The Complex Genetic Architecture of Early Root and Shoot Traits in Flax Revealed by Genome-Wide Association Analyses

Demissew Sertse1,2, Frank M. You2, Sridhar Ravichandran2 and Sylvie Cloutier1,2*

1 Department of Biology, University of Ottawa, Ottawa, ON, Canada, 2 Ottawa Research and Development Center, Agriculture and Agri-Food Canada, Ottawa, ON, Canada

Roots are fundamental organs for water and nutrient uptake as well as for signal transduction in response to biotic and abiotic stresses. Flax has a shallow tap root system that relies mostly on top soil nutrient and moisture resources. The crop can easily be outcompeted by weeds or other crops in intercropping systems, especially in moisture deficit conditions. However, there is a wide range of variation among genotypes in terms of performance under scarce resources such as moisture limitation. Here we phenotyped 15 root, two shoot traits and shoot to root dry weight ratio on 115 flax accessions grown in a hydroponic pouch system and performed a genome-wide association study (GWAS) based on seven different models to identify quantitative trait loci underlying these traits. Significant variation among genotypes was observed for the two shoot and 12 of the 14 root traits. Shoot dry weight was correlated with root network volume, length, surface area, and root dry weight ($r > 0.5, P < 0.001$) but not significantly correlated with root depth ($r = 0.033, P > 0.05$). The seven GWAS models detected a total of 228 quantitative trait nucleotides (QTNs) for 16 traits. Most loci, defined by an interval of 100 kb up and downstream of the QTNs, harbored genes known to play role(s) in root and shoot development, suggesting them as candidates. Examples of candidate genes linked to root network QTNs included genes encoding GRAS transcription factors, mitogen-activated protein kinases, and auxin related lateral organ boundary proteins while QTN loci for shoot dry weight harbored genes involved in photomorphogenesis and plant immunity. These results provide insights into the genetic bases of early shoot and root development traits in flax that could be capitalized upon to improve its root architecture, particularly in view of better withstandng water limiting conditions during the cropping season.

Keywords: flax, root, shoot, genome-wide association study, quantitative trait nucleotides, candidate genes

INTRODUCTION

Roots are vital organs in terrestrial higher plants for acquisition of essential nutrients and water. Because roots function in a bio-physico-chemically dynamic rhizosphere, they play an important role in controlling and regulating the impacts of various edaphic factors through internal physiological adjustments (Jackson et al., 1990; Hodge 2004) and signal transduction (Batool et al., 2018).
Adaptation of plants to a scarcity of resources and associated edaphic factors is therefore governed by their root system. The architectural features of root systems are crucial in efficiently tapping the available resources such as water and nutrients in the rhizosphere (Yue et al., 2006) and inadequate root system development may lead to significant yield losses in water-limiting conditions (Henry et al., 2011). Understanding root traits and resource use efficiencies of the root system is key to crop yield improvement (Kell, 2011). However, the inaccessibility of the rhizosphere has made root trait studies challenging; hence these traits have been scantily considered in varietal improvement.

Flax, one of the founder crops of agriculture (Weiss and Zohary, 2011), has been grown as both fiber and oilseed crops for nearly the entire span of its cultivation history (Herbig and Maier, 2011). The crop is adapted to diverse ecologies, from the warm Indian subcontinent to the cool temperate areas in Eurasia (Casa et al., 1999; Sertse et al., 2019). A wide range of uses are derived from its stem fibers and its oil-rich seeds (Singh et al., 2011). Flax production, however, is constrained by low yield (Wittkop et al., 2009). The meager flax yield improvements of ~0.5 ton/Ha since the 1960s, obtained through breeding and agronomic practices, have not been sufficient to impact its production in a major way and production continues to decline as growers are shifting to better yielding crops such as soybean, canola (http://www.fao.org/faostat/en/#data). High and stable yielding cultivars are urgently needed to rekindle growers' interest and meet market potential.

Flax, like other oilseed crops, is a tap-rooted plant. Compared to canola, sunflower, and safflower, flax has a shallower root system and, as such, it mainly relies on moisture and nutrient resources available in the soil’s top layers (Hocking et al., 1997; Kar et al., 2007) mainly within 70 cm depth (Flax Council of Canada, 2015; Hall et al., 2016). Flax roots can grow to depth of 90–120 cm with a lateral spread of ~ 30 cm in light soil (Gill 1987). However, the proportion of its root deeper than 60 cm is only 4–7% (Hall et al., 2016) and the roots rarely grow beyond 80 cm (Flax Council of Canada, 2015). Unlike other crops such as canola, mustard and wheat that have aggressive root growth before early flowering the fastest root growth in flax is between early flowering to late flowering where it declines late bolling stage onwards (Liu et al., 2011). Under limited water, flax can be easily outcompeted by many weeds, exacerbating the competition for resources (Bell and Nalewaja, 1968; Gruenhagen and Nalewaja, 1969; Alessi and Power, 1970; Klimek-Kopyra et al., 2015). Flax is also less competitive than fibrous rooted cereals such as wheat (Morillon and Lassalles, 2002) that are with more extensive roots than flax in top soil layer (Gill, 1987). Despite the fact that flax performed well being intercropped with chickpea in irrigated field when flax grown in the furrow and chickpea on the ridge, poor flax performance was noticed when both the crop grown on flat field suggesting the low competence of flax (Ahlawat and Gangaiah, 2010). However, a wide range of performance among flax genotypes under different moisture regimes (Foster et al., 1998; Diederichsen et al., 2006) can be attributed to variations of their root system (Cattivelli et al., 2008).

Prompted by the current large sets of genetic data from genome-wide association studies (GWAS) and advancements in imaging and data processing, high throughput digital root phenotyping techniques have become attractive. Hund et al. (2009) used a pouch system to phenotype root traits in maize and, this system has subsequently been applied in other crops such as wheat (Atkinson et al., 2015), Brassica (Thomas et al., 2016), and barley (Canto et al., 2018) for examples.

Here we applied this technique to study the early root and shoot development of a flax mini-core collection (n = 115) that comprises representative genotypes from all major flax growing regions of the world. A GWAS was performed using a set of single nucleotide polymorphism (SNP) markers obtained from shotgun short-read re-sequencing data of the germplasm collection. The objectives of this research were 1) to assess the extent of the variation in root traits among genotypes, 2) to identify quantitative trait nucleotides (QTNs) associated with the genetic architecture of selected root and shoot traits, and, 3) to identify candidate genes for the traits harbored at the QTN loci.

**MATERIALS AND METHODS**

**Plant Materials**

A flax mini-core collection (n = 115) that comprised >95% of the genetic diversity (Soto-Cerda et al., 2013) of the flax core collection (n = 407) (Diederichsen et al., 2013) was used. The 75 linseed, 33 fiber flax, and 7 accessions of unknown type of the mini-core collection were collected from 20 countries that represent all major flax growing regions of the world.

**Phenotyping**

Early development root and shoot phenotyping was performed in a hydroponic pouch system modified from Hund et al. (2009). The system comprised two large plastic bins each containing 108 L of nutrient solution that was transferred every 3 h from one bin to the other using timer-operated peristaltic pumps. Aluminum frames mounted on top of the two opaque bins were used to hang the pouches. The bins, but not the plants, were covered with black polyethylene sheets to prevent algal growth. This system was installed in a growth chamber (Conviron PGC20, Serial No. 150342, Controlled Environment Ltd, Canada) maintained at 21/18°C with a 16 h day/8 h night photoperiod.

Sterile 24x30 cm blue germination blotting papers (SGB1924B, Anchor Paper Company, St Paul, MN, USA) were inserted in Ziploc bags. The assemblies were attached to rust-proofed rods using fold back clips that were suspended in the bins with the open side of the Ziploc bags down to allow blotting up of the nutrient solution. Assemblies were labeled and randomized for each of the three consecutive biological replicates performed (Figure 1B).

The Hoagland nutrient solution (HOP1, Hoagland’s No. 2 basal salt mixture, Caisson Labs, Smithfield, UT, USA) was made at 25% strength with deionized water and adjusted to pH 6.3. The pouches were randomly distributed in the two bins. After complete moistening of the blotting paper, three seeds were inserted directly in the paper using tweezers, ~2 cm below the top edge and slightly spaced out near the center.
The siphons that transferred the nutrient solution were set ~3 cm from the bottom of the bins, i.e., touching the bottom of the blotting papers, in order to keep them moist at all time. Five days after seeding, germinated seeds were thinned out to one seedling per pouch. The retained seedlings were allowed to grow for 18 days. On the 19th day, plants were carefully removed from their assemblies for imaging and measurements. Shoots, cut at the shoot and root junction, were measured with a ruler to estimate shoot length prior to being transferred to labeled envelopes and dried for 3 days at 60°C to measure dry weight. The root system of each plant was imaged using a Canon camera (EOS Rebel T5i) mounted on a custom stand to ensure a consistent 40 cm distance between lens and roots. After imaging, the roots were gently peeled from the paper and processed as the shoots to measure their dry weight. Three biological replicates were thus performed consecutively.

Root images were processed using the General Image Analysis of Roots (GIA Roots) software (Galkovskyi et al., 2012). Each image was scaled using the 30 cm edge of the paper as reference. Scaled images were then analyzed for 14 root traits (Table 1, see details of the traits at http://www.rootnet.biology.gatech.edu/giaroots/download/recent/gia_roots_manual.pdf).

Shoot dry weight, shoot length, root dry weight, shoot to root dry weight ratio, and the 14 root traits measured by the GIA software were analyzed. Basic statistics of the 16 traits were computed for 111 genotypes (Table 1). Four genotypes with a single replicate

FIGURE 1  |  Early root phenotyping experiment set-up and representative root images. (A) Diagrammatic sketch of the phenotyping system, arrows show the water flow direction, (B) close-up of the experimental set-up showing a partial view of the upper section of the blotting papers, labels, and plantlets, (C) extensive root system of TMP-2530 (U_MAR_C_CN98193), and (D) root system of flax cultivar Hanley.
TABLE 1 | Phenotypic traits and their basic statistics of three replicates of the 111 accessions of the flax mini-core collection.

| Trait                        | Abbreviation | Description                                                                 | Unit¹ | Range          | Median       | Mean ± SD² |
|------------------------------|--------------|------------------------------------------------------------------------------|-------|----------------|--------------|------------|
| Average root diameter        | ARD          | Average individual root diameter                                            | cm    | 0.01–0.04      | 0.03         | 0.035 ± 0.0001** |
| Maximum number of roots      | MaxR         | After sorting the number of roots crossing a horizontal line from smallest to largest, the maximum number is considered to be the 84th-percentile value (one standard deviation) | count | 2–19          | 8.33         | 8.94 ± 3.37**                            |
| Median number of roots       | MedR         | The result of a vertical line sweep in which the number of roots that crossed a horizontal line was estimated, and then the median of all values for the extent of the network was calculated. | count | 1–10          | 3.67         | 3.99 ± 1.65** |
| Network area                 | NWA          | Total area covered by all roots                                             | cm²   | 1.92–14.67     | 7.31         | 7.27 ± 2.15**                            |
| Network depth                | NWDep        | The maximum depth reached by the root                                      | cm    | 10.95–24.91    | 20.29        | 20.04 ± 2.51                              |
| Network distribution         | NWDist       | The fraction of root network in the lower two third of the network (analogy of root depth density) | na    | 0.26–1.98      | 0.79         | 0.865 ± 0.331                             |
| Network length               | NWL          | Total length of the entire network (~half perimeter)                       | cm    | 156.2–513.7    | 255.9        | 245.3 ± 77.5**                            |
| Network perimeter            | NWPer        | Total length following all root surfaces                                   | cm    | 111.7–1061.7   | 539.1        | 510.2 ± 163.8**                           |
| Network surface area         | NWSA         | The sum of surface area of all roots in the network                        | cm²   | 7.03–53.93     | 26.79        | 26.49 ± 7.94**                            |
| Network volume               | NWV          | The total volume of all roots in the network                               | cm³   | 0.08–0.48      | 0.25         | 0.252 ± 0.074*                            |
| Network width                | NWW          | The maximum linear width attained by the root                             | cm    | 4.17–21.33     | 12.42        | 12.43 ± 3.12**                            |
| Network width to depth       | NWW_Dep      | Ratio of NWV to NWDep (NWV/NWDep)                                          | na    | 0.203–1.316    | 0.64         | 0.632 ± 0.159*                            |
| Root dry weight              | RDWt         | Oven dried root weight                                                     | g     | 0.008–0.045    | 0.02         | 0.024 ± 0.006                             |
| Shoot dry weight             | SDWt         | Oven dried shoot weight                                                    | g     | 0.015–0.064    | 0.03         | 0.026 ± 0.008*                            |
| Shoot length                 | SL           | Length of shoot from root collar to tip of the shoot                      | cm    | 4.95–13.27     | 7.67         | 7.91 ± 1.30*                             |
| Shoot : Root ratio           | S_RDWt       | Ratio of SDWt to RDWt (SDWt/RDWt)                                          | na    | 0.70–3.00      | 1.00         | 1.30 ± 0.44                              |
| Specific root length 1       | SRL          | Ratio of NWL to NWL/NWL                                                   | cm²   | 686.4–1249     | 990.4        | 976.6 ± 116.4**                           |
| Specific root length 2       | SRL2         | Ratio of N WL to RDW (NWL/RDW)                                             | cm/g  | 5,615–19,685.5 | 10877        | 10,765.99 ± 2,695.79                     |

¹cm = centimeter; g = gram; na = not applicable.
²SD = standard deviation; * and ** = significant at P < 0.05 and 0.001, respectively.

were not included in the analysis. A one-way analysis of variance (ANOVA) was performed for each trait using R. Pearson pairwise correlation coefficients between traits were calculated and summarized using R package sjPlot (Lüdecke, 2017). To have insight into the potential effect of geography of origin, variation due to root network length among geographic regions was illustrated in boxplot using R.

Genotyping and Genetic Data Analysis
SNP data of the entire flax core collection (n = 407) was previously generated after resequencing each genotypes using Illumina HiSeq 2000 platform in 100 bp paired-end mode to at an average coverage of 17X. The alignment, SNP call, and quality control such as removal of SNPs in long terminal repeat region were performed as previously described (Sertse et al., 2019). The SNP data for the mini-core collection was extracted from this core collection dataset.

Single Nucleotide Polymorphism Filtering and Preparation of Datasets
For this study, only SNPs with no missing data were used. Because the minor allele frequency (MAF) of an SNP could differ between the core and the mini-core collections, we used two datasets for GWAS. The first dataset (7K) included all SNPs with no missing data regardless of their MAF in the mini-core collection because these had already previously been included based on MAF > 5% criteria in the core collection. The second dataset was smaller (3K) because SNPs with MAF < 5% in the mini-core collection per se were removed. The two datasets were analyzed separately and results were compared.

Genetic Structure and Variation Analysis
To estimate the possible number of ancestral populations (K), a cross-validation technique (Alexander and Lange, 2011) was applied. Analysis of ancestral proportion of each genotype (Pritchard et al., 2000) was performed for K values ranging from 2 to 20 using sparse non-negative matrix factorization (sNMF) (Frichot et al., 2014) of the R package LEA (Frichot and François 2015) with default parameters except that the number of runs was increased from 10 to 20. The K value producing the lowest cross-validation error was accepted as the number of ancestral populations. The same package was used to visualize the cross-validation and to generate the structure plot. Neighbor-joining (NJ) phylogenetic and principal component (PC) analyses were performed using TASSEL v 5.2 (Bradbury et al., 2007).
of the two analyses were summarized using Tree of Life (iTOL) v3 (Letunic and Bork, 2016) and R, respectively. Genotypes were assigned to the suggested ancestral populations based on their Q-matrix. Populations were named based on the passport data of their members indicating geography of origin.

**Phenotype-Genotype Association Analyses and Mapping**

The use of multi-locus methods that capture small effect loci in complex polygenic traits such as in plant roots and shoots has recently become a feasible approach. To benefit the algorithmic merits of different models and support results of one by another, it is also advantageous to apply multiple methods (Zhang et al., 2019). To assess the genetic variants underlying each root and shoot traits in this study, the multi-locus GWAS methods FASTmrEMMA (Wen et al., 2016), FASTmrMLM (Tamba and Zhang, 2018), ISIS EM-BLASSO (Tamba et al., 2017), mrMLM (Wang et al., 2016), pKWmEB (Ren et al., 2018), and pLARmEB (Zhang et al., 2017) included in the R package multi locus random-SNP mixed linear model (mrMLM) (Wen et al., 2017) were applied. The single locus genome scan method latent factor mixed linear model (LFMM) in the R package lmm4 (Frichot et al., 2013) was also used. To control the type I error in multiple comparison, false discovery rate (FDR) correction at α = 0.05 was applied (Benjamini and Hochberg, 1995) for all models to identify significant QTNs. For LFMM, Bonferroni correction factor at α = 0.05 (0.05/n, n = the number of total SNPs) was also used as a comparative method. Quantitative trait loci (QTL) regions spanning 100 Kb up and downstream of all associated QTNs were examined for the predicted coding genes they harbored using the flax reference genome (You et al., 2018). Predicted functions of genes identified within each QTL were investigated based on their Arabidopsis orthologues (www.arabidopsis.org). Strongly associated SNPs and their putative underlying genes were illustrated on the flax pseudomolecules (You et al., 2018) using MapChart 2.3 (Voorrips, 2002). The favorability of alleles at QTNs detected by at least two of the multi-locus models with high phenotypic variance explained (PVE) (R² > 5%), was illustrated using box plot based on mean phenotypic value of genotypes with each allele.

For genes previously identified to encode for interacting proteins, protein interaction networks were constructed using the tool STRING V11 (https://string-db.org) (Szklarczyk et al., 2018). The interaction networks were constructed based on protein matchings searches in flax (Linum usitatissimum) with a minimum regulatory confidence of 0.95.

**RESULTS**

**Phenotypic Variation**

From the two shoot and 14 root targeted traits, significant variation (P < 0.05) among genotypes was detected for all except for root dry weight, network depth, and network distribution (Table 1). In multiple comparison, genotype TMP-2530 (U_MAR_C_CN98193) was significantly outperformed at least one genotype for 10 of the 13 traits (Supplementary Table 1). TMP-2530 was also the only genotype that had significantly higher shoot dry weight and, it displayed a distinctive heavy root network (Figure 1C) compared to other genotypes such as Hanley for example (Figure 1D).

Several traits were significantly correlated (Table 2). Shoot length and dry weight were strongly correlated with root dry weight and root network volume; the latter being highly correlated with root and shoot dry weights with r values of 0.78 and 0.58, respectively. East Asian genotypes appeared superior in network length (NWl) genotypes from Americas, the Middle-East, and South east Asia (Figure 2).

| TABLE 2 | Pairwise correlation of the traits. |
|----------|------------------|
| Traits   | ARD | MaxR | MedR | NWA | NWDep | NWDis | NWL | NWPer | NWSA | NWV | NWW | NWW Dep | RDWt | SDWt | SL | S_RDWt | SRL1 |
| MaxR     | -0.52** |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
| MedR     | -0.48** | 0.87* |      |      |      |      |      |      |      |      |      |      |      |      |      |
| NWA      | -0.37** | 0.75** | 0.68** |      |      |      |      |      |      |      |      |      |      |      |      |
| NWDep    | -0.07 |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
| NWDis    | 0.25 | -0.30** | -0.45** | -0.11 | 0.32** |      |      |      |      |      |      |      |      |      |      |
| NWL      | -0.48** | 0.80** | 0.72** | 0.98** | 0.34** | -0.16 |      |      |      |      |      |      |      |      |      |
| NWPer    | -0.51** | 0.84** | 0.75** | 0.98** | 0.34** | -0.16 | 0.99** |      |      |      |      |      |      |      |      |
| NWSA     | -0.37** | 0.75** | 0.67** | 0.99** | 0.36** | -0.12 | 0.98** | 0.97** |      |      |      |      |      |      |      |
| NWV      | -0.22** | 0.64** | 0.59** | 0.97** | 0.38** | -0.08 | 0.93** | 0.91** | 0.98** |      |      |      |      |      |      |
| NWW      | -0.36** | 0.63** | 0.53** | 0.82** | 0.34** | 0.05 | 0.82** | 0.82** | 0.81** | 0.77** |      |      |      |      |      |
| NWW Dep  | -0.35** | 0.62** | 0.62** | 0.65** | -0.17 | -0.15 | 0.67** | 0.66** | 0.65** | 0.59** | 0.85** |      |      |      |      |
| RDWt     | -0.17 | 0.48** | 0.43** | 0.76** | 0.33** | 0.12 | 0.72** | 0.71** | 0.76** | 0.76** | 0.61** | 0.45** |      |      |      |
| SDWt     | -0.08 | 0.44** | 0.43** | 0.62** | 0.14 | 0.05 | 0.60** | 0.59** | 0.63** | 0.63** | 0.56** | 0.50** | 0.56** |      |      |
| SL       | -0.14 | 0.27** | 0.17 | 0.41** | 0.33** | 0.24 | 0.39** | 0.40** | 0.39** | 0.39** | 0.19 | 0.45** | 0.50** |      |      |
| S_RDWt   | 0.16 | -0.21 | -0.16 | -0.35** | -0.35** | -0.04 | -0.33** | -0.33** | -0.34** | -0.33** | -0.29** | -0.10 | -0.63** | 0.21** | -0.14 |
| SRL1     | -0.80** | 0.54** | 0.51** | 0.34** | 0.13 | -0.24 | 0.47** | 0.50** | 0.32** | 0.14 | 0.38** | 0.33** | 0.12 | 0.03 | 0.15 | -0.21 |
| SRL2     | -0.4** | 0.39** | 0.35** | 0.27** | 0.02 | -0.33 | 0.35** | 0.35** | 0.20** | 0.21 | 0.26 | 0.29** | -0.37** | 0.05 | -0.08 | 0.44** | 0.454** |

*p < 0.05; **p < 0.01; ***p < 0.001.
Single Nucleotide Polymorphisms and Genetic Structure

A total of 7707 SNPs with MAF > 5% in the core collection had no missing data and, 3,243 of these had a MAF > 5% in the mini-core per se. These datasets are henceforth referred to as the 7K and 3K datasets, respectively. Phenotypic data was incomplete for 4 genotypes and additional 10 genotypes had poor SNP call; hence the following analyses were carried out on 101 accessions of the mini-core collection using the 7K and 3K datasets independently. Population structure analysis based on both SNP datasets clustered the genotypes into six ancestral populations (Figure 3A) grouped as follows: Canadian cultivars (CANC), Canadian Russian (CA_RU), temperate (TEMP), Asian (ASIA), Admixture (ADM), and mini_Indian (MIND) with only two accessions (Figure 3B). Principal component analysis (PCAs) with PC1/PC2 and PC1/PC3 produced similar population structure patterns (Figures 3C, D). The NJ clustering slightly deviated from this pattern by assigning more than half of the fiber type accessions to a clade (FIB), splitting the TEMP into two clades (TEMP1 and TEMP2), distributing the CAN_RU, ADM, and MIND into the TEMP clades but with distinguishable high branch lengths of MIND. CANC and ASIA were each in a separate clade for a total of five clades and six distinguishable populations (Figure 3E).

Quantitative Trait Nucleotide-Trait Association

From the two datasets, the six multi-locus (mrMLMs) and the LFMM methods identified a grand total of 228 QTNs associated with at least one of the 16 traits, of which 33 QTNs were detected in both datasets (Supplementary Table 2). A total of 35 large effect QTNs (high PVE) with $R^2 > 5\%$ for at least one trait were discovered, of which, 15 were identified by two or more models (Table 3). Overall, 14 QTNs were significantly associated with at least one trait with the LFMM model and the stringent Bonferroni 0.05/n threshold (Supplementary Table 3). The extent of the phenotypic variations for the traits at these 15 QTNs is illustrated (Figures 4A, B). Large effect QTNs at Chr4:17242614 and Chr5:15312783 positions were consistently associated with root network depth explaining 22.7 and 19.3% of variation for this trait respectively. Both were detected by at least four of the six multi-locus models and the LFMM model where the latter was also significant in the 3K dataset using the Bonferroni correction threshold (Supplementary Table 3, Supplementary Figure 1). A QTN associated with network perimeter with the highest PVE ($R^2 = 24.20$) at Chr9:19061342 using the multilocus FASTmrMLM (Table 3, Supplementary Table 2) was also significantly associated with the same trait based the Bonferroni criterion in both dataset (Supplementary Table 3, Supplementary Figures 1 and 2).

As expected, QTN associations with multiple correlated traits were observed. For example, a QTN at Chr15:11371216 position was associated with the following six correlated root network traits: network area, NWL, network perimeter, network surface area, network volume, and network width (Supplementary Table 3). This QTN was one of the large effect QTNs detected by multiple models for all the traits in both datasets (Table 2 and Supplementary Table 2). Two QTNs (Chr6:3310382 and Chr11:8154007) were associated with shoot dry weight (SDWt), each explaining more than 15% of the phenotypic variance of the trait (Table 3). QTN Chr11:8154007 was also significant for SDWt in both datasets using the Bonferroni correction criterion (Figure 5).

Genes Linked to Quantitative Trait Nucleotides

Most loci within 100 kb up and downstream of the detected QTNs harbored genes that had previously been reported to play role(s) in root and/or shoot development in plants (Figure 6, Supplementary Table 2). The genes were primarily related to auxin efflux, nutrient transport, and plant immunity. Loci corresponding to large effect QTNs detected by multiple methods harbored genes for the growth and development of plant organs. For instance, the locus defined by QTN Chr15:11371216 harbored genes predicted to encode a lateral organ boundary (LOB) protein and a mitogen-activated protein kinase (MAPK).
Network analysis indicated that this predicted MAPK gene likely interacts with other MAPK genes typical of MAPK cascades (Supplementary Figure 3A). QTN locus Chr5:15312783, also detected by multiple methods but in this case for its association with root depth, comprised genes predicted to encode GRAS [collective name for gibberellic acid insensitive (GAI), repressor of GA1 (RGA), and Scarecrow (SCR)] transcription factors (Gao et al., 2004). This locus also harbored genes that were predicted to encode ARM repeats, a GATA-type zinc finger transcription factor family protein and YUCCA6 (YUC6). The other large effect QTN associated with root depth (Chr4:17242614) was linked to Arabidopsis orthologue genes AT5G37020 and AT5G10360 that encode auxin response factor-8 (ARF8) and sucrose synthase-6 (SUS6), respectively.

Most QTNs associated with SDWt were at loci containing genes predicted to encode pentatricopeptide repeat (PPR), photomorphogenesis, ubiquitin-related and plant defense proteins such as multi and toxic compound extrusion (MATE) (Figure 6, Supplementary Table 2). The large effect QTNs Chr6:3310382 and Chr11:8154007 associated with SDWt were

**FIGURE 3**: Population structure. (A) Estimate of the number of ancestral populations indicating six as the best fit, (B) population structure plot showing the six populations in different colors, (C) principal component analysis (PCA) plot of the first two principal components (PCs) where the percentages in parentheses represent the variance explained by the PCs, (D) PCA plot of the first and third PCs, (E) neighbor-joining (NJ) dendrogram where naming convention indicates the type (O, oil; F, fiber; U, unknown), the country of origin, the breeding status (C, cultivar; B, breeding material; L, landrace) followed by the accession name.
located in relatively high gene density regions involved in plant immunity, development processes, and plant growth regulation (Figure 6, Supplementary Table 2). Among others, a gene predicted to function as a suppressor of phytochrome A-105 (SPA3) and assumed to regulate plant growth by suppressing photomorphogenesis (Laubinger and Hoecker 2003), was duplicated at QTN locus Chr6:3310382. The SPA3 and UBA2 proteins (Figure 6, Supplementary Table 2) were predicted to interact via COP1 (Supplementary Figure 3B).

**DISCUSSION**

Root trait measurements provide essential information to facilitate varietal improvement in breeding programs to select superior genotypes especially in nutrient and moisture deficit areas (Comas et al., 2013; Ndour et al., 2017). Root traits have already been used in breeding schemes to select elite genotypes in crops such as wheat for example (Wasson et al., 2014). Applications of root phenotype-genotype association through GWAS has enabled the identification of important QTL for root traits (Hochholdinger et al., 2018) that impact shoot traits including yield (Reinert et al., 2016). Our study provides insights into phenotype-genotype associations for early root and shoot traits of flax genotypes from over 20 countries by identifying QTL and proposing plausible candidate genes for further investigations.

**Phenotypic Variation**

The significant variations observed among genotypes for most of the early root system and shoot development traits evaluated points to the genetic diversity of flax for such traits and the potential for genetic improvement. The highest and lowest morphometric values reflect the level of diversity within the gene pool, promising useful materials for improvement through breeding. The wide range of variation...
for most agronomic traits among flax genotypes (Richards and Diederichsen 2003) is well represented in the core collection (Diederichsen et al., 2013; You et al., 2017). Most genotypes in the flax core collection exhibit both fiber and oilseed features that can attributable to selection processes for dually elite ones (You et al., 2017).

Correlations between root and shoot traits reflect the notion of balance between roots and shoots referring to plants partitioning their resource allocations between the two plant parts (Davidson 1969; Garnier, 1991). A high correlation between shoot and root dry weights has also been reported in rice (Zhao et al., 2019); this is not surprising considering the role of roots in supplying nutrients to the above-ground parts. The higher correlations between root network traits such as network perimeter, network length, network surface area, and network volume with shoot traits compared to that with root depth suggest the crop’s reliance on the top layer rooting system for resource uptake (Hocking et al., 1997; Kar et al., 2007).

The inverse relationship between average root diameter and root network related traits such as total network length, perimeter, and surface area have been reported not only in flax (Soto-Cerda et al., 2019) but also in other plants such as Arabidopsis thaliana (Qian et al., 2015). The strong negative correlation between specific root length and average root diameter agrees with similar trade-offs between these traits in plants such as maize (Zhu and Lynch, 2004) and various tree species (Bauhus and Messier, 1999; Kramer-Walter et al., 2016). The superiority of East-Asian genotypes in root network length is consistent with the performance of these materials under drought condition (data not shown) that may reflect their breeding importance.
Genome Wide Association and Candidate Genes

The genetic structure of the mini-core collection used herein is in accordance with previous reports describing the extent of the genetic and agronomic trait diversity of the flax core collection and indicating its suitability for GWAS (Soto-Cerda et al., 2013). Most QTN loci harbored genes predicted to play a role in organ development. QTN Chr15:11371216 was identified based on its association with multiple root traits, its large effect, and its detection by multiple models. The LOB and MAPK orthologous genes at this locus have already been shown to contribute to root development in *Medicago trunculata* (Liu et al., 2005; Ariel et al., 2010; Han et al., 2014). LOBs are plant-specific proteins known for their involvement in lateral organ development (Shuai et al., 2002) including lateral root formation (Jeon et al., 2017). LOB proteins not only mediate a number of root and shoot development processes but can also respond to environmental stimuli (Shuai et al., 2002; Xu et al., 2016). MAPKs regulate several physiological processes of all eukaryote organisms including microbes and metazoans (Nishihama et al., 1995). In plants, they are involved in diverse developmental processes including response to abiotic stresses (Nakagami et al., 2005) but they have also been reported to play a role in regulating plant root growth via auxin signaling (Zhao et al., 2013).

The consistent association of Chr5:15312783 with root depth based on most of the models used and the occurrence of multiple genes predicted to function as GRAS, ARM, ATA, and YUC genes at its locus hint at a possible role in determining root depth. GRAS family genes play important roles in plant organ development (Bolle, 2016). The predicted GRAS gene at this locus appeared to be an orthologue of *A. thaliana*’s AT5G66770 that encodes SCR (Gao et al., 2004), thereby supporting the candidacy of this gene for the QTL. SCR and SHORTROOT (SHR) proteins of the GRAS family are known for their regulatory functions of root development in *Arabidopsis* (Benfey et al., 1993; Kamiya et al., 2003; Sbabou et al., 2010). These two proteins interact and work in tandem, i.e., SCR regulates the movement of SHR (Cui et al., 2007). *Arabidopsis* orthologues ARM (Coates et al., 2006), GATA (Behringer and Schwechheimer, 2015), and YUC (Woo et al., 2007; Cha et al., 2015) at this locus could also be important for the development of plant organs in general and roots in particular. The specific YUC6 orthologue gene at this locus was reported to have important role in drought tolerance in *Arabidopsis* (Cha et al., 2015). The large effect Chr4:17242614 QTN for root depth may be attributable to an ARF8 orthologous gene that can regulate auxin-mediated process and influence primary root elongation (Biliou et al., 2005) and possibly lateral roots as well (Wang et al., 2015; Lee et al., 2019). ARF8 has been demonstrated to have clear effect on root growth habit where the roots of wild type plants grew slanted and those of mutants had a vertical downward elongation (Tian et al., 2004). This gene can also regulate flower maturation in later developmental
FIGURE 6 | Physical map of the 15 chromosomes of flax illustrating the position of the quantitative trait nucleotides (QTNs) and their proximal candidate genes. QTNs with \( R^2 > 5\% \) in the 3K, 7K or in both datasets are in green, purple, and red, respectively. QTNs indicated in blue have \( R^2 < 5\% \) and were detected in the 3K dataset. Numbers on the left of the chromosomes represent physical distances in megabases.
stage, affecting fertility and seed production (Nagpal et al., 2005; Ghelli et al., 2018). The existence of a SUS orthologous gene as a regulator of primary root development (Sturm et al., 1995) may also be entertained as a candidate QTL for flax root depth.

The consistent occurrence of PPR, MATE, and ubiquitin related orthologous genes at large effect QTN loci associated with shoot traits, especially with SDWt, make them strong candidates. The PPR genes are involved in plant growth and stress tolerance in several plant species (Laluk et al., 2011; Wu et al., 2016; Xing et al., 2018). MATE efflux proteins on the other hand play a vital role in plant immunity against different toxins (Diener et al., 2001) including secondary metabolites (Gomez et al., 2009), xenobiotics such as heavy metals (Li et al., 2002) and aluminum (Li et al., 2017). Ubiquitin orthologues linked to SDWt QTNs are candidate loci on the ground of their known roles in shoot development (Stirnberg et al., 2002; Yang et al., 2007). The SPA3 orthologous gene at the Chr6:3310382 locus, the QTN associated with SDWt, is a member of a gene family that acts as a suppressor of phytochrome A and regulates morphogenesis (Hoecker et al., 1999; Laubinger and Hoecker, 2003). In Arabidopsis, SPA3 has a pronounced effect on seedling elongation (Laubinger and Hoecker, 2003) and, interestingly, it is expressed in all above ground tissues while showing no detectable expression in roots (Zhu et al., 2008). The SPA3 protein in Arabidopsis and rice has a conserved DWD motif that reflects its role in modulating a protein involved in photomorphogenesis repression and in activation of etiolation through CONSTITUTIVELY PHOTOMORPHOGENIC1 (COP1); the latter also possesses E3 ubiquitin ligase activity (Lee et al., 2008). A secondary role for SPA3 that would be related to UBA at this locus through COP1 can therefore be considered (Supplementary Figure 3).

Most genes linked to the large effect QTNs associated with root and shoot traits are responsive to abiotic stresses such as drought, salt, and temperature. For instance, some GRAS family genes have useful roles in drought and salt tolerance (Ma et al., 2010; Xu et al., 2015). Some members of MAPK genes positively regulate low-temperature tolerance while decreasing drought and salt resistance (Jia et al., 2016). The role played by auxin-related genes (e.g., LOBs) in various abiotic stresses, including drought, have been reported for many plant species (Jain and Khurana, 2009; Wang et al., 2010; Jung et al., 2015; Huang et al., 2016). The QTN trait associations in these studies may further implicate variation in local adaptation to different environmental conditions given that our germplasm represented all flax growing regions of the world and represented more than 95% of the genetic diversity of whole flax core collection.

CONCLUSIONS

Early root and shoot trait phenotyping and GWAS of flax have provided insights into the complex relationships of these traits and their associated QTN loci were mined to hypothesize candidate genes. Early root network traits are interrelated, positively impact shoot traits and, consequently, seedling vigor. As such, early root establishment may also affect downstream yield performance. Our results suggest that, in flax, the extent of the root network is more important than root depth per se during the early growth stages. Root development studies spanning all growth stages would be necessary to quantify the relative importance of both traits over the entire cropping season and under different moisture regimes.

The GWAS yielded QTNs associated with most of the root traits and both shoot traits and the candidate genes identified at the major loci provide grounds for further investigations, particularly as they relate to stress tolerance. Some of the QTNs have pleiotropic effects that can either stem from linked genetic features at the loci or from single genes affecting multiple root traits. SDWt associated loci harbored genes that regulates physiological processes in above ground plant parts such as photosynthesis. Genes expressed at different stages and in different tissues may be tested for their specific role in controlling agronomic traits or imparting stress tolerance. Therefore, some of these loci co-locate with QTL for other traits not measured herein. Given the polygenic nature of several agronomic traits, consideration must be give to QTL of small effects because their cumulative impact is important in pre-breeding and positive selection can be achieved through genomic selection and other marker-assisted breeding schemes.

DATA AVAILABILITY STATEMENT

This manuscript contains previously unpublished data. The name of the repository and accession number(s) are not available.

AUTHOR CONTRIBUTIONS

DS and SC conceived the project. DS, SC, and SR participated in designing and implementing the phenotyping experiment. SC and FY made the genetic data available. DS performed the phenotypic and GWAS data analysis. DS prepared the manuscript. SC, FY, and SR worked on revision of the manuscript.

FUNDING

This Agri-Innovation Project (J-000279) was funded by the Flax Council of Canada and Agriculture and Agri-Food Canada.

ACKNOWLEDGMENTS

We would like to thank Alain Saumure for help in designing and installation of the phenotyping system. Assistance of several individuals from the Cloutier lab during phenotyping is acknowledged. A special thanks to Shanise Reddekopp for daily monitoring of the experiments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01483/full#supplementary-material
Hocking, P., Kirkegaard, J., Angus, J., Gibson, A., and Koetz, E. (1997). Comparison of canola, Indian mustard and Linola in two contrasting environments. I. Effects of nitrogen fertilizer on dry-matter production, seed yield and seed quality. *Field Crops Res.* 49, 107–125. doi: 10.1016/S0378-4290(96)01603-5

Hodge, A. (2004). The plastic plant: root responses to heterogeneous supplies of nutrients. *New Phytol.* 162, 9–24. doi: 10.1111/j.1469-8137.2004.01015.x

Hoecker, U., Tepperman, J. M., and Quail, P. H. (1999). SPA1, a WD-repeat protein specific to phytochrome A signal transduction. *Science* 284, 496–499. doi: 10.1126/science.284.5431.496

Huang, X., Bao, Y., Wang, B., Liu, L., Chen, J., Dai, L., et al. (2016). Identification and expression of Aux/IAA, ARF, and LBD family transcription factors in *Boehmeria nivea*. *Biol. Plant* 60, 245–250. doi: 10.1007/s10535-016-0584-8

Hund, A., Trachsel, S., and Stamp, P. (2009). Growth of axial and lateral roots of maize: development of a phenotyping platform. *Plant Soil* 325, 335–349. doi: 10.1007/s11104-009-9884-2

Jackson, R., Manwaring, I., and Caldwell, M. (1990). Rapid physiological adjustment of roots to localized soil enrichment. *Nature* 344, 58. doi: 10.1038/344058a0

Jain, M., and Khurana, P. J. (2009). Transcript profiling reveals diverse roles of auxin-responsive genes during reproductive development and abiotic stress in rice. *FEBs J.* 276, 3148–3162. doi: 10.1111/j.1746-4689.2009.07033.x

Jeon, E., Young Kang, N., Cho, C., Joon Seo, P., Chung Soo, M., and Kim, J. (2017). Identification and expression of Aux/IAA, ARF, and LBD family transcription factors in *Boehmeria nivea*. *Biol. Plant* 60, 245–250. doi: 10.1007/s10535-016-0584-8

Kramer-Walter, K. R., Bellingham, P. J., Millar, T. R., Smissen, R. D., Richardson, D. R., and Donald, S. L., I., and Laughlin, D. C. (2016). Root traits are multidimensional: specific root length is independent from root tissue density and the plant economic coefficient of dry season oilseed crops. *Agric. Water Manage.* 152–167. doi: 10.1007/s11104-009-9984-2

Li, L., He, Z., Pandey, G. K., Tsuchiya, T., and Luan, S. (2002). Functional cloning and characterization of a plant efflux carrier for multidrug and heavy metal detoxification. *J. Biol. Chem.* 277, 5360–5368. doi: 10.1074/jbc.M108777200

Li, N., Meng, H., Xing, H., Liang, L., Xiao, Y., and Luo, K. (2017). Genome-wide analysis of MATE transporters and molecular characterization of aluminum resistance in *Populus*. *J. Exp. Bot.* 68, 5669–5683. doi: 10.1093/jxb/erx370

Liu, H., Wang, S., Yu, X., Yu, J., He, X., Zhang, S., et al. (2005). AR1L, a LBD-domain protein required for adventitious root formation in rice. *Plant J.* 43, 47–56. doi: 10.1111/j.1365-313X.2005.02434.x

Liu, L., Gan, Y., Bueckert, R., and Van Rees, K. (2011). Rooting system of oilseed and pulse crops I: Temporal growth pattern across the plant developmental periods. *Field Crops Res.* 122, 256–263. doi: 10.1016/j.fcr.2011.04.002

Ma, H.-S., Liang, D., Shuai, P., Xia, X.-L., and Yin, W.-L. (2010). The salt-and-drought-inducible poplar GRAS protein *SCL7* confers salt and drought tolerance in *Arabidopsis thaliana*. *J. Exp. Bot.* 61, 4011–4019. doi: 10.1093/jxb/ern217

Morrillon, R., and Lasalles, J.-P. (2002). Water deficit during root development: effects on the growth of roots and osmotic water permeability of isolated root protoplasts. *Planta* 214, 392–399. doi: 10.1007/s004250100626

Nagpal, P., Ellis, C. M., Weber, H., Ploense, S. E., Barkawi, L. S., Guilleford, T. J., et al. (2005). Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. *Development* 132, 4107–4118. doi: 10.1242/dev.01935

Nakamura, H., Pitschke, A., and Hirt, H. (2005). Emerging MAP kinase pathways in plant stress signalling. *Trends Plant Sci.* 10, 339–346. doi: 10.1016/j.tplants.2005.05.009

Ndour, A., Vadez, V., Pradal, C., and Lucas, M. (2017). Virtual plants need water too: Functional-structural root system models in the context of drought tolerance breeding. *Front. Plant Sci.* 8, 1577. doi: 10.3389/fpls.2017.01577

Nishihama, R., Banno, H., Shibata, W., Hirano, K., Nakashima, M., Usami, S., et al. (1995). Plant homologues of components of MAPK (mitogen-activated protein kinase) signal pathways in yeast and animal cells. *Plant Cell Physiol.* 36, 749–757. doi: 10.1093/oxfordjournals.pcp.a078818

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–958.

Qian, H., Lu, H., Ding, H., Lavoie, M., Li, Y., Liu, W., et al. (2015). Analyzing *Arabidopsis thaliana* root proteome provides insights into the molecular bases of endoselositive imazethapyryl toxicity. *Sci. Rep.* 5, 11975. doi: 10.1038/srep11975

Reinert, S., Kortz, A., Léon, I., and Naz, A. A. (2016). Genome-wide association mapping in the global diversity set reveals new QTL controlling root system and related shoot variability in barley. *Front. Plant Sci.* 7, 1061. doi: 10.3389/fpls.2016.01061

Richards, K., and Diederichsen, A. (2003). Cultivated flax and the genus *Linum: a review*. *Flax. CRC Press* 172, 43, 56. doi: 10.1093/jxb/erx370

Soto-Cerda, B. J., Diederichsen, A., Ragupathy, R., and Cloutier, S. (2013). Genetic characterization of a core collection of flax (*Linum usitatissimum L*) suitable for association mapping studies and evidence of divergent selection between fiber and linseed types. *Plant J.* 74, 9–24. doi: 10.1111/1365-313X.12562

Sudre, D., You, F. M., Ravichandran, S., and Cloutier, S. (2019). The genetic structure of flax illustrates environmental and anthropogenic selections that gave rise to its eco-geographical adaptation. *Mol. Phylogenet. Evol.* 137, 22–32. doi: 10.1016/j.ympev.2019.04.010

Shuai, B., Reynaga-Pena, C. G., and Springer, P. S. (2002). The lateral organ boundaries gene defines a novel, plant-specific gene family. *Plant Physiol.* 129, 747–761. doi: 10.1104/pp.101.00926

Singh, K. K., Mirudla, D., Rehal, J., and Barmwal, P. (2011). *Flaxseed*: a potential source of food, feed and fiber. *Crit. Rev. Food Sci. Nutr.* 51, 210–222. doi: 10.1080/01605952.2010.478958

Soto-Cerda, B. J., Cloutier, S., Gajardo, H. A., Aravena, G., and Quian, R. (2019). Identifying drought- resilient flax genotypes and related-candidate genes...
