium wilt; resistance; flax; GWAS; immune response; disease severity index



Citation: Kanapin, A.; Bankin, M.; Rozhmina, T.; Samsonova, A.; Samsonova, M. Genomic Regions Associated with Fusarium Wilt Resistance in Flax. *Int. J. Mol. Sci.* **2021**, *22*, 12383. <https://doi.org/10.3390/ijms222212383>

Academic Editors: Yuriy L. Orlov, Ming Chen and Oxana B. Dobrovolskaya

Received: 19 October 2021

Accepted: 15 November 2021

Published: 17 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Flax (*Linum usitatissimum* L.) is a valuable crop cultivated for oil and fiber. One of the major flax pathogens affecting the world's crop production is the fungus of the genus *Fusarium*, *Fusarium oxysporum* f.sp. *lini*. The primary fungal infection occurs through the roots. Next, the pathogen colonizes the xylem and blocks the flow of water and nutrients causing the yellowing and wilting of the leaves, vascular tissue damage and, ultimately, plant death [1]. As a result of the disease outbreak, 80–100% of flax harvest could be lost. Moreover, the fungus chlamydospores can survive for up to 50 years in the infected soil and are extremely difficult to eliminate [2].

Flax wilt management is achieved through various agricultural practices, pesticides [3,4] and breeding efforts that aim to cultivate resistant crop varieties [5–8]. Pesticides used in excess and irresponsibly are hazardous for human health [9]. Furthermore, they reduce biodiversity; therefore, they are detrimental to the structure and function of an ecosystem [9,10].

In view of the above, the best approach for effective disease control is a cultivation of resistant genotypes in rotation with other crops. The majority of modern flax varieties show

high or moderate resistance towards Fusarium wilt [7,8]. However, there is a significant risk of disease development due to variations of pathotypes in fungal populations and a substantial genetic erosion of varietal material. This situation is further aggravated by global warming. Climate change and other natural, human-induced stressors, may lead to a rise in aggressiveness of individual pathogen races, as well as to the loss of resistance by the varieties determined by one or two genes. Therefore, it is of utmost importance, to take proactive steps, and to create varieties with different effective pathogen-resistant genes, as well as with optimal combinations of these genes to ensure long-term and efficient crop protection.

The mechanisms of flax resistance to Fusarium wilt have never been fully understood, although resistance to the disease was developed by selection and recombination [7,8]. The cornerstone of a plant's immune system is represented by two classes of receptors that recognize pathogens and protect plant tissues against fungal invasions [11]. At the initial stage of the plant–pathogen interaction, pattern recognition receptors (PRRs) play an important role in pathogen perception and early immune response. All well-characterized PRRs are evolutionary, conserved, membrane-localized, receptor-like kinases. They trigger the PAMP, i.e., pattern-triggered immunity (PTI) by recognizing pathogen-/microbial-associated molecular patterns. PTI plays important role in basal resistance that generally develops after the successful infection of an adapted pathogen [12]. The second class of receptors involved in the establishment of the next level of defense, i.e., the effector-triggered immunity (ETI) is activated when a plant R-gene protein recognizes an effector encoded by the avirulence (*Avr*) gene of the pathogen [13,14]. R genes belong to five classes formed by genes encoding nucleotide-binding domain, leucine-rich, repeat-containing (NBS-LRR) proteins, receptor-like kinase genes, genes producing receptor-like transmembrane proteins or serine threonine kinases, and atypical R genes [15]. To date, over 300 R genes have been characterized and cloned. Of these, more than 60% contain NBS and LRR domains [16,17]. Both PTI- and ETI-resistant regulatory cascades overlap in downstream signaling pathways, which include MAP kinase (MAPK) cascades, calcium fluxes, reactive oxygen species (ROS) production and alterations in hormone networks [18]. Pathogen-related proteins, components of the cell wall, transcription factors, secondary metabolites and antioxidants all play essential roles in the response of flax to the *F. oxysporum* f.sp. *lini* infection, as demonstrated by several transcriptomics experiments [19–21].

Identification of resistance genes and quantitative trait loci are key elements of a successful breeding program. Due to its increased economic importance in the past couple of decades flax has become a centre spot of attention and an object of intensive research in the field of genomics. Recently numerous GWAS studies carried out in flax which interrogated traits related to yield, phenology and fatty acid content were reported [22–29]. In addition, flax omics data analytics broadened our knowledge of wilt resistance genes obtained with classical genetics approaches by adding data on QTL loci to a known gene set (~10 genes) conferring wilt resistance in flax [7,8,30]. Specifically, Spielmeier et al. [31,32] identified two QTLs explaining 38% and 26% of phenotypic variations, located within linkage groups 6 and 10. These loci were discovered in a population of double haploid lines developed from crosses between flax wilt-resistant and susceptible parents. Overall, this plethora of information could be further used in follow-up experiments to decipher resistance mechanisms [33] or in breeding programs for the development of superior cultivars [34].

In this study we aim to exploit flax natural variations to uncover genomic regions controlling resistance to Fusarium wilt. We performed the whole-genome sequencing of 297 flax accessions at 10× depth coverage to characterize genetic diversity and population structure specific to the most Russian varieties [35]. We evaluated the accessions for disease resistance under controlled conditions and applied GWAS algorithms which relied on different statistical models to reveal SNP trait associations. To our knowledge, this is the first time that the GWAS approach was applied to interrogate Fusarium wilt resistance in flax.

2. Results

2.1. Evaluation Resistance to Fusarium Wilt in Flax

Flax resistance to highly virulent *F. oxysporum* MI39 was evaluated under greenhouse conditions by calculating the Disease Severity Index (DSI) (Table S1). The DSI is a normalized proportion of genotypes with identical disease symptoms (see Section 4). The accessions were classified as resistant (42%), weakly susceptible (7%), moderately susceptible (17%) and susceptible (34%) using average DSI values, across 3 years. DSI values are strongly correlated across all three years of investigation (Figure 1), thus attesting to the consistency of the phenotypic dataset. The DSI values obtained for flax cultivars and breeding lines significantly differ from landraces (p -values for 2019–2021 years of experiment were 7.1×10^{-6} , 1.96×10^{-7} and 0.52×10^{-3} , correspondingly) which may be indicative of breeding efforts aimed at the development of wilt-resistant genotypes (Figure 1).

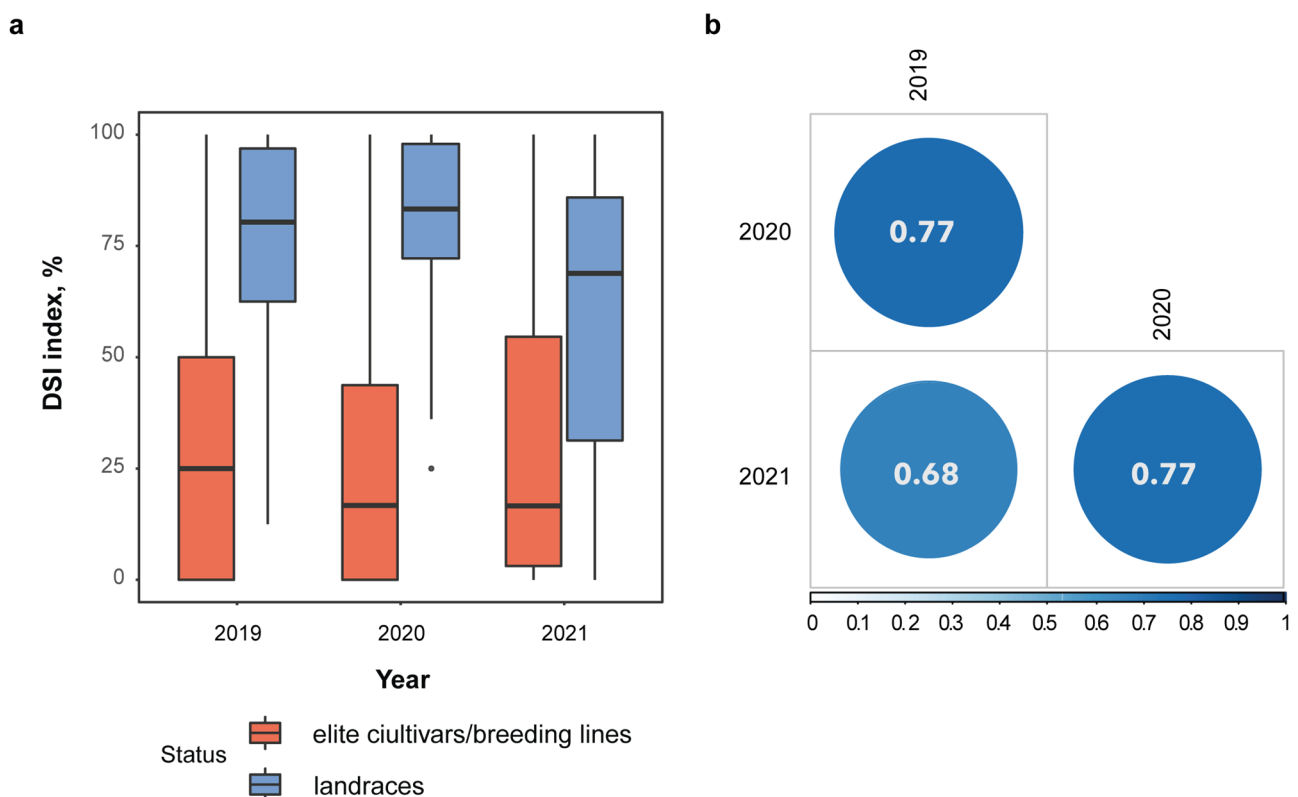


Figure 1. Fusarium wilt resistance in the dataset. (a) DSI index values in elite cultivars/breeding lines and landraces. Black dots represent outlier values, (b) The between year correlation of DSI index.

2.2. Association Mapping of Fusarium Wilt Resistance

A total of 72,526 variants retained after SNP calling and filtering (thresholds for MAF (Minor Allele Frequency) and call rate were 0.05 and 0.85, respectively) were widely distributed over 15 flax chromosomes (Figure S1a). Patterns of population differentiation were analyzed using a principal components analysis, which did not produce a clear separation of fiber and linseed genotypes (see bi-plot in Figure S1b,c).

The precision of the genetic association test depends on the scale of LD (Linkage Disequilibrium), which implicitly defines the size of marker set required for analysis. With a genome size of about 370 Mb and the mean LD decay of 8.6 Kb [35], the number of markers in our dataset (i.e., 72,526 loci) was twice the minimal number of SNPs necessary for an accurate genetic test.

Genome-wide association analyses were performed using the following models: GLM, MLM, CMLM, FarmCPU, SUPER, Blink as implemented in GAPIT3 R package [36]. Using both the GAPIT estimation of Bayesian information content and scree plot analysis, (Figure S1b) we retained a total of five principal components to be used in GWAS.

The six aforementioned methods yielded 38 QTNs identified throughout the experiment (FDR adjusted p -value 0.05, see Table S2). All models performed well in controlling population and family structures, as confirmed by respective Q–Q plots where the observed p -values deviated from the expected values at the end of distribution. Most of the QTNs with the exception of Chr8:22560236, Chr11:6013057 and Chr13:4884610 were discovered by several models. Fifteen QTNs were detected in at least two years (Table 1), while eight QTNs were found in all years (Figure 2). Importantly, ten of these stable QTNs occupy a region of 640 Kb in length at the beginning of chromosome 1 and explain about 7–10% of the phenotypic variation for resistance to Fusarium wilt. The rest of QTNs are located on chromosomes 8, 11 and 13. The majority of the identified QTNs demonstrate a significant allelic effect across all 3 years (Figures 3, S2 and S3).

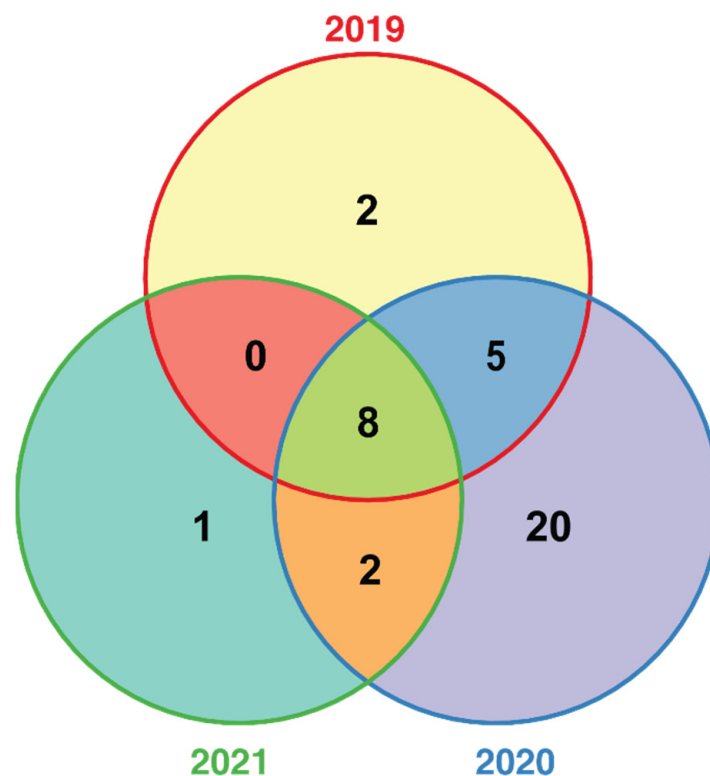


Figure 2. Venn diagram representing the number of unique and shared QTNs associated with DSI index in between-year comparisons.

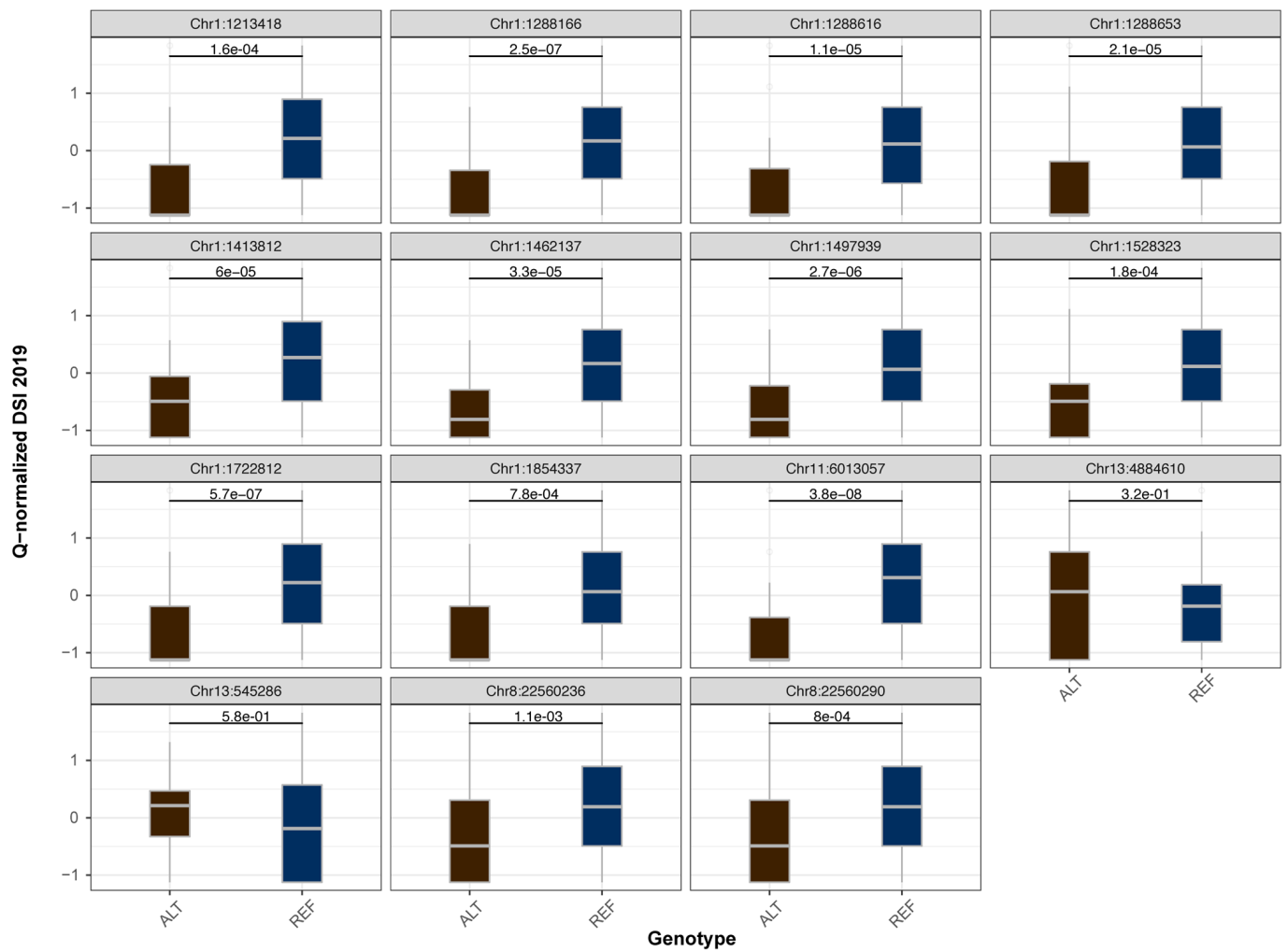


Figure 3. Box plots of the allelic effects observed in stable QTNs in 2019 year. Numbers are p -values according to the Mann–Whitney–Wilcoxon non-parametric test.

Negative alleles decrease the DSI index, while positive alleles increase it. The DSI index value progressively decreases with the increase in the negative-effect allele number in accession genotypes (see Figure 4), thus reflecting a stronger resistance to the pathogen.

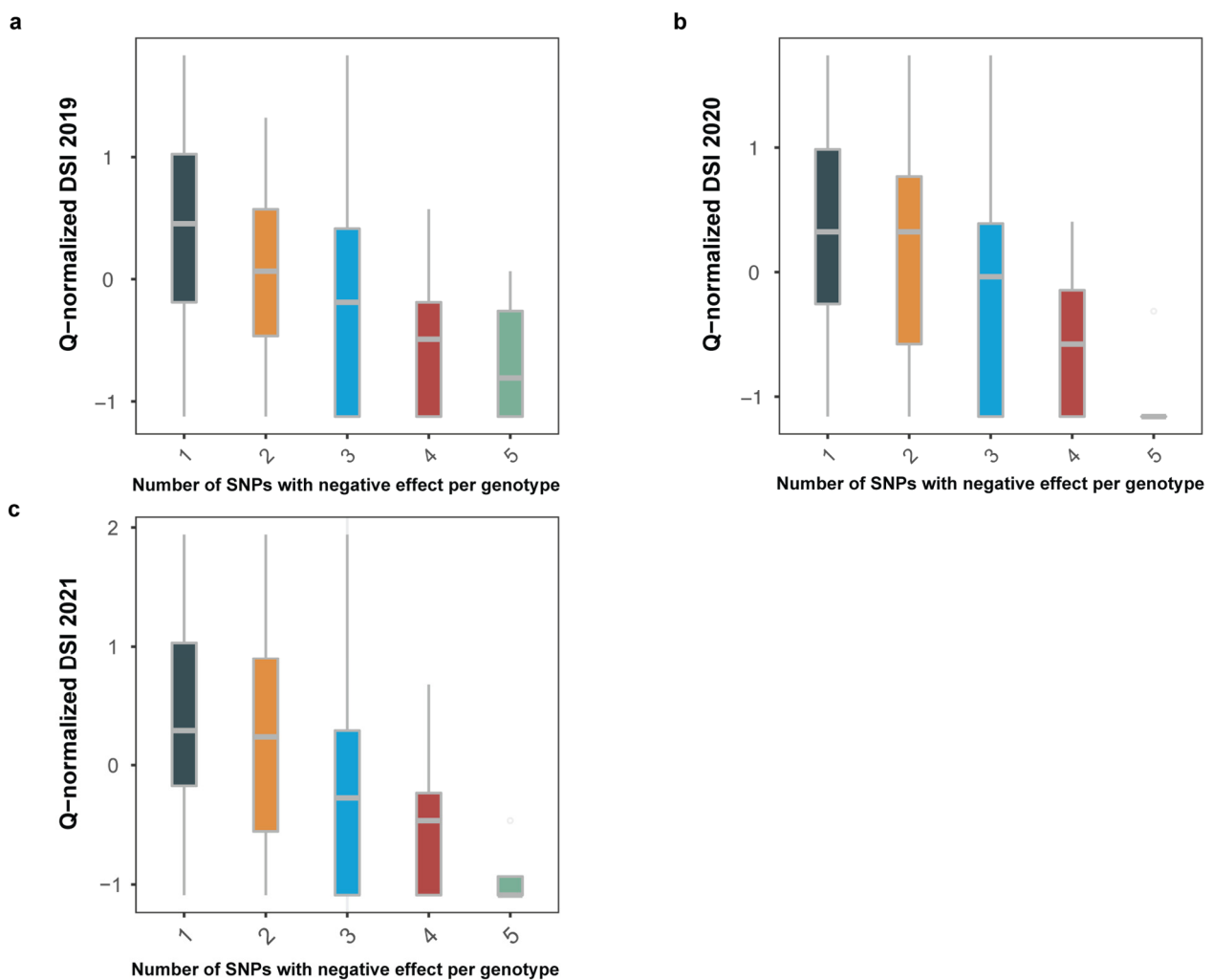


Figure 4. The DSI index value in accessions with different number of QTNs with negative-effect alleles across years of experiment, namely (a) year 2019, (b) year 2020, and (c) year 2021.

2.3. *Fusarium* Wilt Resistance Candidate Genes

Candidate genes involved in various processes associated with pathogen response were identified within predefined regions with length adapted to the extent of LD decay for a relevant chromosome (Figure S4) and centered on stable QTNs (Tables 1 and 2). Many QTNs detected in this study are orthologous to well-characterized *Arabidopsis* genes. For instance, Chr1:1288653 QTN is a missense variant of *Lus10025717* gene orthologous to *A. thaliana*'s *AT2G22560* gene, which encodes the KIP1-like protein harboring the actin binding domain. Likewise, QTN Chr1:1462137 is an upstream variant of *Lus10025756* gene, whose *Arabidopsis* ortholog encodes cytochrome P450 monooxygenase from the subfamily CYP709B. Additionally, a missense variant Chr1:1528323 QTN is within the *Lus10025773* gene, which is an ortholog of the *AT1G53050* gene for protein kinase superfamily protein. Chr1:1722812 QTN is an intron variant within *Lus10025823* gene orthologous to the *AT1G79750* gene encoding NADP-malic enzyme 4. Finally, Chr1:1854337 QTN is an upstream transcript variant of *Lus10025853* gene, where the *A. thaliana* ortholog encodes exportin 1A. About 2.5 kb from this QTN is *Lus10025852* gene, where an ortholog in *Arabidopsis* (*AT5G17890*) encodes the DA1-related protein.

Table 1. List of stable QTNs identified by several models and in at least two years.

Chromosome ID	Number	QTN Position	REF/ALT	Average PVE	Average Effect	Year	MAF	QTN Annotation
CP027619.1	1	1213418	C/T	0.106	−0.53	2019 2020 2021	0.138	upstream_transcript_variant
CP027619.1	1	1288166	A/G	0.09	−0.74	2019 2020 2021	0.093	synonymous_variant
CP027619.1	1	1288616	C/A	0.10	−0.72	2019 2020 2021	0.112	synonymous_variant
CP027619.1	1	1288653	T/C	0.103	−0.76	2019 2020 2021	0.103	missense_variant
CP027619.1	1	1413812	T/A	0.102	−0.69	2019 2020 2021	0.131	intergenic_variant
CP027619.1	1	1462137	A/G	0.082	−0.66	2019 2020	0.099	upstream_transcript_variant
CP027619.1	1	1497939	G/C	0.075	−0.62	2019 2020	0.103	intergenic_variant
CP027619.1	1	1528323	C/G	0.078	−0.64	2019 2020	0.101	missense_variant
CP027619.1	1	1722812	C/G	0.105	−0.68	2019 2020 2021	0.131	intron_variant
CP027619.1	1	1854337	T/C	0.082	0.56	2019 2020	0.101	upstream_transcript_variant
CP027632.1	8	22560236	G/A	0.064	−0.53	2019 2020	0.185	upstream_transcript_variant
CP027632.1	8	22560290	C/A	0.077	0.50	2019 2020 2021	0.19	synonymous_variant
CP027621.1	11	6013057	A/G	0.097	−0.64	2019 2020 2021	0.144	intron_variant
CP027623	13	545286	A/G	0.07	0.52	2020 2021	0.127	intergenic_variant
CP027623	13	4884610	T/G	0.067	1.35	2020 2021	0.457	intergenic_variant

Table 2. Candidate genes located near stable QTNs.

QTN	Candidate Gene	Position	<i>A. thaliana</i> Orthologue	Protein
Chr11:288653	<i>Lus10025717</i>	Chr1:1285554-1289474	<i>AT2G22560</i>	KIP1-like protein
Chr1:1462137	<i>Lus10025756</i>	Chr1:1463118-1465114	<i>AT2G46950</i> , <i>CYP709B2</i>	Cytochrome P450, family 709, subfamily B, polypeptide 2
Chr1:1528323	<i>Lus10025773</i>	Chr1:1527916-1531752	<i>AT1G53050</i>	Protein kinase superfamily protein
Chr1:1722812	<i>Lus10025823</i>	Chr1:1722251-1726433	<i>AT1G79750</i> , <i>NADP-ME4</i>	NADP-malic enzyme 4
Chr1:1854337	<i>Lus10025852</i>	Chr1:1851791-1853596	<i>AT5G17890</i> , <i>CHS3</i> , <i>DAR4</i>	DA1-related protein 4, nucleotide-binding, leucine-rich repeat protein
Chr1:1854337	<i>Lus10025853</i>	Chr1:1854958-1870647	<i>AT5G17020</i> , <i>XPO1A</i>	Exportin 1A
Chr8:22560236	<i>Lus10015356</i>	Chr8:22561110-22562338	<i>AT5G62740</i> , <i>HIR1</i>	SPFH/Band 7/PHB domain-containing, membrane-associated protein family protein
Chr8:22560236	<i>Lus10015357</i>	Chr8:22562785-22565427	<i>AT5G62740</i> , <i>VDAC1</i>	Voltage-dependent anion channel 1
Chr8:22560236	<i>Lus10015344</i>	Chr8:22554684-22556269	<i>AT5G40010</i> , <i>AATP1</i>	AAA-ATPase 1
Chr8:22560236	<i>Lus10015339</i>	Chr8:22569641-22570539	<i>AT3G01270</i>	Pectate lyase family protein
Chr11:6013057	<i>Lus10035917</i>	Chr11:6011724-6014460	<i>AT1G80230</i>	Rubredoxin-like superfamily protein
Chr13:4884610	<i>Lus10009330</i>	Chr13:4886276-4888681	<i>AT1G71400</i> , <i>RLP12</i>	Receptor-like protein 12, RLK
Chr13:4884610	<i>Lus10009332</i>	Chr13:4890865-4893541	<i>AT5G66880</i> , <i>SNRK2.3</i>	Sucrose nonfermenting 1(SNF1)-related protein kinase 2.3

Chromosome 8 contains Chr8:22560236 QTN, which is an upstream transcript variant of the *Lus10015356* gene orthologous to the *A. thaliana* HIR1 gene encoding a hypersensitive induced reaction protein. Upstream of this QTN, at a distance of 5.5 kb is the *Lus10015354* gene orthologous to *AT5G40010* gene encoding AAA-ATPase 1. Downstream of Chr8:22560236 QTN are two genes, *Lus10015357* and *Lus10015359*. The closest gene (2.5 Kb), is the ortholog of the *AT5G62740* gene encoding the voltage-dependent anion channel (VDAC) 1 protein, while the other gene, which is 9.4 Kb further downstream, is orthologous to the *AT3G01270* gene for pectate lyase family protein. The *Lus10035917* gene with the Chr11:6013057 QTN is located in the intron of an ortholog of *AT1G80230* gene for rubredoxin-like superfamily protein (Table 2). On chromosome 13, two candidate genes were detected. The *Lus10009330* encoding receptor-like kinase is 1.2 Kb downstream of QTN Chr13:4884610 and is the ortholog of the *A.thaliana* RLP12 gene. The second proposed candidate gene, *Lus10009332*, is the ortholog of *AT5G66880* encoding sucrose nonfermenting 1 (SNF1)-related protein.

3. Discussion

In contrast to flax rust which, being a model system for H.H. Flor's series of elegant experiments and a basis for the development of gene-for-gene theory [37], was extensively studied in the past, flax wilt, as well as mechanisms leading to the infection, currently remain unknown.

Recently, several high-throughput transcriptomics experiments provided a general insight into flax resistance to pathogens [19–21]. However, the identification of potential key candidate genes in the control of resistance is challenging due to a systemic effect inflicted by infection and, consequently, an unfeasibly large number of genes to be interrogated in focused follow-up studies. To gain an ultimate understanding of biology and mechanisms underlying flax wilt disease, and to prioritize targets for further in-depth experiments, it is essential to make the most of the available genomic and genetic resources, as well as the analytical approaches such as GWAS.

To gain insight into the mechanisms of Fusarium wilt resistance in flax we applied GWAS methods to a collection of 297 flax genotypes which encompassed elite cultivars, breeding lines and landraces. Importantly, the fiber morphotype dominated over the linseed by fifty percent. The disease resistance was estimated by means of the Disease Severity Index (DSI). The DSI shows the normalized proportion of specimens with identical disease symptoms. The genotypes analyzed in this study exhibited a wide variation in resistance to the highly virulent MI39 *F. oxysporum* f.sp *lini* strain in all three successive years of the experiment, thus providing a solid basis for dissecting the genetic architecture of the trait. As expected, elite cultivar and breeding line accessions demonstrated a more than three-fold decrease in DSI as compared to the genotypes with a different selection status, thereby acknowledging the overall success of various breeding programs.

The accurate deciphering of the genetic architecture of a trait in a diverse population is only possible if a statistical model accounts for the spurious associations arising from population structure and family relatedness. Currently, the most popular approach to control for false positives is the incorporation of the population structure and a kinship matrix, which captures family relatedness as covariates in the mixed linear models (MLM). However, all of the MLMs developed to date are single-locus models testing one marker at a time. Such approaches are one-dimensional approximations of a true genetic model of a complex trait controlled by many loci simultaneously and, therefore, they could be inaccurate in controlling false positive associations and estimations of marker effects [38]. To cope with this problem, several single-locus models are generally applied and the associations detected by several models are deemed as significant. We performed association analyses for every year of trait evaluation with six different models: GLM, MLM, CMLM, FarmCPU, SUPER, Blink, as implemented in GAPIT R package [26].

The functional annotation of candidate genes was inferred from their *Arabidopsis* orthologs annotated in the TAIR database [39], as well as from the function of the orthologous genes in other plant species as reported in the literature (Table 2).

The plant response to a pathogen attack involves several tiers of defense, including pathogen sensing by receptor-like kinases (RLK) and R-gene proteins followed by downstream signaling via MAP-kinase cascades, G-proteins and calcium fluxes, which leads to activation of defense and secondary metabolite genes, as well as to the rewiring of hormonal networks. Changes in the calcium status of the cells trigger an oxidative burst. The accumulation of reactive oxygen species (ROS) leads to a hypersensitive response, cell wall protein cross-linking, phytoalexin production, callose deposition and systemic acquired resistance [12,16]. The 640 Kb region on chromosome 1 contains several genes potentially involved in *Fusarium oxysporum* recognition and plant immunity responses. Of these, the *Lus10022557* gene encodes the KIP1-like protein and contains three significant SNPs singled out by the largest negative effects (Table 1). The KIP1-like proteins bind actin [40] and may relay immune signals to inactivate pathogens and interact with RLKs. Therefore, we consider this gene as the most promising candidate for further in-depth downstream analysis [41].

Other potentially promising candidate genes encode the NBS-LRR protein (*Lus10025852*), protein kinase superfamily protein (*Lus10025773*), exportin A1 (*Lus10025853*), NADP-malic enzyme 4 (*Lus10025823*) and cytochrome P450 monooxygenase from CYP709B subfamily (*Lus10025756*). These five proteins are known as significant players in pathogen responses in other plants. First, protein kinases play a central role in signalling during pathogen recognition and the subsequent activation of plant defence mechanisms [42]. The second protein, exportin Xpo1 in *N. benthamiana* is required for elicitor-induced phytoalexin production and the induction of cell death [43]. The third protein, NADP-malic enzyme (NADP-ME), provides the building blocks and energy for the synthesis of two defence-related secondary metabolites, flavonoids and the lignin precursor, monolignol. Moreover, NADP-ME can produce NADPH for synthesis of ROS [44]. Finally, cytochrome P450 monooxygenases from subfamily CYP709B play important roles in plant defence through their involvement in biosynthesis of phytoalexin and some other secondary metabolites [45].

On chromosome 8, we identified four candidate genes in the vicinity of Chr8:22560236 QTN. Two of them, *Lus10015356* and *Lus10015357*, encode a hypersensitive induced reaction (HIR) protein and voltage-dependent anion channel (VDAC) 1 protein, respectively. These genes are likely to be involved in programmed cell death induced by a pathogen attack [46–48]. Of the two other genes, *Lus10015354* encodes AAA-ATPase 1, which plays an important role in the salicylic acid-mediated defence response against the blast fungus *Magnaporthe oryzae* [49] in rice. Finally, *Lus10015359* encodes the pectate lyase superfamily protein. This enzyme degrades pectin, a major component of the plant cell wall. During infection, pathogens not only secrete pectin-degrading enzymes, but also hijack the host signalling pathways to induce cell remodelling by plant-derived enzymes [50].

Lus10035917, a candidate gene containing Chr11:6013057 QTN, encodes rubredoxin-like superfamily protein (Table 2). In plants, it regulates the reactive oxygen species balance. As a result of the plant response to a pathogen attack and abiotic stresses their levels are increased [51–53].

Two candidate genes, *Lus10009330* and *Lus10009332*, were detected downstream of Chr13:4884610 QTN. *Lus10009330* encodes the receptor-like kinase. RLp12, the *A. thaliana* ortholog of this gene, encodes the CLAVATA2-related protein implicated in innate immunity to microbe and nematode infections [54]. The other gene, *Lus10009332* encodes the sucrose nonfermenting 1 (SNF1)-related protein kinase 2.3. (SnRK). SnRK1 regulates carbon metabolism and responds to hormonal signals. In soybean, SnRK promotes resistance to oomycete *Phytophthora sojae* potentially working through the accumulation of salicylic acid (SA) and induction of SA-related genes [55].

Overall, the candidate genes we pinpointed are involved in pathogen recognition, downstream signaling and plant immunity response and, as such, provide excellent targets for further downstream analyses.

4. Materials and Methods

4.1. Plant Material Collection and Phenotyping

A total of 297 flax genotypes were selected from the collection of the Federal Research Center for Bast Fiber Crops (Torzhok, Russia) (Table S1). The dataset included 179 fiber flax accessions, 117 linseed flax accessions and 1 accession of unknown morphotype. Oilseed group was further subdivided into the following subgroups: 98 intermediate accessions, 4 large-seeded accessions and 15 crown accessions. There were landraces, elite cultivars and breeding lines in the dataset.

The accessions were evaluated for *Fusarium oxysporum* resistance under greenhouse conditions in randomized complete block design. Three evaluations were performed during 2019–2021. In each evaluation, each specimen was replicated 16 times by sowing all the seeds in cross-container rows. The container dimensions were 550 × 85 × 20 cm. Two genotypes, AP5 and I-7, were used as susceptible and resistant genotypes to *Fusarium* wilt controls. The infection background was created by introducing 400 g of a pure culture of MI39 *Fusarium oxysporum* f.sp. *lini* strain into a container. The seeds were planted on a 12 day after the inoculation with the pure fungal culture.

The pure culture inoculum was prepared by first growing MI39 on beer-wort agar-agar medium with a subsequent incubation on the oat grain substrate (the grain to water ratio of 1 to 1.75) for 3–4 weeks; such a time period was sufficient for macro and microconidia development. After three to four weeks, when oats were completely infected by the fungus, the pathogen was introduced into the soil. The required amount of the introduced infection for a container (550 × 85 × 20 cm) was established experimentally. The indicator of the reliability of the infectious background was the standard varieties (resistant and susceptible genotypes), which were sown along the edges and in the middle of each container (16 seeds per row).

The evaluation of the disease severity was carried out during the harvesting period in the phase of early yellow ripeness. The DSS (Disease Severity score) grades ranged from 0 to 3, where the 0 value stood for a healthy plant, 1 indicated a partial plant browning or stem browning from one side, 2 indicated a fully browned plant with bolls, and, finally, 3 corresponded to a fully browned plant that collapsed prior to the formation of bolls. Proceeding from these grades the Disease Severity index (DSI) was calculated using the standard formula accepted in phytopathology [56]:

$$DSI = \frac{\sum ab}{AK} 100\%,$$

where a is the number of plants with identical DSS, b is the estimated DSS; A is the total number of plants and K is the highest DDS grade (i.e., 3).

The plants were considered resistant if $DSI \leq 20\%$, as weakly susceptible in the case of $20\% < DSI \leq 30\%$, as moderately susceptible if $30\% < DSI \leq 50\%$ and, finally, as completely susceptible when DSI value exceeded 50%. The phenotype values were quantile normalized prior to GWAS analyses.

4.2. DNA Sequencing and Variant Calling

DNA was extracted from collected leaves with DNeasy Plant Mini Kit (Qiagen, Stanford, CA, USA). DNA sequencing was performed at the BGI (Hong Kong, China) using Illumina protocol, generating paired-end reads 150 bp in length. 1143.850625 GB of raw data comprising 7.626 billion reads with an average of $9.3\times$ coverage or 3.7 Gbp per sample were generated. Processed reads were aligned to NCBI flax reference genome assembly ASM22429v2 with bwa-mem using default parameters [57]. Variant calling was run using NGSEP [58] (v. 4.0) and identified 3416829 biallelic SNPs, which were further filtered using

VCFtools [59] to obtain a minor allele frequency (i.e., MAF) > 5% and genotype call rate >85%. A total of 72,526 SNPs passed all filters and remained for further analysis.

4.3. Genetic Data Analyses

Principal component analysis (PCA) was conducted using the PCA tools R package [60]. SNP depth and distribution across chromosomes was plotted using rMVP package [61].

The linkage disequilibrium (LD) decay was evaluated using squared Pearson's correlation coefficient (r^2). The PopLDdecay [62] version 3.4.1 was run to calculate r^2 in a 500 kb window. The LD decay was calculated based on r^2 and the distance for each pair of SNPs using an R script in accordance with Hill–Weir approximation [63].

We applied Mann–Whitney–Wilcoxon test [64] to make group comparisons.

4.4. GWAS

The genome-wide association analyses were performed with GAPIT3 R package [36]. We selected the following models: GLM, MLM, CMLM, FarmCPU, SUPER, Blink according to the workflow proposed by the package authors for multiple models testing. The FDR adjusted p -value threshold for candidate QTNs selection was set at 0.05.

4.5. Candidate Genes

A function of the candidate genes containing the identified QTNs was inferred from the function of their *Arabidopsis* orthologs in the TAIR database, from flax genome annotations kindly provided by Cloutier group (Ottawa Research and Development Centre, Canada) [65], as well as from the function of the homologous genes in other plant species, as described in the literature. We also searched for candidate genes within a window surrounding detected QTNs, where width was defined as proceeding from the LD decay estimated for each chromosome (Figure S4).

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ijms222212383/s1>.

Author Contributions: M.S., T.R.; Methodology, M.B., A.K.; Software, A.K.; Investigation, T.R., M.B.; Resources, T.R.; Data Curation, M.B.; Writing—M.S., A.S. and A.K. Writing—Review and Editing, A.K., A.S.; Visualization, M.B. and A.S.; Supervision, M.S.; Funding Acquisition, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by RSF fund, grant 19-16-00030.

Data Availability Statement: The data that support the findings of this study are openly available at European Variation Archive (EVA) Project: PRJEB46073 Analyses: ERZ2775743 (genetic variants).

Acknowledgments: The authors gratefully acknowledge Peter the Great St. Petersburg Polytechnic University Centre for Supercomputing (scc.spbstu.ru) for providing excellent computational resources and support for this project. We would like to express our special appreciation and thanks to members of the Cloutier group (Ottawa Research and Development Centre, Canada) for sharing the latest version of flax genome annotation [22].

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

References

1. Boyle, L.W. *Histological Characters of Flax Roots in Relation to Resistance to Wilt and Root Rot*; Technical Bulletin, No. 458; United States Department of Agriculture: Washington, DC, USA, 1934; pp. 1–18.
2. Houston, B.R.; Knowles, P.F. Fifty-years survival of Flax *Fusarium* wilt in the absence of flax culture. *Plant Dis. Rep.* **1949**, *33*, 38–39.
3. Saharan, G.; Mehta, N.; Sangwan, M. Fungal Diseases of Linseed. In *Diseases of Oilseed Crops*; Indus Publishing Company: New Delhi, India, 2005; pp. 176–201.
4. Rashid, K.Y.; Kenaschuk, E.O. Effect of trifluralin on *Fusarium* wilt in flax. *Can. J. Plant Sci.* **1993**, *73*, 893–901. [[CrossRef](#)]
5. Liu, X.Y.; Chen, S.L.; Sun, Q.A.; He, D.T.; Wu, Y.N. Evaluation of *Fusarium* wilt resistance of flax varieties. *Sci. Agric. Sin.* **1993**, *26*, 44–49.

6. Ondrej, M. Evaluation of flax genepool according to resistance to *Fusarium* wilt of flax and to mildew. *Plant Genet. Resour.* **1993**, *92*, 54–58.
7. Rozhmina, T.A.; Loshakova, N.I. Samples of fiber and oil flax (*Linum usitatissimum* L.)—Sources of effective genes for resistance to *Fusarium* wilt and its dependence on temperature. *Agric. Biol.* **2016**, *51*, 310–317.
8. Rozhmina, T.A. Identification of effective genes for resistance to *Fusarium* wilt in fiber flax varieties. *Biol. Agric.* **2017**, *4*, 10–12.
9. Sánchez-Martín, J.; Keller, B. Contribution of recent technological advances to future resistance breeding. *Theor. Appl. Genet.* **2019**, *132*, 713–732. [[CrossRef](#)]
10. Sánchez-Bayo, F. Impacts of Agricultural Pesticides on Terrestrial Ecosystems. In *Ecological Impacts of Toxic Chemicals*; Sánchez-Bayo, F., van den Brink, P.J., Mann, R.M., Eds.; Bentham Science Publishers Ltd.: Sharjah, United Arab Emirate, 2011; pp. 63–87.
11. Kourelis, J.; van der Hoorn, R.A.L. Defended to the nines: 25 Years of resistance gene cloning identifies nine mechanisms for R protein function. *Plant Cell* **2018**, *30*, 285–299. [[CrossRef](#)]
12. Jones, J.D.G.; Dangl, J.L. The plant immune system. *Nature* **2006**, *444*, 323–329. [[CrossRef](#)]
13. Tsuda, K.; Katagiri, F. Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. *Curr. Opin. Plant. Biol.* **2010**, *13*, 459–465. [[CrossRef](#)]
14. Künstler, A.; Bacsó, R.; Gullner, G.; Hafez, Y.M.; Király, L. Staying alive—Is cell death dispensable for plant disease resistance during the hypersensitive response? *Physiol. Mol. Plant Pathol.* **2016**, *93*, 75–84. [[CrossRef](#)]
15. Zhong, Y.; Cheng, Z.-M. A unique RPW8-encoding class of genes that originated in early land plants and evolved through domain fission, fusion, and duplication. *Sci. Rep.* **2016**, *6*, 32923. [[CrossRef](#)]
16. Balint-Kurti, P. The plant hypersensitive response: Concepts, control and consequences. *Mol. Plant Pathol.* **2016**, *20*, 1163–1178. [[CrossRef](#)]
17. Ma, Y.; Chhapekar, S.S.; Lu, L.; Oh, S.; Singh, S.; Kim, C.S.; Kim, S.; Choi, G.J.; Lim, Y.P.; Choi, S.R. Genome-wide identification and characterization of NBS-encoding genes in *Raphanus sativus* L. and their roles related to *Fusarium oxysporum* resistance. *BMC Plant Biol.* **2021**, *21*, 47. [[CrossRef](#)]
18. Block, A.; Alfano, J.R. Plant targets for *Pseudomonas syringae* type III effectors: Virulence targets or guarded decoys? *Curr. Opin. Microbiol.* **2011**, *14*, 39–46. [[CrossRef](#)]
19. 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* (Suppl. 2), 253. [[CrossRef](#)]
20. Galindo-González, 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](#)]
21. Boba, A.; Kostyn, K.; Kozak, B.; Zalewski, I.; Szopa, J.; Kulma, A. Transcriptomic profiling of susceptible and resistant flax seedlings after *Fusarium oxysporum lini* infection. *PLoS ONE* **2021**, *16*, e0246052. [[CrossRef](#)]
22. Chandrawati, S.N.; Kumar, R.; Kumar, S.; Singh, P.K.; Yadav, V.K.; Ranade, S.A.; Yadav, H.K. Genetic diversity, population structure and association analysis in linseed (*Linum usitatissimum* L.). *Physiol. Mol. Biol. Plants Int. J. Funct. Plant Biol.* **2017**, *23*, 207–219. [[CrossRef](#)]
23. 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](#)]
24. 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](#)]
25. 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.* **2019**, *286*, 98–107. [[CrossRef](#)]
26. 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.* **2020**, *10*, 1682. [[CrossRef](#)]
27. 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.* **2021**, *296*, 877–891. [[CrossRef](#)] [[PubMed](#)]
28. 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. *Appl. Genet.* **2021**, *134*, 191–212. [[CrossRef](#)] [[PubMed](#)]
29. 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.* **2018**, *8*, 2232. [[CrossRef](#)]
30. Knowles, P.; Houston, B. Inheritance of Resistance to *Fusarium* Wilt of Flax in Dakota Selection 48–94. *Agron. J.* **1955**, *47*, 131–135. [[CrossRef](#)]
31. Spielmeyer, W.; Green, A.; Bittisnich, D.; Mendham, N.; Lagudah, E. Identification of quantitative trait loci contributing to *Fusarium* wilt resistance on an AFLP linkage map of flax (*Linum usitatissimum*). *Appl. Genet.* **1998**, *97*, 633–641. [[CrossRef](#)]

32. Spielmeyer, W.; Lagudah, E.; Mendham, N.; Green, A. Inheritance of resistance to flax wilt (*Fusarium oxysporum* f.sp. *lini* Schlecht) in a doubled haploid population of *Linum usitatissimum* L. *Euphytica* **1998**, *101*, 287–291.
33. Caballo, C.; Castro, P.; Gil, J.; Millan, T.; Rubio, J.; Die, J.V. Candidate genes expression profiling during wilting in chickpea caused by *Fusarium oxysporum* f.sp. *ciceris* race 5. *PLoS ONE* **2019**, *14*, e0224212. [[CrossRef](#)]
34. Gupta, P.K.; Kulwal, P.L.; Jaiswal, V. Association mapping in plants in the post-GWAS genomics era. *Adv. Genet.* **2019**, *104*, 75–154.
35. Duk, M.; Kanapin, A.; Surkova, S.; Bankin, M.; Rozhmina, T.; Samsonova, A.; Samsonova, M. The genetic landscape of fiber flax. *Front. Plant Sci.* **2021**, in press.
36. Wang, J.; Zhang, Z. GAPIT Version 3: Boosting Power and Accuracy for Genomic Association and Prediction. *Genom. Proteom. Bioinform.* **2021**, in press. [[CrossRef](#)]
37. Flor, H. Tests for allelism of rust-resistance genes in flax. *Crop Sci.* **1965**, *5*, 415–418. [[CrossRef](#)]
38. Wang, S.B.; Feng, J.Y.; Ren, W.L.; Huang, B.; Zhou, L.; Wen, Y.J.; Zhang, J.; Dunwell, J.M.; Xu, S.; Zhang, Y.M. Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. *Sci. Rep.* **2016**, *6*, 19444. [[CrossRef](#)] [[PubMed](#)]
39. Berardini, T.Z.; Reiser, L.; Li, D.; Mezheritsky, Y.; Muller, R.; Strait, E.; Huala, E. The Arabidopsis information resource: Making and mining the “gold standard” annotated reference plant genome. *Genesis* **2015**, *53*, 474–485. [[CrossRef](#)]
40. Porter, K.; Day, B. From filaments to function: The role of the plant actin cytoskeleton in pathogen perception, signaling and immunity. *J Integr. Plant Biol.* **2016**, *58*, 299–311. [[CrossRef](#)] [[PubMed](#)]
41. Skirpan, A.L.; McCubbin, A.G.; Ishimizu, T.; Wang, X.; Hu, Y.; Dowd, P.E.; Ma, H.; Kao, T. Isolation and characterization of kinase interacting protein 1, a pollen protein that interacts with the kinase domain of PRK1, a receptor-like kinase of petunia. *Plant Physiol.* **2001**, *126*, 1480–1492. [[CrossRef](#)] [[PubMed](#)]
42. Romeis, T. Protein kinases in the plant defence response. *Curr. Opin. Plant Biol.* **2001**, *4*, 407–414. [[CrossRef](#)]
43. Mizuno, Y.; Imano, S.; Camagna, M.; Suzuki, T.; Tanaka, A.; Sato, I.; Chiba, S.; Kawakita, K.; Takemoto, D. Nicotiana benthamiana exportin 1 is required for elicitor-induced phytoalexin production, cell death induction, and resistance against potato late blight pathogen *Phytophthora infestans*. *J. Gen. Plant Pathol.* **2019**, *85*, 347–355. [[CrossRef](#)]
44. Sun, X.; Han, G.; Meng, Z.; Lin, L.; Sui, N. Roles of malic enzymes in plant development and stress responses. *Plant Signal. Behav.* **2019**, *14*, e1644596. [[CrossRef](#)] [[PubMed](#)]
45. Jun, X.U.; Wang, X.Y.; Guo, W.Z. The cytochrome P450 superfamily: Key players in plant development and defense. *J. Integr. Agric.* **2015**, *14*, 1673–1686.
46. Zhou, L.; Cheung, M.Y.; Li, M.W.; Fu, Y.; Sun, Z.; Sun, S.M.; Lam, H.M. Rice hypersensitive induced reaction protein 1 (OsHIR1) associates with plasma membrane and triggers hypersensitive cell death. *BMC Plant Biol.* **2010**, *10*, 29. [[CrossRef](#)]
47. Choi, H.W.; Kim, Y.J.; Hwang, B.K. The hypersensitive induced reaction and leucine-rich repeat proteins regulate plant cell death associated with disease and plant immunity. *Mol. Plant-Microbe Interact.* **2011**, *24*, 68–78. [[CrossRef](#)] [[PubMed](#)]
48. Kusano, T.; Tateda, C.; Berberich, T.; Takahashi, Y. Voltage-dependent anion channels: Their roles in plant defense and cell death. *Plant Cell Rep.* **2009**, *28*, 1301–1308. [[CrossRef](#)] [[PubMed](#)]
49. Liu, X.; Inoue, H.; Tang, X.; Tan, Y.; Xu, X.; Wang, C.; Jiang, C.-J. Rice *OsAAA-ATPase1* is Induced during Blast Infection in a Salicylic Acid-Dependent Manner, and Promotes Blast Fungus Resistance. *Int. J. Mol. Sci.* **2020**, *21*, 1443. [[CrossRef](#)] [[PubMed](#)]
50. Shin, Y.; Chane, A.; Jung, M.; Lee, Y. Recent Advances in Understanding the Roles of Pectin as an Active Participant in Plant Signaling Networks. *Plants* **2021**, *10*, 1712. [[CrossRef](#)]
51. Li, Y.; Liu, P.; Takano, T.; Liu, S. A chloroplast-localized rubredoxin family protein gene from *Puccinellia tenuiflora* (PutRUB) increases NaCl and NaHCO₃ tolerance by decreasing H₂O₂ accumulation. *Int. J. Mol. Sci.* **2016**, *17*, 804. [[CrossRef](#)]
52. Li, Y.; Liu, P.P.; Ni, X. Molecular evolution and functional analysis of rubredoxin-like proteins in plants. *BioMed Res. Int.* **2019**, *2019*, 2932585. [[CrossRef](#)]
53. O’Brien, J.A.; Daudi, A.; Butt, V.S.; Paul Bolwell, G. Reactive oxygen species and their role in plant defence and cell wall metabolism. *Planta* **2012**, *236*, 765–779. [[CrossRef](#)]
54. Pan, L.; Lv, S.; Yang, N.; Lv, Y.; Liu, Z.; Wu, J.; Wang, G. The multifunction of CLAVATA2 in plant development and immunity. *Front. Plant Sci.* **2016**, *7*, 1573. [[CrossRef](#)]
55. Wang, L.; Wang, H.; He, S.; Meng, F.; Zhang, C.; Fan, S.; Wu, J.; Zhang, S.; Xu, P. GmSnRK1.1, a sucrose non-fermenting-1(SNF1)-related protein kinase, promotes soybean resistance to *Phytophthora Sojae*. *Front. Plant Sci.* **2019**, *10*, 996. [[CrossRef](#)]
56. Loshakova, N.I.; Krylova, T.V.; Kudryavtseva, L.P. *Guidelines for the Phytopathological Assessment of the Resistance of Fiber Flax to Diseases*; Publishing house of the Russian Academy of Agricultural Sciences: Moscow, Russia, 2000.
57. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)]
58. Tello, D.; Gil, J.; Loaiza, C.D.; Riascos, J.J.; Cardozo, N.; Duitama, J. NGSEP3: Accurate variant calling across species and sequencing protocols. *Bioinformatics* **2019**, *35*, 4716–4723. [[CrossRef](#)]
59. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. 1000 Genomes Project Analysis Group. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [[CrossRef](#)] [[PubMed](#)]

60. Blighe, K.; Lun, A. PCAtools: Everything Principal Components Analysis. Available online: <https://github.com/kevinblighe/PCAtools> (accessed on 14 September 2021).
61. Yin, L.; Zhang, H.; Tang, Z.; Xu, J.; Yin, D.; Zhang, Z.; Yuan, X.; Zhu, M.; Zhao, S.; Li, X. rMVP: A Memory-efficient, Visualization-enhanced, and Parallel-accelerated tool for Genome-Wide Association Study. *Genom. Proteom. Bioinform.* **2021**, in press. [[CrossRef](#)] [[PubMed](#)]
62. Zhang, C.; Dong, S.-S.; Xu, J.-H.; He, W.-M.; Yang, T.-L. PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **2019**, *35*, 1786–1788. [[CrossRef](#)] [[PubMed](#)]
63. Hill, W.G.; Weir, B.S. Variances and covariances of squared linkage disequilibria in finite populations. *Theor. Popul. Biol.* **1988**, *33*, 54–78. [[CrossRef](#)]
64. Mann, H.B.; Whitney, D.R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* **1947**, *18*, 50–60. [[CrossRef](#)]
65. You, F.; Cloutier, S. Mapping quantitative trait loci onto chromosome-scale pseudomolecules in flax. *Methods Protoc.* **2020**, *3*, 28. [[CrossRef](#)]