



NRC Publications Archive Archives des publications du CNRC

The LuWD40-1 gene encoding WD repeat protein regulates growth and pollen viability in flax (*Linum Usitatissimum* L.)

Kumar, Santosh; Jordan, Mark C.; Datla, Raju; Cloutier, Sylvie

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. / La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

For the publisher's version, please access the DOI link below. / Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

Publisher's version / Version de l'éditeur:

<https://doi.org/10.1371/journal.pone.0069124>

PLoS ONE, 8, 7, pp. e69124:1-e69124:10, 2013-07-30

NRC Publications Record / Notice d'Archives des publications de CNRC:

<https://nrc-publications.canada.ca/eng/view/object/?id=950609e5-505d-482f-9b94-ff51d2191f91>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=950609e5-505d-482f-9b94-ff51d2191f91>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.



The *LuWD40-1* Gene Encoding WD Repeat Protein Regulates Growth and Pollen Viability in Flax (*Linum usitatissimum* L.)

Santosh Kumar^{1,2}, Mark C. Jordan², Raju Datla³, Sylvie Cloutier^{1,2*}

1 Department of Plant Science, University of Manitoba, Winnipeg, Manitoba, Canada, **2** Cereal Research Centre, Agriculture and Agri-Food Canada, Winnipeg, Manitoba, Canada, **3** National Research Council, Saskatoon, Saskatchewan, Canada

Abstract

As a crop, flax holds significant commercial value for its omega-3 rich oilseeds and stem fibres. Canada is the largest producer of linseed but there exists scope for significant yield improvements. Implementation of mechanisms such as male sterility can permit the development of hybrids to assist in achieving this goal. Temperature sensitive male sterility has been reported in flax but the leakiness of this system in field conditions limits the production of quality hybrid seeds. Here, we characterized a 2,588 bp transcript differentially expressed in male sterile lines of flax. The twelve intron gene predicted to encode a 368 amino acid protein has five WD40 repeats which, *in silico*, form a propeller structure with putative nucleic acid and histone binding capabilities. The LuWD40-1 protein localized to the nucleus and its expression increased during the transition and continued through the vegetative stages (seed, etiolated seedling, stem) while the transcript levels declined during reproductive development (ovary, anthers) and embryonic morphogenesis of male fertile plants. Knockout lines for *LuWD40-1* in flax failed to develop shoots while overexpression lines showed delayed growth phenotype and were male sterile. The non-viable flowers failed to open and the pollen grains from these flowers were empty. Three independent transgenic lines overexpressing the *LuWD40-1* gene had ~80% non-viable pollen, reduced branching, delayed flowering and maturity compared to male fertile genotypes. The present study provides new insights into a male sterility mechanism present in flax.

Citation: Kumar S, Jordan MC, Datla R, Cloutier S (2013) The *LuWD40-1* Gene Encoding WD Repeat Protein Regulates Growth and Pollen Viability in Flax (*Linum usitatissimum* L.). PLoS ONE 8(7): e69124. doi:10.1371/journal.pone.0069124

Editor: Randall P. Niedz, United States Department of Agriculture, United States of America

Received: April 11, 2013; **Accepted:** June 11, 2013; **Published:** July 30, 2013

Copyright: © 2013 Kumar et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The project was funded through Agriculture and Agri-Food Canada A-base grant No. 1785. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: Sylvie.J.Cloutier@agr.gc.ca

Introduction

Male sterility in flax (*Linum usitatissimum* L.) was first documented in the early nineteen hundreds from F_2 of a cross between a blue flower oilseed type and a white flower fibre type where both parents were hermaphrodite with no male sterility [1]. The anthers in these male sterile F_2 lines sometimes developed to a stage where little viable pollen was produced suggesting an incomplete sterility mechanism. The male sterility phenotype was heritable because mostly male sterile plants gave rise to plants exhibiting the same phenotype. The sterile F_2 plants also had smaller petals that seldom opened. In a reciprocal cross with the blue flower oilseed type as the female parent, one out of four F_2 plants were male sterile [1].

In the present study we observed the occurrence of male sterility in an F_2 cross between oilseed type flax accession Double Low and cultivar AC McDuff. Both Double Low and AC McDuff are fertile parents. A gene with homology to EST LuP1225D10 (EB713752) encoding a tryptophan-aspartate (WD) repeat protein from AC McDuff was found to be expressed exclusively in temperature sensitive male sterile plants of flax [2]. Hence, we decided to characterize this gene in order to determine its role in male sterility.

WD proteins are present in bacteria such as *Thermomonospora curvata* [3] and *Cyanobacterium synechocystis* [4] but are more prevalent in eukaryotes. The earlier hypothesis suggested intragenic duplication and recombination events as the source of WD-repeats [5] but recent evidence hints at divergence of members of this protein family at different time points during evolution [6]. The WD proteins contain varying numbers of WD repeats and act as protein-protein and protein-nucleic acid interaction domain. The WD domain consists of 40 amino acids with a glycine-histidine (GH) di-peptide near the N-terminal, a conserved aspartic acid located before the WD repeats and the signature WD di-peptide at the C-terminal end [7,8]. The first identified, and so far the best characterized WD protein, is the β -subunit of heterotrimeric G protein [9–11]. Based on crystal structure studies of G proteins, the WD proteins acquire a highly symmetrical β -propeller fold structure with each repeat containing a small four-stranded β sheet [10,11]. A single β -propeller may contain 4 to 8 repeats but 7 or 8 repeat propellers are the most common [12]. The large β -propeller of WD40 proteins are composed of ~300 amino acids which can interact through the top, the bottom and the circumference of the propeller [6]. Studies in yeast interactome demonstrated that WD40 proteins are involved in more protein-protein interactions than any other domains [13,14]. The

functional versatility of WD40 proteins is owed to their ability, (i) to target different substrates selectively similar to F-box proteins [15] (ii) to recruit different substrates in binding modes similar to the peptide-in-groove binding of clathrin [16] or by distinct binding modes through the top and side of the WD domain as in the G_β of the G proteins [11] (iii) to interact through insertion motifs as in MAD3 protein [17] or inter-blade binding grooves of WD40 domains for ligand binding as in Sro7 protein [18].

The WD40 proteins are involved in diverse cellular functions including cell division [19], cytoskeletal organization [20,21], vesicle formation and trafficking [22] and transcriptional regulation [23,24]. WD40 protein mediated transcriptional regulation by histone recognition [25] and arrangement of histone tails for the spreading of chromatin [26] impart silencing and thus has epigenetic consequences in regulating growth and development. WD40 protein malfunctions can lead to disease development [5]. Although much of the earlier evidence on the role of WD40 proteins came from animal studies, research in the model species *Arabidopsis* has advanced the understanding of their diverse functions in plants [27,28]. Studies in *Arabidopsis* and rice revealed the presence of 85 and 78 WD40 proteins, respectively [29]. The majority of the WD repeat containing genes discovered have crucial functions in plants such as the F-box protein transport inhibitor response1 which is an auxin hormone receptor [30]. The *fertilization-independent endosperm* gene regulates flowering and seed dormancy [31]. The WD40 domain containing G protein coupled receptors play a vital role as hormone receptors and signalling components [32]. Cyclophilin71 interaction with histone H3 is essential for chromatin-based gene silencing affecting organogenesis [33]. A WD40 protein encoding gene, the promoter of which interacts with the dehydration response element, may be involved in stress tolerance in foxtail millet [34]. The *Arabidopsis* Yaozhe (YAO) is a nucleolar WD40 protein involved in embryogenesis and gametogenesis [35]. Rice Immature Pollen1 containing five WD40 repeats regulates late pollen development in rice [36]. *OsLIS-L1* gene encoding the Lissencephaly Type-1-Like protein is involved in plant height and male gametophyte formation in rice [37].

In this study, we report on the cloning and characterization of a novel WD40 protein encoding gene *LuWD40-1* corresponding to the EST LuP1225D10, a homologue of which was initially identified to be expressed exclusively in Chinese male sterile flax accession [2].

Materials and Methods

Gene sequence analysis

Clone LuP1225D10 (EB713752), obtained from a cDNA library constructed from 12 days after flowering flax bolls [38], was completely sequenced. Extended promoter and 5'UTR sequences were obtained from the assembly of AC McDuff short Illumina reads [39] mapped onto the whole genome shotgun (WGS) sequence of flax cv CDC Bethune (LinUsi v1.1, NCBI genome project #68161) [40]. The transcription start site (TSS) was identified using the bioinformatics pipeline for TSS signal analysis (http://fruitfly.org/seq_tools/promoter.html) [41]. Open reading frame (ORF) prediction was done using the ORF Finder tool from NCBI and the DNAMAN software (<http://www.lynnon.com/>). Promoter analysis was performed with PLant Cis-acting regulatory DNA Elements (PLACE) [42], PLANT Promoter Analysis Navigator (PlantPAN) and Weight Matrix patterns of PLant regulatory sequences (ScanWM-PL) available on the Softberry web portal (<http://linux1.softberry.com/berry.phtml?topic=scanwmp&group=programs&subgroup=promoter>). Sequence homologies were

calculated with ClustalW2 [43] and a dendrogram was constructed using the neighbour-joining method implemented in the Molecular Evolutionary Genetics Analysis (MEGA) software version 5.1 [44].

RNA-Seq and data normalization

Tissue samples of flax cv CDC Bethune were collected from embryos (globular, heart, torpedo, mature), seeds, etiolated seedlings, stems, ovaries and anthers. Total RNA was isolated from each of these tissue samples using RNeasy® Plant Mini Kit (Qiagen Inc., Mississauga, Ontario, Canada) to construct the nine tissue-specific libraries. RNA-Seq was performed on the Illumina platform (Genome Analyzer II) at the National Research Council (Saskatoon, SK, Canada). Paired-end tagged (PET) sequence reads were generated for each library using one sample per lane. The PET RNA sequences from each tissue were aligned to the WGS sequence of flax [40] using Tophat [45] with default settings. Transcript assembly and transcript abundance were analysed using the cuffdiff modules of the Cufflinks package with default settings [46,47]. To calculate expression levels from read counts, the Cufflinks package was used to count the reads that map to a transcript and then to normalize the mapped read counts to the length of the transcript. In addition, to compare expression of transcripts across runs, the read counts were normalised for the total sequence yield of the machine [46]. These two normalization steps were implemented by calculating the Fragments Per Kilobase of transcripts per Million mapped reads (FPKM) [46] to reduce variability arising from assembly of PET reads, length of transcript, RNA composition and variations in library preparation. The normalized RNA-Seq dataset is available through the Total Utilization Flax GENomics (TUFGEN) project website (www.linum.ca) and can be downloaded from <http://linum.ca/downloads/RNAseq>.

Constructs for plant transformation

Cloning was carried out using the Gateway system (Invitrogen, Carlsbad, CA, USA). Forward (5'-ATGCGCATGGACGCGACGAAC-3') and reverse primers with sequences encoding the human influenza Hemagglutinin tag (5'-TCAAGCGTAGTCTGGACGTCGTATGGGTACCCCGGCTGCAGCG-3') were designed to amplify the 1,107 bp full length cDNA of the gene. An additional reverse primer R₁ (5'-CATATGCTGGTCCGT-CATGGC-3') was designed at the 5' end spanning introns and exons to distinguish the amplification product arising from cDNA or genomic DNA templates and was used in combination with the forward primer to amplify a 274 bp fragment (Fig. S1A) for generating the RNAi construct, for gene amplification in confirmation of transgenic lines and for semi-quantitative RT-PCR experiments. The full length gene without a stop codon was amplified using the same forward primer and the following reverse primer (5'-CCCCCGGCTGCAGCGGAAT-3') and cloned into the CD3-685 vector (pEarleyGate103, ABRC) containing a GFP tag at the C-terminal end. Cloning and shuttling were performed as per manufacturer's instructions (Gateway® Cloning Technology, Invitrogen, Carlsbad, CA, USA). The PCR amplified gene fragments were cloned into pDONR221 entry clone and positive transformants were sequenced to confirm their accuracy. The gene constructs were then incorporated into pIPKb004 binary destination vector for overexpression (OE) and pIPKb009 for RNAi based knockout (KO) constructs [48]. The transgenic flax lines were tested for the presence of the transgene using a CaMV 35S promoter forward primer (5'-GATGACGCACAATCCCCTACTT-3') [48] and the gene specific internal reverse primer R₁. The following forward (5'-TAGAGCTGACCAGGACAAACA-3') and reverse

(5'-GTTTATGAATGCGCTTGTCTCA-3') primers were used to amplify the house-keeping gene adenine phosphoribosyl-transferase1 (*apt1*) [49] that served as the internal control for semi-quantitative RT-PCR [50].

Transformation

Flax cultivar Prairie Grande was transformed according to the protocol of Wijayanto and McHughen [51] with the following modifications. The sterilized seeds were grown in the dark on Murashige and Skoog (MS) medium with 2% sucrose and 0.4% phytigel (MS2P). Sets of thirty hypocotyl segments per plate in multiple replicates were maintained at 24°C under a 16 h light and 8 h dark cycle for 4 days before bombardment. Plasmid DNA (5 µl at 1 µg/µl) was precipitated on sterilized 0.6 micron gold particles using CaCl₂ and protamine as previously described [52]. The particle bombardment of 5 µg plasmid DNA, per plate at 650 psi pressure and a distance of 6 cm was done using a biolistic PDS-1000/He delivery system (Bio-Rad Laboratories, Mississauga, Ontario, Canada). The bombarded hypocotyls were maintained for an additional 4 days in the same conditions as the pre-culture before transferring them to MS2P medium, containing 5 mg/l hygromycin for 4 weeks followed by subculture with 20 mg/l hygromycin for 2–3 months until regeneration of shoots. Regenerating shoots were excised and transferred to Magenta jars containing 5 mg/l hygromycin in MS2P medium. The shoots that survived and formed roots were transferred to soil and grown in growth cabinets at 18°C under a 16 h light and 8 h dark cycle. The putative transformed flax plants were allowed to mature, flower, self-pollinate and set seeds.

For subcellular localization studies, onion epidermal cells were transformed with the GFP-tagged *LuWD40-1* gene using the same particle bombardment conditions as for the flax hypocotyls. The transfected epidermal layers were incubated in the dark for 48 h before microscopy.

PCR and semi-quantitative RT-PCR

DNA was extracted from 10 mg lyophilised leaf tissue of the untransformed Prairie Grande and the transgenic lines using the QiagenDNeasy 96 plant kit (Qiagen Sciences, Maryland, USA). RNeasy® Plant Mini Kit (Qiagen Inc., Mississauga, Ontario, Canada) was used for RNA extraction from leaf tissue of the same genotypes. Total RNA was treated with TURBO DNA-freeDNase (Ambion, Austin, Texas, USA) and cDNA synthesis was performed using Superscript™ II reverse transcriptase (Invitrogen, Carlsbad, California, USA) followed by RNase H (Invitrogen, Carlsbad, California, USA) treatment. The cDNA was quantified using Quanti-iT™ Ribogreen RNA reagent kit (Molecular probes, Eugene, Oregon, USA) [53]. Approximately 30 ng of genomic DNA extracted from leaf tissues of the transgenic lines was used as template to confirm the presence of the transgene construct. Semi-quantitative RT-PCR was performed with 4 ng cDNA from the leaf tissue of transgenic lines using the forward and R₁ primers described above. PCR of DNA and cDNA samples were performed with melting temperature set to 62°C for 36 cycles and 28 cycles for PCR and semi-quantitative RT-PCR, respectively. Semi-quantitative RT-PCR is based on the relative expression measured during the linear phase of the reaction [54]. We performed amplification of the *LuWD40-1* target and the *apt1* control at 28 and 31 cycles and compared the intensity ratios by densitometry measurement using the AlphaImagerHP software (version 3.4, proteinsimple, Santa Clara, CA, USA). At 28 cycles, variations in gene to control ratios were obtained between the untransformed Prairie Grande and the transgenic lines, however, at 31 cycles, the gene to control ratios suggested near saturation of

amplification (Fig. S2). The semi-quantitative RT-PCR experiment was repeated three times at 28 cycles. All protocols were performed according to the manufacturer's instructions.

Microscopy

The flax pollen cells from the untransformed Prairie Grande and the transgenic lines stained with 2,5-diphenyl tetrazolium bromide (MTT) (Sigma, Oakville, Ontario, Canada) [55] were counted at 50× magnification on a Wild Heerbrugg M8 stereo microscope (Wild Heerbrugg, Heerbrugg, Switzerland) equipped with a fiber optic light source and illuminator cables. The transformed onion epidermal cells expressing GFP were visualized through a Zeiss AxioScope.A1 microscope (Carl Zeiss, Oberkochen, Germany) using EC Plan-NEOFLUAR lenses under 20× and 40× magnifications. The images were captured through an AxioCam ICc 1 camera mounted on the microscope and the Axio Vision software (Release 4.8.2) was used to process the images.

Results

Sequence analysis of *LuWD40-1*

A study on dominant genic male sterile flax identified a transcript that had 89% sequence identity with LuP1225D10 and was expressed exclusively in male sterile lines of Chinese flax accession [2]. The EST clone LuP1225D10 was generated from AC McDuff which is a fertile plant. However, when AC McDuff is crossed to another fertile parent, Double Low, the progeny shows male sterility which does not follow Mendelian inheritance. Here, we presented the molecular and functional characterization of the gene.

Sequencing of the cDNA clone LuP1225D10 generated a fragment of 1,638 bp including an 84 bp 5' UTR, a 1,107 bp ORF and a 447 bp 3'UTR ending in a polyA tail. This sequence was submitted to Genbank as *LuWD40-1* (KC686833). The 5'UTR region extended using the AC McDuff Illumina short reads [39] confirmed that LuP1225D10 contained the full length ORF of a WD40 gene. The promoter sequences of CDC Bethune and AC McDuff were identical but the coding sequence of the CDC Bethune homologue *Lus10011938* was predicted to be 6 bp shorter than that of LuP1225D10. The *Lus10011938* sequence was 2,588 bp long with a 1,101 bp predicted ORF. This shorter predicted ORF for CDC Bethune was the result of an error in the prediction of the true start codon of the *ab initio* prediction and RNA-Seq data (Datla, unpublished) confirmed that the AC McDuff and CDC Bethune ORFs were both 1,107 bp in length. The 5'UTR of *LuWD40-1* was 1,034 bp with a TATA box located 30 bp upstream of the TSS. The complete transcript including the extended 5' UTR sequence of *LuWD40-1* (2,588 bp) with its deduced amino acid sequence is shown (Fig. S1A). The genomic clone corresponding to the ORF of *LuWD40-1* from the flax WGS sequence is 4,471 bp and is composed of thirteen exons and twelve introns (Fig. S1B & S1C).

The *LuWD40-1* gene encodes a protein of 368 amino acids with a theoretical isoelectric point (pI) of 5.66 and an estimated average molecular mass of 39.7 kDa (Compute pI/Mw tool, www.expasy.org). A BLASTP search of the deduced amino acid sequence against the non-redundant protein database of NCBI showed high similarity to WD40 proteins such as the putative dead box ATP-dependent RNA helicase from *Ricinus communis* (E-value 4e-158; XP_002514133), the *Arabidopsis lyrata* transducin family protein (XP_002886042.1), DWA-1 like protein from *Cucumis sativus* and the *Arabidopsis thaliana* THO complex subunit 6 (NP_849989.1) proteins. Various unnamed proteins with WD40 domains but unknown functions from *Vitis* (CBI27884.3), *Glycine* (ACU18349.1)

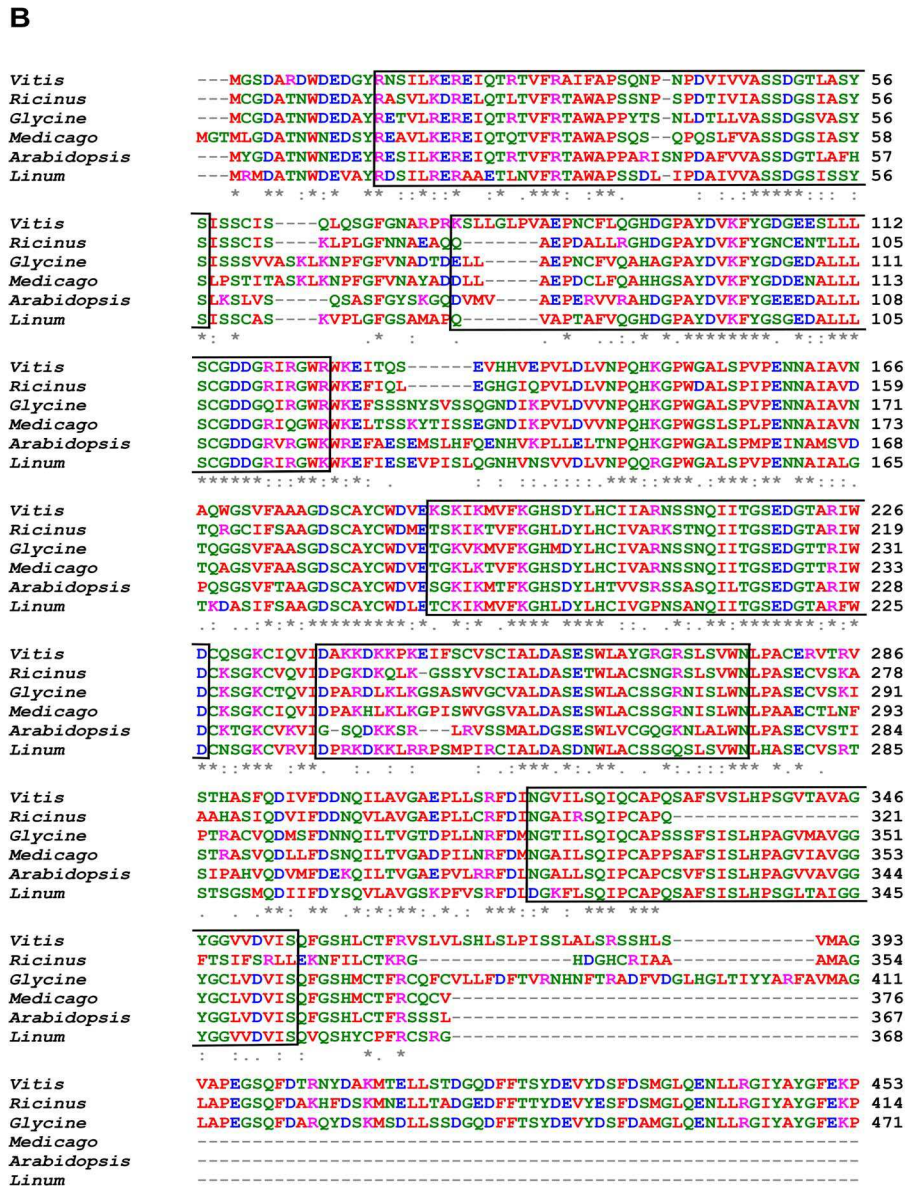
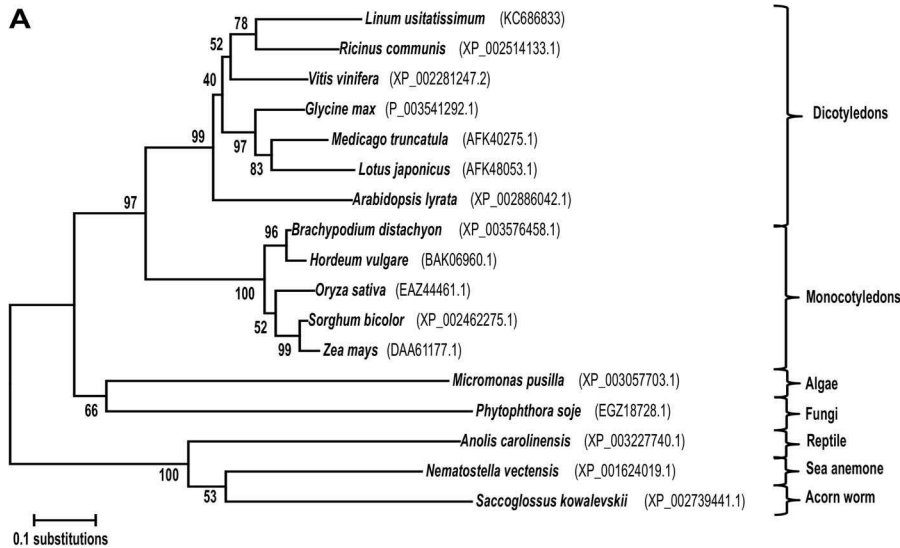


Figure 1. Phylogenetic and structural analysis of WD40 proteins. (A) An amino acid sequence based phylogenetic tree of WD40 proteins from various organisms. The phylogenetic tree was generated by the MEGA 5 software using the neighbour-joining algorithm, and bootstrap values based on 1000 replications are represented at the branch points. The relative amount of change is represented by the scale bar. (B) Deduced amino acid sequence alignment of WD40 repeat proteins from dicot species was generated by ClustalW2. WD40 repeats are shown in boxes. Sequences from other species that extended beyond *Linum* sequence at the 3' end and did not contain any WD40 domains were trimmed. doi:10.1371/journal.pone.0069124.g001

and *Medicago* (AFK40275.1) showed similarity ranging between 60–65%. The BLAST results also included putative orthologs from algae, fungi, reptiles and worms, all of which had WD40 domains. A protein sequence based phylogenetic tree showed clear divergence between WD40 proteins from the plant kingdom compared to species from other kingdoms (Figure 1A). Within plants, monocot and dicot WD40 proteins displayed evolutionary divergence. The flax WD40 protein was most closely related to the protein from castor bean (Figure 1A). A Simple Modular Architecture Research Tool (SMART) analysis of the LuWD40-1 encoded protein (<http://smart.embl-heidelberg.de/>) identified five WD40 domains of 38 to 43 amino acids residues in the dicot orthologs (Figure 1B). Variation in the protein sequence and indels were present, even within the WD40 domains (Figure 1B). The *lotus japonicus* sequence (AFK40275.1) was not included in the alignment because it was partial.

An *ab initio* analysis of LuWD40-1 secondary structure was performed on the I-TASSER server [56]. The model with the highest C-score (−0.78) is shown (Figure 2A). The model was confirmed through the SWISS-MODEL workspace [57] with a QMEAN Z-Score of −5.89. LuWD40-1 with five WD40 repeats formed a seven blade propeller structure similar to proteins with seven WD repeats. Other WD40 domain containing proteins of the I-TASSER database also possessed nucleotide or histone binding characteristics such as the transcriptional repressor TUP1 [58] and the histone binding pRB associated protein p46 (RBBP7) [59] (Figure 2B). LuWD40-1 is predicted to have a histone binding site identical to the histone binding protein RBBP7.

In silico promoter analysis and transcript profile of LuWD40-1 in flax tissues

The LuWD40-1 sequence of AC McDuff is identical to the predicted gene *Lus10011938* from CDC Bethune. A 1,500 bp sequence upstream of the TSS of LuWD40-1 and *Lus10011938* was extracted from the assembly of AC McDuff and CDC Bethune genome sequences and signal scan searches were

performed. The list of reported signals is compiled (Table S1). Signal sequences for induction of gene expression in stems and roots and repression in leaves (E2FCONSensus) were identified. Several elements for light and stress responsive activation of genes were also present. Multiple copies of GTGANTG10 and POLLEN1LELAT52 known to regulate pollen specific expression were detected. The tissue specific transcript profiling based on normalised number of transcripts from tissue specific RNA-Seq data showed results consistent with the predicted promoter signal scan results (Figure 3). The *Lus10011938/LuWD40-1* transcripts had relatively lower expression at all stages of embryo development but were abundant in seed, etiolated seedling, stem, ovary and anther tissues.

Cellular localization of LuWD40-1 in onion epidermal cells

The LuWD40-1 sequence was analysed through Subnuclear Compartments Prediction System (<http://array.bioengr.uic.edu/subnuclear.htm>) which suggested its localization in nuclear lamina. To confirm this prediction, onion epidermal cells were transformed with a *LuWD40-1-GFP* construct and an empty vector-GFP construct as control. Diffused green fluorescence signal was throughout the cells bombarded with the empty vector-GFP construct (Figure 4 A–B) but the GFP signal was localised exclusively in the nucleus in cells transiently expressing the *LuWD40-1-GFP* construct (Figure 4C–D).

Phenotypic analysis of LuWD40-1 transformed lines in flax

The flax cultivar Prairie Grande was transformed with *LuWD40-1* constructs to develop OE and KO lines. A total of 1,770 and 2,940 hypocotyls were shot with the *LuWD40-1* OE and RNAi based KO constructs, respectively. None of the calli transformed with the *LuWD40-1* KO construct produced shoots, even after ten rounds of selection on hygromycin. Three out of the four T₁ *LuWD40-1* OE lines that produced shoots were positive for presence of the transgene (Fig. S3) and were propagated through selfing to produce T₂ lines. The OE T₁ lines had low

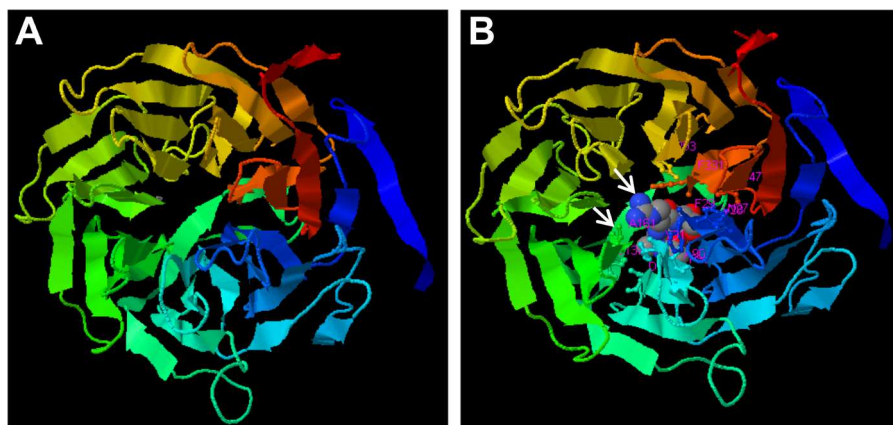


Figure 2. Structure of LuWD40-1 based on I-Tasser prediction. (A) Predicted propeller structure of LuWD40-1 protein based on 10 highly similar templates. The beta sheets and turns are drawn as arrows and lines, respectively. (B) Predicted binding sites on LuWD40-1 based on six most probable templates are shown with arrows. The LuWD40-1 protein is depicted as a cartoon and the binding proteins are illustrated as spheres. doi:10.1371/journal.pone.0069124.g002

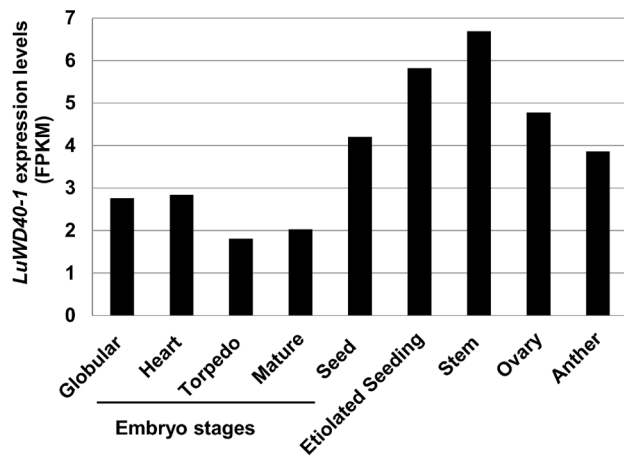


Figure 3. Transcript profile of *Lus10011938/LuWD40-1* based on RNA-Seq data. The RNA extracted from wild type plants were sequenced from various tissues. The abundance of *Lus10011938/LuWD40-1* transcripts in different tissue samples are represented by normalized FPKM values. RNA-Seq and related data normalization is described in the Materials and Methods section. doi:10.1371/journal.pone.0069124.g003

pollen viability and it was difficult to obtain viable seeds from these lines for further propagation. Only 2–5 viable seeds were obtained from each OE line and which were then grown to produce the T₂ generation. The three T₂ lines OE-1, -2 and -5 all showed overexpression of *LuWD40-1* as compared to the untransformed *Prairie Grande* (Figure 5A).

Pollen viability was estimated by the MTT test where viable pollen cells stain dark purple due to the conversion of the yellow

MTT by the mitochondrial succinate dehydrogenase into a purple formazan product [60] while the non-viable pollen remains colourless (Figure 5B). The *Prairie Grande* pollen was dense and sank to the bottom of the tube upon immersion in water. The pollen grains of the OE lines mostly floated on the water surface. When visualized under the microscope with light source below the tissue, the *Prairie Grande* pollen was visualized as dark bodies due to the absorbance of the light by the active tissue within the cells whereas the OE pollen formed clear circular structures through which light was transmitted (Figure 5B). Pollen viability was estimated to be 96% in *Prairie Grande* but ranged from 12% to 16% in the T₂ *LuWD40-1* OE lines (Figure 5C). Attempts were made to test the viability of the female reproductive organs in the OE lines. Cross pollination of 30 flowers in the three OE lines with *Prairie Grande* pollen resulted in seed set in ten OE flowers (~33% success) suggesting retention of female fertility.

The male sterility observed in the *LuWD40-1* OE lines limited our ability to test sufficient numbers of individuals for each transgenic line generated. Hence, the physiological differences were compared between *Prairie Grande* and the three transgenic OE lines and not within each lineage. The OE lines showed clear signs of delayed growth. The slower growth habit of OE plants translated into significantly delayed flowering phenotype that and took ~80 days to flower compared to *Prairie Grande* which reached that stage in ~57 days (Fig. S4A). Fewer adventitious buds led to the production of fewer branches in OE lines compared to *Prairie Grande* which produced ~10 branches per plant whereas the OE lines only had three (Fig. S4B). With fewer branches, the OE plants grew taller averaging 70 cm whereas *Prairie Grande* averaged 50 cm (Figure 6A; Fig. S4C). The petals of the *Prairie Grande* flowers open completely and fell prior to boll development (Figure 6B) while flowers of the OE lines remained closed and the petals persisted in that shape until maturity.

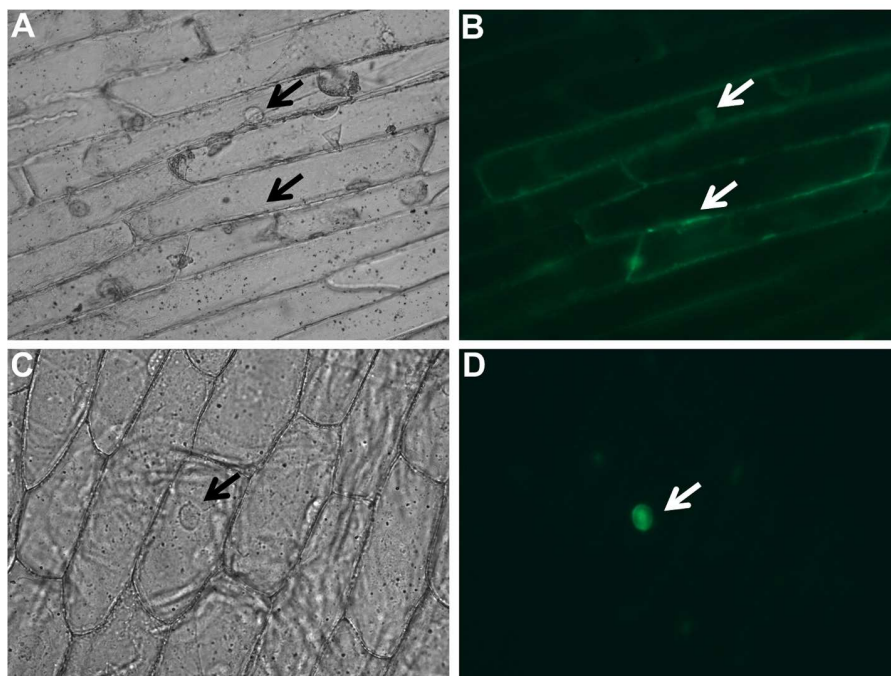


Figure 4. Subcellular localization of *LuWD40-1*-GFP in onion epidermal cells. The CaMV 35S promoter driven empty vector CD3-685-GFP and the CD3-685-*LuWD40-1*-GFP construct were transformed into onion epidermal cells using particle bombardment. The subcellular compartmentalization of the empty vector constructs under bright field (A) and fluorescence (B) shows diffusion throughout the cells (arrows). Transient expression of the *LuWD40-1*-GFP construct under bright field (C) and fluorescence (D) indicates localization in the nucleus (arrows). doi:10.1371/journal.pone.0069124.g004

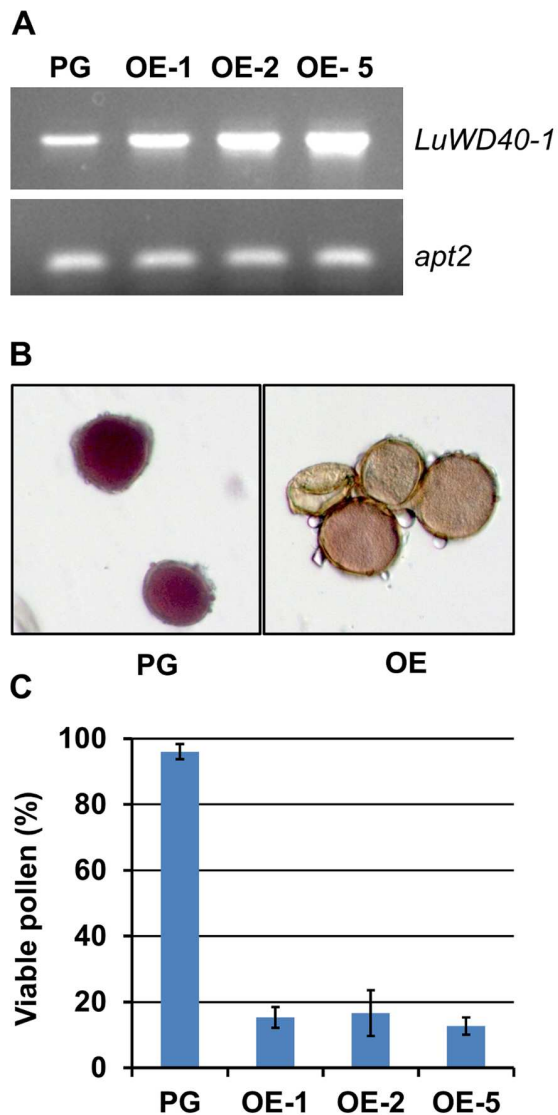


Figure 5. Transgenic lines overexpress *LuWD40-1* and show pollen sterility. (A) Semi-quantitative RT-PCR of *LuWD40-1* and the *apt1* control shows overexpression of *LuWD40-1* in the three T₂ transgenic lines compared to the untransformed Prairie Grande (PG). (B) The PG lines with viable pollen (left) and the OE lines with non-viable pollen (right) as visualized by MTT staining for pollen viability. (C) Pollen viability in PG and OE lines. Pollen viability was measured using pollen grains from three flowers per plant. doi:10.1371/journal.pone.0069124.g005

Removal of the petals revealed normal development of reproductive organs in Prairie Grande but in OE lines, the anthers were shrivelled and had empty pollen cells (Figures 5B and 6C). Numerous plump dark brown seeds developed on Prairie Grande whereas the seed set and quality was severely impeded in OE lines (Figure 6D). The majority of the seeds that developed on OE lines were shrivelled, bleached and had poor germination.

Discussion

Flax is a dual purpose crop grown for its omega-3 fatty acid rich seed oil (linseed) and its stem fibre (fibre flax). The increasing importance and value of the crop have justified the development of genetic resources to assist plant breeding efforts directed towards

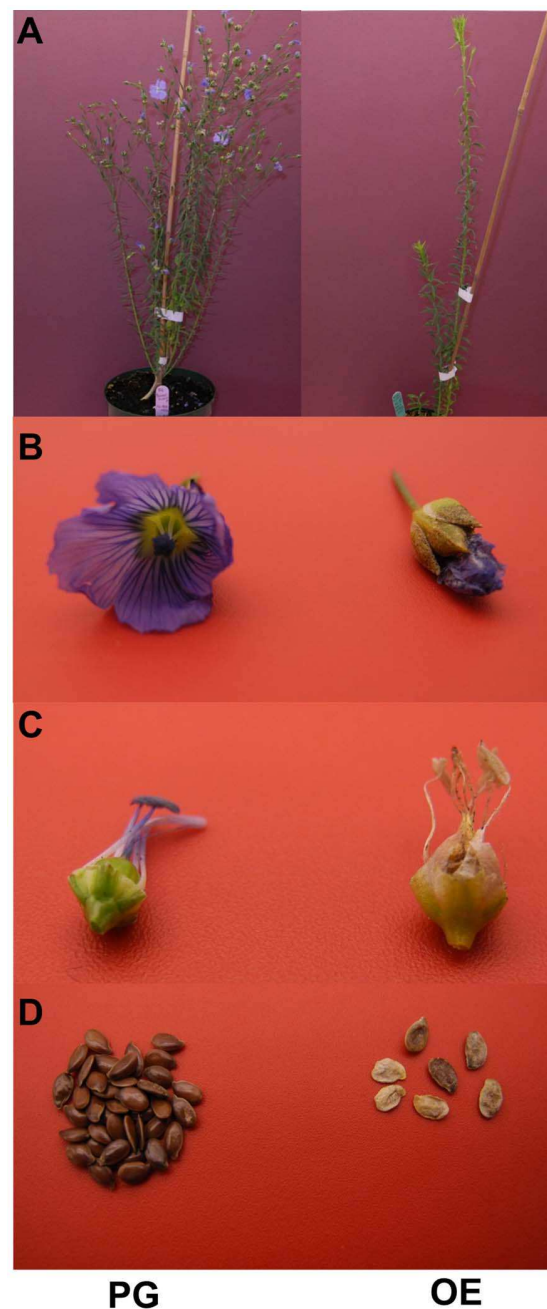


Figure 6. Phenotypic characterization of *LuWD40-1* transgenic lines. Images comparing the phenotypes of Prairie Grande (PG) and OE lines for (A) whole plant architecture, (B) flower morphology, (C) reproductive structures and (D) seed quality. doi:10.1371/journal.pone.0069124.g006

higher yielding varieties. The current advances include the first flax genome sequence [40], a collection of expressed sequence tags (ESTs) from more than 10 different tissues [61], a consensus genetic map [62], a physical map [63] and thousands of SNP markers [39].

Linseed and fibre types belong to the same species, *Linum usitatissimum*, a diploid self-pollinated crop with an estimated genome size of ~375 Mb and 15 chromosomes ($2n = 2 \times = 30$). In flax, self-pollination occurs just prior to flower petal opening. The flower only lasts part of a day during which non-fertilized

ovaries are still receptive, providing an opportunity for cross-pollination, although limited because pollen flow and insect-mediated pollination are minimal in flax [64]. Except for the recent development of a hybrid in China [65], flax varieties around the world are pure lines. Three independent and essential elements require investigations to justify the establishment of flax as a hybrid crop: (1) the quantification of its heterotic potential, (2) the development of an efficient male sterility system and (3) the establishment of a cost effective and reliable hybrid seed production system. Here, we investigated the *LuWD40-1* gene in flax for its potential use in a flax male sterility system.

The *LuWD40-1* homologue was found to be expressed exclusively in male sterile lines of Chinese flax accession tested only at the flower bud stage [2]. In our study, however, the *LuWD40-1* transcripts were present in all tested stages of CDC Bethune (Figure 3) as well as developing bolls of AC McDuff (EST Lu1225D10), both being male fertile flax accessions. The overexpression of *LuWD40-1* in male fertile accession Prairie Grande caused male sterility.

Analysis of the transcript features showed the presence of a large 5'UTR (Fig. S1A) possibly due to the retention of an intron during alternative splicing upstream of the ATG start codon. Suggestive evidence of intron selection comes from large scale analysis of human and mouse annotated genes where intron retention in UTR regions was demonstrated in ~15% and 22% of transcripts, respectively [66,67]. Large 5'UTRs are characteristics of regulatory genes encoding growth factors, homeodomain proteins and various receptors involved in post-transcriptional modifications [68]. They are also known to affect the expression pattern of their genes. In Arabidopsis, the long intron in the 5'UTR of elongation factor 1 α -A3 enhances its gene expression in a length dependent manner [69].

The WD40 protein has putative orthologs in worms and reptiles, as well as lower and higher plants (Figure 1A), however, considerable differences exist between the protein sequences, even within dicot species (Figure 1B). Variations within the WD40 motifs account for their functional versatility as the most active protein interactor. Similar to the well characterized transducin family with seven WD repeats [11], *LuWD40-1* is also predicted to form a seven blade propeller structure even though it harbours only five WD repeats (Figure 2A). Although uncommon, other WD40 proteins with five WD repeats also form seven bladed propeller structures such as the coronin family of proteins [70].

The localization of *LuWD40-1* into the nucleus (Figure 4) is consistent and supports the *in silico* prediction for interaction with histone binding proteins (Figure 2B). Several WD40 proteins such as WDR5, NURF55 and RbBP4/7 are classified as chromatin readers because they can recruit chromatin modifiers to specific sites on the chromosome by direct interaction with histones [71]. In addition to protein-protein interactions, WD40 repeat proteins such as BUB3 and Rae1 also act as docking proteins for assembly of protein complexes that regulate mitotic checkpoints [72,73]. The Arabidopsis WD40 repeat protein NEDD1 regulates microtubule development during cell proliferation [74]. A recent report identified the WD40 protein interaction with nucleic acid where a non-coding RNA transcript from the *HOXA* locus interacted directly with WDR5 [75]. These diverse WD40 proteins and their nucleic acid binding characteristics make them key regulatory candidates in transcriptional and post-transcriptional regulation during gametogenesis and growth.

The *Lus10011938/LuWD40-1* transcripts were expressed in all analysed developmental stages of flax (Figure 3). The

Lus10011938/LuWD40-1 transcript was highly expressed in stems (Figure 3) and the fact that the KO lines failed to produce shoots from calli suggests a putative essential role in shoot formation as also found with WD40 protein mutants that showed reduced shoot growth resulting in shorter plants [37]. The flax lines overexpressing *LuWD40-1* had delayed flowering, severely reduced pollen viability, branching and growth (Figure 5 and Fig. S4). These phenotypes are likely the result of the constitutive overexpression including ectopic misexpression of the *LuWD40-1*. The presence of promoter elements such as GTGANTG10 and POLLEN1LE-LAT52 that cause anther specific expression [76] is consistent with transcript abundance of *Lus10011938/LuWD40-1* in anthers (Figure 3), however, constitutive overexpression of the same gene affects male gametophyte development leading to the production of sterile pollen. Imbalance in overall transcript profile could result in improper protein interactions of *LuWD40-1* protein in pollen tissues. This hypothesis is supported by studies where pollen and anthers showed expression of approximately 15,000 genes, all of which were highly regulated [77].

Near complete male sterility (~90% non-viable pollen) in a population arising from an F_2 cross between oilseed type flax accession Double Low and cultivar AC McDuff was found in approximately 4% of the lines while 8% were partially sterile (~50% non-viable pollen). The occurrence of male sterility in the F_2 generation did not follow any simple ratios of gene inheritance or interaction. Epigenetic mechanisms involving small RNAs and/or transposable elements (TE) which are highly expressed in pollen vegetative nucleus cells but not in female gametophytes must be investigated [78].

Although other WD40 proteins like YAO and OsLIS-L1 have been implicated in embryo development, plant height and pollen formation [35,37], they differ considerably from *LuWD40-1* by their additional functional motifs. The Arabidopsis YAO, with its nuclear localization signal followed by seven WD40 repeats, is expressed ubiquitously in all plant tissues [35]. The rice OsLIS-L1 contains a LisH motif, a CTLH domain and nine WD40 repeats, out of which, four form the G_β homology domain [37] indispensable for inducing male sterility. No other domains but the WD40 repeats could be identified in *LuWD40-1*. Moreover, we observed the phenotypes (Figure 6A–D) in OE lines whereas our knockout lines did not produce shoots, suggesting conserved functions of different WD40 proteins in diverse pathways of plant morphogenesis and reproduction.

To explore the possibility of using *LuWD40-1* OE lines to develop hybrids, we crossed the male sterile flowers of OE lines with pollen from Prairie Grande. A total of ten out of 30 cross-pollinated flowers set seeds suggesting the fertile state of ovaries. The low rate of cross-pollination efficiency could be due to pollen incompatibility mechanisms. Constitutive expression of the *LuWD40-1* gene may also interfere with cellular processes affecting pollen tube growth in the style of OE lines similar to the Pollen tube Defective WD40 (PD40) gene in tobacco [79]. More likely, it reflects the inherent difficulties associated with controlled crossing such as ovary receptivity and pollen developmental stage at the time of pollination.

The current study lays a foundation for the generation of male sterile flax lines providing basic physiological characterization of *LuWD40-1*. More insights can be gained by studying the pollen specific expression of this gene in flax to separate the negative effects on plant growth from the desired male sterility phenotype. It is imperative to understand the inheritance of this system for its implementation in the field.

Supporting Information

Figure S1 Sequence and schematic representation of the LuWD40-1 protein encoded by cDNA clone LuP1225D10. (A) Nucleotide and deduced amino acid sequence of LuWD40-1 with 5' and 3' untranslated regions. The second in frame ATG selected in the *ab initio* annotation of *Lus10011938* is italicized. The 5' region amplified for semi-quantitative RT-PCR experiment is underlined (B) Schematic representation of the genomic clone corresponding to *LuWD40-1* ORF drawn to scale. The boxes represent exons and lines represent introns. (C) DNA sequence of the genomic clone representing the open reading frame of *LuWD40-1*.

(PDF)

Figure S2 Semi-quantitative RT-PCR. Amplification of the target gene *LuWD40-1* and the control gene *apt1* from the three transgenic lines overexpressing *LuWD40-1* and the untransformed Prairie Grande shows that the reactions had not reached saturation at 28 cycles but that saturation was achieved at 31 cycles. Ratios (*LuWD40-1:apt1*) at the bottom were calculated by measuring the amplicon intensity by densitometry using the AlphaImagerHP software version 3.4 (proteinsimple, Santa Clara, CA, USA). Note: the Gateway recombination tags in the *LuWD40-1* primers increase the amplification size to 332 bp (274+29 bp tags in each forward and reverse primers). Molecular marker (M) is the 1KB Plus DNA ladder (Invitrogen, Carlsbad, California, USA).

(PDF)

References

- Bateson W, Gairdner A (1921) Male-Sterility in flax, subject to two types of segregation. *J Genet* 11: 269–275.
- Hui Z, Bateer S, Xiao-yun J, Zi-qin L, Qiang L (2010) Study on genes related to male sterility in dominant genic male sterile flax In: Hammond JJ, editor. Proceedings of the 63rd flax institute of the United States. Fargo, North Dakota: Flax Institute of the United States. 112–122.
- Janda L, Tichy P, Spizek J, Petricek M (1996) A deduced *Thermomonospora curvata* protein containing serine/threonine protein kinase and WD-repeat domains. *J Bacteriol* 178: 1487–1489.
- Grigorieva G, Shestakov S (1982) Transformation in the cyanobacterium *Synechocystis* sp. 6803. *FEMS Microbiol Lett* 13: 367–370.
- Andrade MA, Perez-Iratxeta C, Ponting CP (2001) Protein repeats: structures, functions, and evolution. *J Struct Biol* 134: 117–131.
- Stirnemann CU, Petsalaki E, Russell RB, Muller CW (2010) WD40 proteins propel cellular networks. *Trends Biochem Sci* 35: 565–574.
- Neer EJ, Schmidt CJ, Nambudripad R, Smith TF (1994) The ancient regulatory-protein family of WD-repeat proteins. *Nature* 371: 297–300.
- Smith TF, Gaitatzes C, Saxena K, Neer EJ (1999) The WD repeat: a common architecture for diverse functions. *Trends Biochem Sci* 24: 181–185.
- Fong HK, Hurley JB, Hopkins RS, Miake-Lye R, Johnson MS, et al. (1986) Repetitive segmental structure of the transducin beta subunit: homology with the CDC4 gene and identification of related mRNAs. *P Natl Acad Sci USA* 83: 2162–2166.
- Wall MA, Coleman DE, Lee E, Iniguez-Lluhi JA, Posner BA, et al. (1995) The structure of the G protein heterotrimer $G_{\alpha\beta\gamma}$. *Cell* 83: 1047–1058.
- Lambright DG, Sondek J, Bohm A, Skiba NP, Hamm HE, et al. (1996) The 2.0 Å crystal structure of a heterotrimeric G protein. *Nature* 379: 311–319.
- Paoli M (2001) An elusive propeller-like fold. *Nat Struct Biol* 8: 744–745.
- Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, et al. (2008) Single-molecule DNA sequencing of a viral genome. *Science* 320: 106–109.
- Collins SR, Kemmeren P, Zhao XC, Greenblatt JF, Spencer F, et al. (2007) Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. *Mol Cell Prot* 6: 439–450.
- Colditz F, Nyamsuren O, Niehaus K, Eubel H, Braun HP, et al. (2004) Proteomic approach: identification of *Medicago truncatula* proteins induced in roots after infection with the pathogenic oomycete *Aphanomyces euteiches*. *Plant Mol Biol* 55: 109–120.
- ter Haar E, Harrison SC, Kirchhausen T (2000) Peptide-in-groove interactions link target proteins to the beta-propeller of clathrin. *P Natl Acad Sci USA* 97: 1096–1100.
- Hardwick KG, Johnston RC, Smith DL, Murray AW (2000) MAD3 encodes a novel component of the spindle checkpoint which interacts with Bub3p, Cdc20p, and Mad2p. *J Cell Biol* 148: 871–882.
- Hattendorf DA, Andreeva A, Gangar A, Brennwald PJ, Weis WI (2007) Structure of the yeast polarity protein Sro7 reveals a SNARE regulatory mechanism. *Nature* 446: 567–571.
- Feldman RM, Correll CC, Kaplan KB, Deshaies RJ (1997) A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. *Cell* 91: 221–230.
- Vaisman N, Tsouladze A, Robzyk K, Ben-Yehuda S, Kupiec M, et al. (1995) The role of *Saccharomyces cerevisiae* Cdc40p in DNA replication and mitotic spindle formation and/or maintenance. *Mol Gen Genet* 247: 123–136.
- de Hostos EL, Bradtke B, Lottspeich F, Guggenheim R, Gerisch G (1991) Coronin, an actin binding protein of *Dictyostelium discoideum* localized to cell surface projections, has sequence similarities to G protein beta subunits. *Embo J* 10: 4097–4104.
- Pryer NK, Salama NR, Schekman R, Kaiser CA (1993) Cytosolic Sec13p complex is required for vesicle formation from the endoplasmic reticulum *in vitro*. *J Cell Biol* 120: 865–875.
- Hoey T, Weinzierl RO, Gill G, Chen JL, Dynlacht BD, et al. (1993) Molecular cloning and functional analysis of *Drosophila* TAF110 reveal properties expected of coactivators. *Cell* 72: 247–260.
- Williams FE, Varanasi U, Trumbly RJ (1991) The CYC8 and TUP1 proteins involved in glucose repression in *Saccharomyces cerevisiae* are associated in a protein complex. *Mol Cell Biol* 11: 3307–3316.
- Song JJ, Garlick JD, Kingston RE (2008) Structural basis of histone H4 recognition by p55. *Genes Dev* 22: 1313–1318.
- Margueron R, Justin N, Ohno K, Sharpe ML, Son J, et al. (2009) Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* 461: 762–767.
- van Nocker S, Ludwig P (2003) The WD-repeat protein superfamily in Arabidopsis: conservation and divergence in structure and function. *BMC Genomics* 4: 50.
- Xiang D, Yang H, Venglat P, Cao Y, Wen R, et al. (2011) POPCORN functions in the auxin pathway to regulate embryonic body plan and meristem organization in Arabidopsis. *Plant cell* 23: 4348–4367.
- Lee JH, Terzaghi W, Gusmaroli G, Charron JB, Yoon HJ, et al. (2008) Characterization of Arabidopsis and rice DWD proteins and their roles as substrate receptors for CUL4-RING E3 ubiquitin ligases. *Plant cell* 20: 152–167.
- Dharmasiri N, Dharmasiri S, Estelle M (2005) The F-box protein TIR1 is an auxin receptor. *Nature* 435: 441–445.

Figure S3 Confirmation of T₁ transgenic lines by PCR using CaMV35S and gene specific primer combination on genomic DNA extracted from leaves. Details of primers and PCR conditions are available in the Materials and Methods section.

(PDF)

Figure S4 Physiological characterization of three *LuWD40-1* transgenic lines. Graphs showing differences between Prairie Grande (PG) and overexpressing (OE) lines in physiological parameters: (A) bolting time, (B) branching and (C) height of the plants.

(PDF)

Table S1 Transcription factor binding sites in the putative *LuWD40-1* promoter.

(PDF)

Acknowledgments

The authors wish to acknowledge the technical help of Susan Brown and Natasa Radovanovic. The authors thank Dr. Claudio Stasolla for providing CD3-685 GFP vector and Dr. Colin Hiebert for providing microscopy facilities. The authors are grateful to Andrzej Walichnowski for help with manuscript editing.

Author Contributions

Conceived and designed the experiments: SK SC. Performed the experiments: SK MCJ RD SC. Analyzed the data: SK RD SC. Contributed reagents/materials/analysis tools: MCJ RD SC. Wrote the paper: SK RD SC.

31. Sorensen MB, Chaudhury AM, Robert H, Bancharel E, Berger F (2001) Polycomb group genes control pattern formation in plant seed. *Current Biol* 11: 277–281.
32. Tuteja N, Sopory SK (2008) Plant signaling in stress: G-protein coupled receptors, heterotrimeric G-proteins and signal coupling via phospholipases. *Plant Signal Behav* 3: 79–86.
33. Albert TJ, Molla MN, Muzny DM, Nazareth L, Wheeler D, et al. (2007) Direct selection of human genomic loci by microarray hybridization. *Nat Methods* 4: 903–905.
34. Mishra AK, Puranik S, Bahadur RP, Prasad M (2012) The DNA-binding activity of an AP2 protein is involved in transcriptional regulation of a stress-responsive gene, *SiWD40*, in foxtail millet. *Genomics* 100: 252–263.
35. Li HJ, Liu NY, Shi DQ, Liu J, Yang WC (2010) YAO is a nucleolar WD40-repeat protein critical for embryogenesis and gametogenesis in Arabidopsis. *BMC Plant Biol* 10: 169.
36. Han MJ, Jung KH, Yi G, Lee DY, An G (2006) Rice Immature Pollen 1 (RIP1) is a regulator of late pollen development. *Plant Cell Physiol* 47: 1457–1472.
37. Gao X, Chen Z, Zhang J, Li X, Chen G, et al. (2012) *OxLIS-L1* encoding a lissencephaly type-1-like protein with WD40 repeats is required for plant height and male gametophyte formation in rice. *Planta* 235: 713–727.
38. Cloutier S, Niu Z, Datla R, Duguid S (2009) Development and analysis of EST-SSRs for flax (*Linum usitatissimum* L.). *Theor Appl Genet* 119: 53–63.
39. Kumar S, You FM, Cloutier S (2012) Genome wide SNP discovery in flax through next generation sequencing of reduced representation libraries. *BMC Genomics* 13: 684.
40. Wang Z, Hobson N, Galindo L, Zhu S, Shi D, et al. (2012) The genome of flax (*Linum usitatissimum*) assembled *de novo* from short shotgun sequence reads. *Plant J* 72: 461–473.
41. Reese MG (2001) Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Comp Chem* 26: 51–56.
42. Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res* 27: 297–300.
43. Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, et al. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* 38: W695–W699.
44. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
45. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25: 1105–1111.
46. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* 7: 562–578.
47. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, et al. (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol* 31: 46–53.
48. Himmelbach A, Zierold U, Hensel G, Riechen J, Douchkov D, et al. (2007) A set of modular binary vectors for transformation of cereals. *Plant Physiol* 145: 1192–1200.
49. Gutierrez L, Mauriat M, Guenin S, Pelloux J, Lefebvre JF, et al. (2008) The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. *Plant Biotechnol J* 6: 609–618.
50. Banik M, Duguid S, Cloutier S (2011) Transcript profiling and gene characterization of three fatty acid desaturase genes in high, moderate, and low linolenic acid genotypes of flax (*Linum usitatissimum* L.) and their role in linolenic acid accumulation. *Genome* 54: 471–483.
51. Wijayanto T, McHughen A (1999) Genetic transformation of *Linum* by particle bombardment. *In Vitro Cell Dev Biol* 35: 456–465.
52. Sivamani E, DeLong RK, Qu R (2009) Protamine-mediated DNA coating remarkably improves bombardment transformation efficiency in plant cells. *Plant Cell Rep* 28: 213–221.
53. Libus J, Storchova H (2006) Quantification of cDNA generated by reverse transcription of total RNA provides a simple alternative tool for quantitative RT-PCR normalization. *BioTechniques* 41: 156–160.
54. Guenin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C, et al. (2009) Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *J Exp Bot* 60: 487–493.
55. Firmage DH, Dafni A, editors (2001) Field tests for pollen viability; a comparative approach: *Acta Hort.* (ISHS). 561: 587–594 p.
56. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* 5: 725–738.
57. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22: 195–201.
58. Sprague ER, Redd MJ, Johnson AD, Wolberger C (2000) Structure of the C-terminal domain of Tup1, a corepressor of transcription in yeast. *EMBO J* 19: 3016–3027.
59. Murzina NV, Pei XY, Zhang W, Sparkes M, Vicente-Garcia J, et al. (2008) Structural basis for the recognition of histone H4 by the histone-chaperone RbAp46. *Structure* 16: 1077–1085.
60. Mosmann T (1983) Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *J Immunol Methods* 65: 55–63.
61. Venglat P, Xiang D, Qiu S, Stone SL, Tibiche C, et al. (2011) Gene expression analysis of flax seed development. *BMC Plant Biol* 11: 74.
62. Cloutier S, Miranda E, Ragupathy R, Radovanovic N, Reimer E, et al. (2012) Integrated consensus genetic and physical maps of flax (*Linum usitatissimum* L.). *Theor Appl Genet* 125: 1783–1795.
63. Ragupathy R, Rathinavelu R, Cloutier S (2011) Physical mapping and BAC-end sequence analysis provide initial insights into the flax (*Linum usitatissimum* L.) genome. *BMC Genomics* 12: 217.
64. Jhala AJ, Bhatt H, Topinka K, Hall LM (2011) Pollen-mediated gene flow in flax (*Linum usitatissimum* L.): can genetically engineered and organic flax coexist? *Heredity* 106: 557–566.
65. Hammond JJ (2010) Proceedings of the 63rd flax institute of the United States. In: Hammond JJ, editor; 2010; Fargo, North Dakota. Flax Institute of the United States. 1–133.
66. Galante PA, Sakabe NJ, Kirschbaum-Slager N, de Souza SJ (2004) Detection and evaluation of intron retention events in the human transcriptome. *RNA* 10: 757–765.
67. Sharov AA, Dudekula DB, Ko MS (2005) Genome-wide assembly and analysis of alternative transcripts in mouse. *Genome Res* 15: 748–754.
68. van der Velden AW, Thomas AAM (1999) The role of the 5' untranslated region of an mRNA in translation regulation during development. *Int J Biochem Cell Biol* 31: 87–106.
69. Chung B, Simons C, Firth A, Brown C, Hellens R (2006) Effect of 5'UTR introns on gene expression in *Arabidopsis thaliana*. *BMC Genomics* 7: 120.
70. Appleton BA, Wu P, Wiesmann C (2006) The crystal structure of murine coronin-1: a regulator of actin cytoskeletal dynamics in lymphocytes. *Structure* 14: 87–96.
71. Migliori V, Mapelli M, Guccione E (2012) On WD40 proteins: propelling our knowledge of transcriptional control? *Epigenetics* 7: 815–822.
72. Wang X, Babu JR, Harden JM, Jablonski SA, Gazi MH, et al. (2001) The mitotic checkpoint protein hBUB3 and the mRNA export factor hRAE1 interact with GLE2p-binding sequence (GLEBS)-containing proteins. *J Biol Chem* 276: 26559–26567.
73. Sitterlin D (2004) Characterization of the *Drosophila* Rael1 protein as a G1 phase regulator of the cell cycle. *Gene* 326: 107–116.
74. Zeng CJT, Lee YRJ, Liu B (2009) The WD40 repeat protein NEDD1 functions in microtubule organization during cell division in *Arabidopsis thaliana*. *Plant Cell* 21: 1129–1140.
75. Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, et al. (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472: 120–124.
76. Hobo T, Suwabe K, Aya K, Suzuki G, Yano K, et al. (2008) Various spatiotemporal expression profiles of anther-expressed genes in rice. *Plant Cell Physiol* 49: 1417–1428.
77. Twell D (1994) Molecular and cellular aspects of plant reproduction; Scott RJ, Stead AD, editors: Cambridge University Press.
78. Slotkin RK, Vaughn M, Borges F, Tanurdzic M, Becker JD, et al. (2009) Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* 136: 461–472.
79. Hafidh S, Breznova K, Ruzicka P, Fecikova J, Capkova V, et al. (2012) Comprehensive analysis of tobacco pollen transcriptome unveils common pathways in polar cell expansion and underlying heterochronic shift during spermatogenesis. *BMC Plant Biol* 12: 24.