

RESEARCH ARTICLE

Molecular identification and functional characterization of a cyanogenic glucosyltransferase from flax (*Linum unsitatissimum*)

Michael Kazachkov¹, Qiang Li^{1,2^{aa}}, Wenyun Shen¹, Liping Wang^{1^{ab}}, Peng Gao^{1^{ac}}, Daoquan Xiang¹, Raju Datla^{1^{ac}}, Jitao Zou^{1*}

1 National Research Council Canada, Saskatoon, Saskatchewan, Canada, **2** Department of Plant Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

^{aa} Current address: College of Plant Science and Technology, Huazhong Agricultural University

^{ab} Current address: Department of Molecular and Cellular Biology, University of Guelph, Ontario, Canada

^{ac} Current address: Global Institute of Food Security, University of Saskatchewan, Saskatchewan, Canada

* Jitao.Zou@nrc-cnrc.gc.ca



OPEN ACCESS

Citation: Kazachkov M, Li Q, Shen W, Wang L, Gao P, Xiang D, et al. (2020) Molecular identification and functional characterization of a cyanogenic glucosyltransferase from flax (*Linum unsitatissimum*). PLoS ONE 15(2): e0227840. <https://doi.org/10.1371/journal.pone.0227840>

Editor: Craig Eliot Coleman, Brigham Young University, UNITED STATES

Received: October 7, 2019

Accepted: December 30, 2019

Published: February 5, 2020

Copyright: © 2020 Kazachkov et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: This research was supported by the National Research Council of Canada core research program and by the TUFGEN project from Genome Prairie and Genome Canada, a not-for-profit organization that is leading's national strategy on genomics. There was no additional external funding received for this study. This is National

Abstract

Flax seed has become consumers' choice for not only polyunsaturated alpha-linolenic fatty acid but also nutraceuticals such as lignans and soluble fiber. There is, however, a major drawback of flax as a source of functional food since the seeds contain significant level of cyanogenic glucosides. The final step of cyanogenic glucoside biosynthesis is mediated by UDP-glucose dependent glucosyltransferase. To date, no flax cyanogenic glucosyl transferase genes have been reported with verified biochemical functionality. Here we present a study on the identification and enzymatic characterization of a first flax cyanogenic glucosyltransferase, LuCGT1. We show that LuCGT1 was highly active towards both aliphatic and aromatic substrates. The *LuCGT1* gene is expressed in leaf tissues as well as in developing seeds, and its expression level was drastically reduced in flax mutant lines low in cyanogenic glucosides. Identification of LuCGT1 provides a molecular handle for developing gene specific markers for targeted breeding of low cyanogenic glucosides in flax.

Introduction

Cyanogenic glucosides are found in plant species of more than 130 families and have important functions in plant defense against herbivores and pathogens [1–3]. Unfortunately as constituents in plant-based food cyanogenic glucosides are extremely undesirable because they can be converted to hydrocyanic acid, a highly poisonous compound, when encounter glycosidase or under low pH in the digestive tracks of animals [4–10]. Flax (*Linum unsitatissimum*) is one of the major economic crops known to possess high level of cyanogenic glucosides, linamarin and lotaustralin [11]. Flax seed is a valuable source of functional food enriched with polyunsaturated alpha-linolenic fatty acid and soluble fiber. Owing to concerns of adversary effects of cyanogenic glycosides, it is generally advised that dietary consumption of flax seed is

Research Council Canada publication NRCC# 56332.

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

to be restricted to a limited amount. Flax seed is also one of the richest sources of lignans for the nutraceutical supplements industry. Unfortunately, structural and physical similarities render extraction of lignans free of cyanogenic compounds from flax seed extremely challenging. Hence, a major goal of flax quality improvement is the reduction of cyanogenic glucosides in seed.

The biosynthetic pathway for cyanogenic glucosides has been extensively studied in different plants including flax (*Linum usitatissimum*) [11–13], cassava [14–17], sorghum [6, 9, 10, 14, 17–22], *Trifolium repens* [23], *Prunus serotina* [7–8]. Existing literature [12, 24] established that the last step in the biosynthetic pathway of cyanogenic glucosides involves glycosyltransferase using UDP-glucose as glucosyl donor. The glucosyltransferase reaction attaches a sugar moiety to a cyanogenic glucosyl acceptor through a glucosidic bond. It is a biochemical step frequently found in secondary metabolism and serves to stabilize and store the metabolites that otherwise would be unstable or toxic to the cell. Genome based bioinformatics analysis showed that there are over one hundred putative glucosyltransferase genes in the genome of flax [25]. Glucosyltransferases, despite of well characterized function domains, are known to be promiscuous in terms of substrate specificity. To date, no genes encoding for cyanogenic glucosyltransferase from flax have been reported. Herein we report the molecular and biochemical characterization of the first flax cyanogenic glucosyltransferase, LuCGT1. Identification of this gene provides a molecular handle for breeding efforts to reduce cyanogenic glucoside level through either molecular marker-assisted selection or targeted gene-knockout endeavors.

Materials and methods

Plant materials and growth conditions

Flax cultivar CDC Bethune (wild type) was grown in growth chambers with 8 hour dark / 16 hour illumination ($75\text{--}100\ \mu\text{E m}^{-2}\ \text{s}^{-1}$). The day/night temperature was maintained at $22/17^\circ\text{C}$. All biochemical analyses were performed with leaves harvested from one month old plants except developing seeds collected 6 or 17 days after flowering. The three EMS flax mutant lines low in cyanogenic glucocides [26], UGG102-2, UGG146-1 and Double Low (DL), kindly provided by Dr. Scott Duguid, Morden Research Station, Agriculture and Agrifood Canada, were raised at identical conditions as that for CDC Bethune.

Strains and reagents

Yeast strains: BY4741 (WT, MATa hisΔ leuΔ metΔ uraΔ) were purchased from European *Saccharomyces cerevisiae* archive for functional analysis (EUROSCARF). ^{14}C -Uridine 5'-diphosphoglucose ^{14}C -UDP-glucose was purchased from American Radiolabeled Chemicals Inc. Substrates and other chemicals used in this study were obtained from Sigma-Aldrich. Media and plates were prepared according to Invitrogen's recipe described for the pYES2.1 TOPO TA Cloning Kit.

PCR amplification and construct generation

The full-length cDNA sequence of candidate glucosyltransferase genes were derived from a flax developing seed cDNA library as previously described [27]. For full length cDNA cloning, the complete coding regions of UGT85Q1, UGT74S1 and UGT84G3 were amplified with the following primers incorporating *Sph*I and *Bam*HI sites: UGT85Q1-F2/ UGT85Q1-R2, UGT74S1-F2/ UGT74S1-R2 and UGT84G3-F2/ UGT84G3-R2. The primers are listed in [S1 Table](#). The purified fragments were digested with *Sph*I and *Bam*HI for UGT85Q1, and *Sph*I

and *Bgl*II for UGT74S1 and UGT84G3, respectively. The digested fragments were then ligated to the pQE70 vector for sequencing verification. For the yeast expression system, the coding regions of UGT85Q1, UGT74S1 and UGT84G3 were amplified with UGT85Q1-F1/ UGT85Q1-R1, UGT74S1-F1/ UGT74S1-R1 and UGT84G3-F1/ UGT84G3-R1, respectively. The obtained fragments were ligated directly to the pYES2.1 vector according to manufacturer's instruction (Invitrogen). The resulting plasmids were verified through DNA sequencing and subsequently introduced into yeast strain BY4741. To examine the promoter and coding region of UGT85Q1, five consecutive fragments were amplified with the following primers covering the whole region: Pro-F1/ Pro-R1, Pro-F2/ Pro-R2, Pro-F3/ Pro-R3, Pro-F4/ Pro-R4 and Pro-F5/ Pro-R5 (S1 Table).

Heterologous expression in yeast

All procedures were carried out at 4°C except where indicated. Yeast strains were first grown (28°C) in 15 mL of SC-Ura medium containing 2% glucose. Protein expression induction was carried out as described in the manufacturer manual (Invitrogen). After 24 h of growth (28°C) in SC + 2% galactose + 1% raffinose, the cells were washed sequentially with distilled water and homogenization buffer (50 mM Tris-HCl, 1 mM EDTA, 0.6 M sorbitol, pH 7.4, 1 mM dithiothreitol (DTT)). After centrifugation at 4,000 rpm (Eppendorf 5145C), the cells were resuspended in 1 mL homogenization buffer containing 10 µL yeast protease cocktail (Sigma), and shaken vigorously (1 min X 2 times) in the presence of acid-washed glass beads (diameter 0.5 mm). The resultant homogenate was centrifuged at 12,000 rpm for 10 min. The decanted supernatant was further centrifuged at 100,000 g for 90–120 min. The supernatant was carefully separated from the pellet and collected for further experiments. The pellet was resuspended in homogenization buffer containing 20% glycerol and frozen at -80 °C along with supernatant until use. Protein concentration was measured using a Bio-Rad Protein Assay Kit for final enzyme activity calculation.

In vitro assay of glucosyltransferase activity

Glucosyltransferase kinetic constants and substrate specificity were both determined by measuring incorporation of ¹⁴C-UDP-glucose incorporation into cyanogenic glycosides. Glucosyltransferase substrate specificity assessment was performed in 0.1 mL HEPES (pH 8.0, 0.1 M) buffer containing 100 µg of protein, 5mM of ¹⁴C-UDP glucose (0.25 nCi/nmol) substrates and 100 mM aglycones. All assays were performed at least twice. Reaction was allowed to proceed for 30 min at 30°C with 700 rpm shaking and stopped by adding 20 µL acidic acid (20%). All reactions were linear at least in 1 h range. For determination of apparent *K_m* and *V_{max}*, 100 µg of protein was used in the same reaction system with different concentrations of ¹⁴C-UDP-glucose and Acetone Cyanohydrin. The reaction products were spotted on Merck silica G60 TLC plates and separated in Ethyl Acetate : Acetone : Dichloromethane : Methanol : Water (40 : 30 : 12 : 10 : 8) system. Spots corresponding to different cyanogenic glucoside products were scraped off and ¹⁴C incorporation was scintillation counted. Calculation of concentration acetone cyanohydrin (undissociated) in equilibrium has been performed in accordance with a previous study [20].

Total RNA extraction and qRT-PCR

Isolation and purification of total RNA for quantitative RT-PCR (qRT-PCR) was performed using mini plant RNA extraction Kit according to the manufacturer's introductions (QIAGEN). Approximately 1 µg of total RNA was used for cDNA synthesis using QuantiTect reverse transcription kit (QIAGEN). The obtained cDNA was diluted fifty times to conduct

qRT-PCR. For each reaction, a volume of 15 μL containing 5 μL of the diluted cDNA, 7.5 μL of SYBR Green Master Mix (Bio-Rad) and 0.6 μL of primer (5 μM). The reference gene *Actin* and the target gene *UGT85Q1* were amplified using primers Actin-F/Actin-R and UGT85Q1-F3/UGT85Q1-R3, respectively (S1 Table).

Results

Identification of flax cyanogenic glucosyltransferase

We were specifically interested in cyanogenic glucosyltransferases that are highly expressed in developing seeds. Previous reports using crude enzyme preparations from flax demonstrated that the flax glucosyltransferase possesses properties resembling that of enzyme from sorghum [28]. We thus used the deduced amino acid sequence of the *sorghum bicolor* cyanogenic glucosyltransferase [19, 22] as a reference to conduct BLAST search against a flax EST database generated from developing seed [27]. This resulted in the identification of 38 EST entries that displayed significant sequence homology to the sorghum enzyme (S1 Fig).

Three candidate genes, previously annotated as UGT85Q1 (GenBank accession: ADV36300) [25], UGT74S1 (Genbank: JX011632.1) and UGT84G3 (GenBank: AFJ52991.1) were found to be abundantly represented in the developing seed EST database [27]. Full length sequence of the cDNA was established based on sequencing of multiple independent cDNA clones. UGT74S1 was previously shown to possess UDP-glucosyltransferase activity towards secoisolariciresinol and contribute to lignan biosynthesis in flax [29]. The function of UGT84G3 is currently unknown. This study focuses on UGT85Q1, which is herein designated as *LuCGT1* and encodes a polypeptide of 492 amino acids. The deduced amino acid sequence of *LuCGT1* contains at its C-terminal region a putative nucleotide-diphosphate-sugar binding domain, ³⁶⁹WCPQEDVLNHPAVGGFLTHCGWGSIIESLTAGVPLLCWPFPGDQ⁴¹², commonly known as the “plant Secondary Product Glycosyltransferase” (PSPG) [30]. There are two other close related homologs in flax, UGT85Q3 (AFJ53000) and UGT85Q2 (AFJ52999), to which *LuCGT1* displays 69% and 66% sequence identity, respectively.

LuCGT1 encodes a functional cyanogenic glucosyltransferase

We inserted the full length cDNA of *LuCGT1*, *UGT74S1*, and *UGT84G3* into the yeast expression vector pYES2.1. The yeast expression constructs were subsequently introduced into yeast strain BY4741. We conducted enzymatic assays using homogenate of yeast cultures upon galactose induction. Our initial assays employed acetone cyanohydrin and mandelonitrile as glucosyl acceptors and ¹⁴C-UDP-glucose as glucosyl donor. Fig 1 illustrates one representative TLC plate from our experiments. Repeated assays using lysate of BY4741 harboring *UGT74S1* or *UGT84G3* detected no appreciable enzyme activity with either glucosyl acceptors. These results showed that the yeast expression system possessed no background cyanogenic glucosyltransferase assays, and that despite of high sequence homology, neither *UGT74S1* nor *UGT84G3* encoded the enzyme of our interest. Yeast lysate harboring *LuCGT1*, on the other hand, displayed strong activity towards acetone cyanohydrin. It should be noted, however, assays with the same lysate failed to detect activity when mandelonitrile was used as glucosyl acceptor (Fig 1A, lane 8), indicating an apparent substrate preference of this enzyme. The undissociated acetone cyanohydrin concentration in the *in vitro* enzymatic assay of *LuCGT1*, determined according to Mederacke et al [18], was found to be less than 0.2% in *Km* range of concentration (Fig 1B).

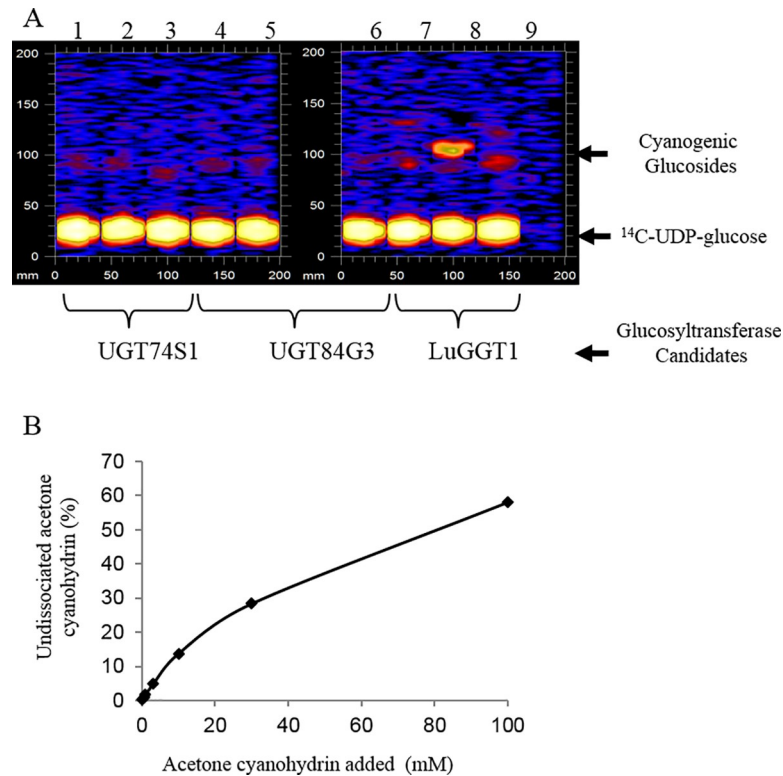


Fig 1. Glucosyltransferase activity of LuCGT1 expressed in yeast. (A) Cyanohydrin UDPG-glycosyltransferase assay using ^{14}C -UDP-glucose in the presence of mandelonitrile (lane 1, 4, 7), acetone cyanohydrin (lane 2, 5, 8), and negative control without any glucosyl acceptor (lane 3, 6, 9). Formation of radiolabeled product is evident in yeast strain expressing LuCGT1 (lane 8). (B) Concentration of acetone cyanohydrin in equilibrium at cyanogenic glucosyltransferase assay. Data for the amount of undissociated acetone cyanohydrin were calculated according to Mederacke et al [15].

<https://doi.org/10.1371/journal.pone.0227840.g001>

Enzyme kinetics of LuCGT1

We further studied the properties of LuCGT1 by assessing the optimal biochemical parameter of the glucosyltransferase reaction. Temperature ranging 5 to 50°C, and pH from 3 to 12 were tested for optimal temperature and pH. Temperature optimum for LuCGT1 was found at 40–42°C (Fig 2A), and the pH optimum for this cyanogenic glucosyltransferase was observed at between pH 7.5 and 8.5 (Fig 2B).

After optimization of the enzymatic assay and determination of the reaction pH and temperature optimum, we were able to assess the apparent K_m and V_{max} of LuCGT1 for acetone cyanohydrin and UDP-Glucose. The apparent K_m for acetone cyanohydrin and uridine diphosphoglucose were at $61.3 \pm 4.5 \mu\text{M}$ and $248 \pm 56 \mu\text{M}$, respectively (Table 1). The four folds difference in K_m between acetone cyanohydrin and UGP-Glucose were consistent with the known physical concentration of these substrates in plant tissues and the fact that the acetone cyanohydrin itself is very unstable, particularly under low pH conditions. Under the optimal pH and at $2x K_m$ concentration, it is estimated that there would be only about 0.2 percent of acetone cyanohydrin that remains in undissociated form suitable for enzyme action.

Substrate specificity of LuCGT1

Our initial enzyme activity assessment using acetone cyanohydrin and mandelonitrile indicated a glucosyl acceptor substrate selectivity of LuCGT1. Given that cyanogenic

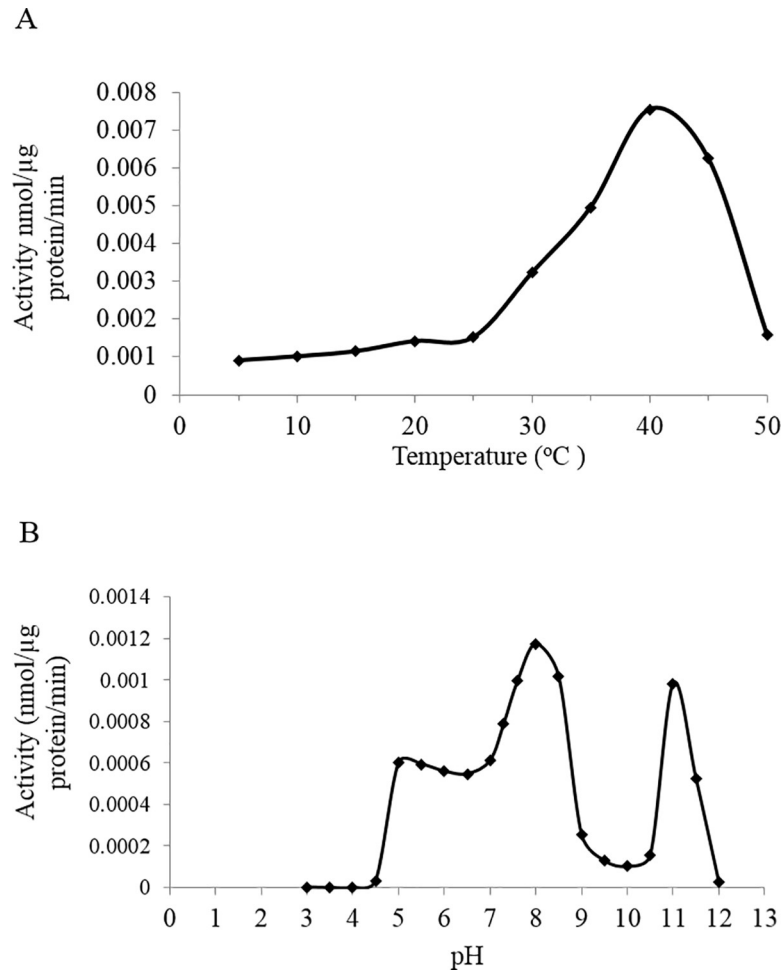


Fig 2. Enzyme kinetics of LuCGT1. (A) Temperature dependence of LuCGT1. Activity of cyanogenic glucosyltransferase was determined by the standard assay, modified by varying temperature. (B) pH dependence of LuCGT1. Activity of cyanogenic glucosyltransferase was determined by standard assay conditions as described in “Materials and methods”, modified by varying pH.

<https://doi.org/10.1371/journal.pone.0227840.g002>

glucosyltransferases often display broad substrate specificity during in vitro assays, we next conducted substrate preference assays with a broader range of potential glucosyl acceptors. High level of glucosyltransferase activity was detected with 3-hydroxypropionitrile, lactonitrile, glyconitrile and 3-hydroxybutyronitrile glycosylation (S2 Fig). The LuCGT1 also displayed significant activity towards benzyl alcohol. The spots on the TLC plate corresponding to glucoside products were scraped off, and specific activities with each glucosyl acceptor relative to that of acetone cyanohydrin (referred as 100%) were assessed. Data presented in Table 2 indicated that this newly discovered flax glucosyltransferase had a broad substrate specificity, which was consistent with previous studies on similar enzymes from other plant species [13,

Table 1. The apparent K_m and V_{max} of LuCGT1 towards acetone cyanohydrin and UDP-glucose.

| Substance | Apparent K_m | Apparent V_{max} |
|--------------------------|----------------------------|------------------------------------|
| Acetone Cyanohydrin | $61.3 \pm 4.5 \mu\text{M}$ | $0.382 \pm 0.021 \text{ nmol/min}$ |
| Uridine diphosphoglucose | $248 \pm 56 \mu\text{M}$ | $0.42 \pm 0.052 \text{ nmol/min}$ |

<https://doi.org/10.1371/journal.pone.0227840.t001>

Table 2. Broad substrate specificity of LuCGT1.

| Substrate | Activity |
|------------------------|--------------|
| salicylic acid | Non detected |
| geraniol | Trace |
| glyconitrile | 58% ± 1.8 |
| 3-hydroxypropionitrile | 69.2% ± 2.4 |
| 3-hydroxybutyronitrile | 57.3% ± 14.4 |
| benzyl alcohol | 30.53% ± 10 |
| 2-hydroxybutyronitrile | 46.64% ± 6.2 |
| lactonitrile | 59.5% ± 4.4 |
| mandelonitrile | Non detected |
| acetone cyanohydrin | 100% |

<https://doi.org/10.1371/journal.pone.0227840.t002>

19]. Due to a lack of a commercial source, we were unable to investigate if LuCGT1 possessed activity towards 2-methylbutyronitrile, which was known as the precursor for lotaustralin.

Expression profile of LuCGT1

To investigate the expression patterns of *LuCGT1* in different tissues of *Linum usitatissimum* plants, quantitative real-time PCR (qRT-PCR) was performed using RNA prepared from 4-week old, roots, stems, leaves, as well as developing seeds from cultivar CDC Bethune at 6 DAF (days after flowering) and 17 DAF (Fig 3A). The *LuCGT1* gene was expressed in all tissues examined. But the highest expression level was detected in leaves and 17 DAF developing seeds. These results were consistent with our findings that EST corresponding to *LuCGT1* was highly abundant in the developing seed and seed coat cDNA libraries [27].

Three flax mutants, UGG 102-2, UGG146-1 and Double Low, have been known to possess low level of cyanogenic glucosides [26]. Two major QTLs governing cyanogenic glucoside contents were mapped to two linkage groups localized on Chromosome 1 and Chromosome 6/13 [26]. *LuCGT1* is located on chromosome Lu10 (12,164,270 bp to 12,165,855 bp) based on latest released *Linum usitatissimum* reference genome (ASM22429v2, GenBank assembly [GCA_000224295.2]). We amplified the genomic DNA sequences corresponding to the promoter region and coding sequences of *LuCGT1* through PCR and conducted sequence comparison with its parent line, cultivar CDC Bethune. No mutations were found in the *LuCGT1* gene in any of the three mutants. We concluded that *LuCGT1* was not the underlying genetic lesion causing low cyanogenic glucoside content in these mutant lines. We then performed transcript level analysis on leaf tissues and developing seeds of the mutants and compared with that of CDC Bethune. Results from qRT-PCR show that in both leaf tissues (Fig 3B) and developing seed (Fig 3C). The transcript levels of *LuCGT1* were found to be significantly lower in the low cyanogenic glucoside flax mutant lines. These results suggest a correlation between *LuCGT1* transcript level and the accumulation of cyanogenic glucoside in flax.

Discussion

Flaxseed has a major quality shortcoming as a source of functional food due to the presence of cyanogenic glucosides. In the present study, we focused on the last step in the biosynthetic pathway of cyanogenic glucosides, the glucosylation of the cyanogenic compounds. A bioinformatics survey of putative UDP-glucosyltransferase genes in flax was previously reported [25]. However, since UDP-glucosyltransferases are known to have broad substrate specificities, a careful biochemical characterization is required to conclusively assign enzymatic function to

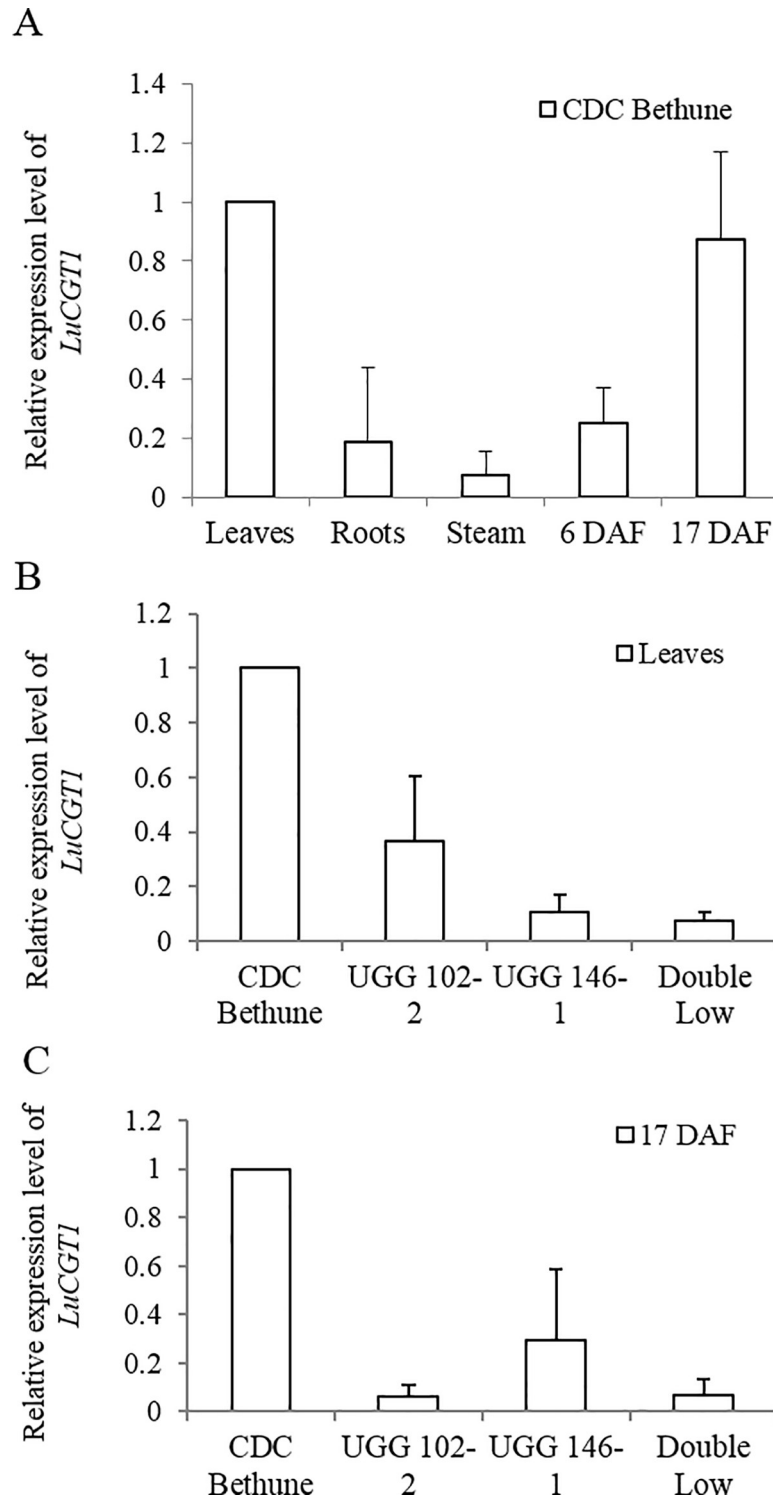


Fig 3. Tissue expression analysis of *LuCGT1* using real-time quantitative RT-PCR. (A) *LuCGT1* expression profile in flax cultivar Bethune; (B) relative expression level of *LuCGT1* in leaf tissues of low cyanogenic glucoside lines; (C) relative expression level of *LuCGT1* in developing seed tissues of low cyanogenic glucoside lines. For comparison, the expression level of *LuCGT1* from CDC Bethune leaves or seeds was normalized to 1 using StepOne software 2.0 (Applied Biosystems). The values represent the average of three independent biological replicates. Each of biological repeat contains for technical replicates. UGG102-2, UGG 146-1, Double Low are low cyanogenic flax lines.

<https://doi.org/10.1371/journal.pone.0227840.g003>

candidate genes. A number of enzymes have been isolated from plant tissues which catalyze the transfer of glucose from UDP-glucose to aliphatic or aromatic hydroxyl groups. We employed the sorghum cyanogenic glucosyltransferase [19] as a reference sequence to conduct bioinformatics search against a flax seed EST database we previously developed [27]. Based on BLAST search results, we were able to select three candidate genes. When expressed heterologously in yeast, only LuCGT1 exhibited cyanogenic glucosyltransferase activity. The other two candidate flax glucosyltransferase, UGT74S1 and UGT84G3, despite of being closely related at the deduced amino acid sequence level, displayed no detectable activity under our assaying conditions. UGT74S1 was shown to be a specialized glucosyltransferase in lignan biosynthesis [29]. Our results show that LuCGT1 is also active towards cyanohydrin intermediates and is highly expressed during seed development. LuCGT1 can utilize aromatic substrates, but it was not active towards mandelonitrile, a highly active substrate for the sorghum glucosyltransferase [22]. LuCGT1 also exhibited activities towards a few other compounds belonging to completely different class. Whether such activities have physiological relevance in flax remains to be studied. The conversions of amino acid to the corresponding oximes have been shown to be associated with microsomes [21, 28, 31] through association with the membrane bound cytochrome P450s [32, 33]. However, in our experiments using yeast lysate, enzyme activity was found to be associated with the soluble fractions of the cell homogenates, which is consistent with findings from the sorghum enzyme [32, 33].

In this study we were also interested in whether there would be mutations in *LuCGT1* in the genome of three low cyanogenic glucoside flax mutant lines. We found no mutations neither in the gene-coding regions nor promoter sequences when compared with that of cultivar CDC Bethune. We cannot rule out the possibility that there are other cyanogenic glucosyltransferase isoforms in flax. However, our RT-PCR experiments designed to assess the transcript level of this gene indicated that the low cyanogenic flax lines had relatively low expression level of *LuCGT1*, in leaves as well as in developing seeds. Hence, even though *LuCGT1* was not the genetic lesion underlying the low cyanogenic glucoside trait, its expression level likely reflected a subdued overall cyanogenic glucoside biosynthesis pathway in the mutants. A notable feature of cyanogenic glucoside synthesis in plants is that of the formation of metabolons between cytochrome P450 and glucosyltransferase. In this context, LcCGT1 will be a valuable molecular tool for probing biochemical mechanisms governing metabolic channeling in flax [31, 32]. The molecular cloning of *LuCGT1* also provides a molecular guide for targeted breeding of low cyanogenic glucosides flaxseed. A low cyanogenic glucosides and/or cyanogenic glucoside-free flax cultivar will bring the full potential and total utilization of flax as a crop, and substantially simplify the extraction procedure of a number of nutraceutical compounds from flax.

Supporting information

S1 Fig. BLAST search alignment. Thirty eight flax EST entries with significant sequence homology to the sorghum enzyme were identified using sorghum glucosyltransferase. (TIF)

S2 Fig. Substrate preference of LuCGT1. Glucosyl acceptor used in the assays were salicylic acid (lane 1), geraniol (lane 2), glyconitrile (lane 3), 3-hydroxypropionitrile (lane 4), 3-hydroxybutyronitrile (lane 5), benzyl alcohol (lane 6), 2-hydroxybutyronitrile (lane 7), lactonitrile (lane 8), acetone cyanohydrin (lane 9, and lane 11), Mandelonitrile (lane 12). Lane 10 as control without glucosyl acceptor substrates. (TIF)

S1 Table. Primers used in this study.
(DOCX)

Acknowledgments

This research was supported by the National Research Council of Canada core research program and by the TUFGEN project from Genome Prairie and Genome Canada, a not-for-profit organization that is leading's national strategy on genomics. There was no additional external funding received for this study. This is National Research Council Canada publication NRCC# 56332.

Author Contributions

Conceptualization: Jitao Zou.

Data curation: Michael Kazachkov, Qiang Li, Wenyun Shen, Raju Datla.

Formal analysis: Michael Kazachkov, Wenyun Shen, Liping Wang, Peng Gao, Daoquan Xiang, Jitao Zou.

Funding acquisition: Raju Datla, Jitao Zou.

Investigation: Michael Kazachkov, Qiang Li, Wenyun Shen, Liping Wang, Peng Gao, Daoquan Xiang.

Methodology: Michael Kazachkov, Wenyun Shen, Jitao Zou.

Project administration: Jitao Zou.

Supervision: Raju Datla, Jitao Zou.

Validation: Michael Kazachkov.

Writing – original draft: Michael Kazachkov, Jitao Zou.

Writing – review & editing: Wenyun Shen, Liping Wang, Peng Gao, Daoquan Xiang, Raju Datla, Jitao Zou.

References

1. Seigler D. Cyanide and cyanogenic glucosides. In Rosenthal G.A., Berenbaum M.R., eds, *Herbivores: Their Interaction with Secondary Plant Metabolites*. Academic Press, San Diego. 1991; 35–77.
2. Forslund K, Morant M, Jorgensen B, Olsen CE, Asamizu E, Sato S, et al. Biosynthesis of the nitrile glucosides rhodiocyanoside A and D and the cyanogenic glucosides lotaustralin and linamarin in *Lotus japonicus*. *Plant Physiol*. 2004; 135:71–84. <https://doi.org/10.1104/pp.103.038059> PMID: 15122013
3. Vetter J. Plant cyanogenic glycosides. *Toxicon* 2000; 38:11–36. [https://doi.org/10.1016/s0041-0101\(99\)00128-2](https://doi.org/10.1016/s0041-0101(99)00128-2) PMID: 10669009
4. Poulton JE. Localization and catabolism of cyanogenic glycosides. *Ciba Found Symp*. 1988; 140:67–91. <https://doi.org/10.1002/9780470513712.ch6> PMID: 3073063
5. Saunders JA, Conn EE. Subcellular localization of the cyanogenic glucoside of sorghum by autoradiography. *Plant Physiol*. 1977; 59: 647–52. <https://doi.org/10.1104/pp.59.4.647> PMID: 16659911
6. Saunders JA, Conn EE. Presence of the cyanogenic glucoside dhurrin in isolated vacuoles from sorghum. *Plant Physiol*. 1978; 61:154–157. <https://doi.org/10.1104/pp.61.2.154> PMID: 16660251
7. Swain E, Poulton JE. Utilization of Amygdalin during Seedling Development of *Prunus serotina*. *Plant Physiol*. 1994; 106:437–445. <https://doi.org/10.1104/pp.106.2.437> PMID: 12232341
8. Swain E, Li CP, Poulton JE. Tissue and Subcellular Localization of Enzymes Catabolizing (R)-Amygdalin in Mature *Prunus serotina* Seeds. *Plant Physiol*. 1992; 100:291–300. <https://doi.org/10.1104/pp.100.1.291> PMID: 16652960

9. Thayer SS, Conn, EE. Subcellular Localization of Dhurrin beta-Glucosidase and Hydroxynitrile Lyase in the Mesophyll Cells of Sorghum Leaf Blades. *Plant Physiol.* 1981; 67:617–622. <https://doi.org/10.1104/pp.67.4.617> PMID: 16661725
10. Wurtele ES, Thayer SS, Conn EE. Subcellular Localization of a UDP-Glucose:Aldehyde Cyanohydrin beta-Glucosyl Transferase in Epidermal Plastids of Sorghum Leaf Blades. *Plant Physiol.* 1982; 70:1732–1737. <https://doi.org/10.1104/pp.70.6.1732> PMID: 16662753
11. Frehner M, Scalet M, Conn EE. Pattern of the Cyanide-Potential in Developing Fruits: Implications for Plants Accumulating Cyanogenic Monoglucosides (*Phaseolus lunatus*) or Cyanogenic Diglucosides in Their Seeds (*Linum usitatissimum*, *Prunus amygdalus*). *Plant Physiol.* 1990; 94:28–34. <https://doi.org/10.1104/pp.94.1.28> PMID: 16667698
12. Hahlbrock K, Conn, EE. The biosynthesis of cyanogenic glycosides in higher plants. I. Purification and properties of a uridine diphosphate-glucose-ketone cyanohydrin beta-glucosyltransferase from *Linum usitatissimum* L. *J. Biol. Chem.* 1970; 245:917–22. PMID: 5417265
13. Cutler AJ, Hosel W, Sternberg M, Conn EE. The in vitro biosynthesis of taxiphyllin and the channeling of intermediates in *Triglochin maritima*. *J. Biol. Chem.* 1981; 256:4253–4258. PMID: 7012151
14. Koch B, Nielsen VS, Halkier BA, Olsen CE, Moller BL. The biosynthesis of cyanogenic glucosides in seedlings of cassava (*Manihot esculenta* Crantz). *Arch Biochem Biophys.* 1992; 292:141–50. [https://doi.org/10.1016/0003-9861\(92\)90062-2](https://doi.org/10.1016/0003-9861(92)90062-2) PMID: 1727632
15. Mederacke H, Bieh B, Selmar D. Characterization of two cyano glucosyltransferases from cassava leaves. *Phytochemistry* 1996; 42:1517–1522.
16. Andersen MD, Busk PK, Svendsen I, Moller BL. Cytochromes P-450 from cassava (*Manihot esculenta* Crantz) catalyzing the first steps in the biosynthesis of the cyanogenic glucosides linamarin and lotaustralin. Cloning, functional expression in *Pichia pastoris*, and substrate specificity of the isolated recombinant enzymes. *J. Biol. Chem.* 2000. 275:1966–1975. <https://doi.org/10.1074/jbc.275.3.1966> PMID: 10636899
17. Kannangara R, Motawia MS, Hansen NK, Paquette SM, Olsen CE, Moller BL, et al. Characterization and expression profile of two UDP-glucosyltransferases, UGT85K4 and UGT85K5, catalyzing the last step in cyanogenic glucoside biosynthesis in cassava. *Plant J.* 2011; 68:287–301. <https://doi.org/10.1111/j.1365-3113X.2011.04695.x> PMID: 21736650
18. Halkier BA, Moller BL. The biosynthesis of cyanogenic glucosides in higher plants. Identification of three hydroxylation steps in the biosynthesis of dhurrin in *Sorghum bicolor* (L.) Moench and the involvement of 1-ACI-nitro-2-(p-hydroxyphenyl) ethane as an intermediate. *J. Biol. Chem.* 1990; 265:21114–21121. PMID: 2250015
19. Jones PR, Moller BL, Hoj PB. The UDP-glucose:p-hydroxymandelonitrile-O-glucosyltransferase that catalyzes the last step in synthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor*. Isolation, cloning, heterologous expression, and substrate specificity. *J. Biol. Chem.* 1999; 274:35483–35491. <https://doi.org/10.1074/jbc.274.50.35483> PMID: 10585420
20. MacFarlane IJ, Lees EM, Conn EE. The in vitro biosynthesis of dhurrin, the cyanogenic glycoside of *Sorghum bicolor*. *J. Biol. Chem.* 1975; 250:4708–4713. PMID: 237909
21. Moller BL, Conn EE. The biosynthesis of cyanogenic glucosides in higher plants. Channeling of intermediates in dhurrin biosynthesis by a microsomal system from *Sorghum bicolor* (linn) Moench. *J. Biol. Chem.* 1980; 255:3049–3056. PMID: 7358727
22. Thorsoe KS, Bak S, Olsen CE, Imberty A, Breton C, Lindberg Moller B. Determination of catalytic key amino acids and UDP sugar donor specificity of the cyanohydrin glycosyltransferase UGT85B1 from *Sorghum bicolor*. Molecular modeling substantiated by site-specific mutagenesis and biochemical analyses. *Plant Physiol.* 2005; 139:664–673. <https://doi.org/10.1104/pp.105.063842> PMID: 16169969
23. Collinge DB, Hughes MA. In vitro characterization of the Ac locus in white clover (*Trifolium repens* L.). *Arch. Biochem. Biophys.* 1982; 218:38–45. [https://doi.org/10.1016/0003-9861\(82\)90318-6](https://doi.org/10.1016/0003-9861(82)90318-6) PMID: 7149740
24. Hosel W, Schiel O. Biosynthesis of cyanogenic glucosides: in vitro analysis of the glucosylation step. *Arch. Biochem. Biophys.* 1984; 229:177–86. [https://doi.org/10.1016/0003-9861\(84\)90142-5](https://doi.org/10.1016/0003-9861(84)90142-5) PMID: 6230992
25. Barvkar VT, Pardeshi VC, Kale SM, Kadoo NY, Gupta VS. Phylogenomic analysis of UDP glycosyltransferase 1 multigene family in *Linum usitatissimum* identified genes with varied expression patterns. *BMC Genomics.* 2012; 13:175. <https://doi.org/10.1186/1471-2164-13-175> PMID: 22568875
26. Chin-Fatt A. QTL mapping and NIRS estimation of cyanogenic glucosides in flaxseed. M.Sc. thesis. University of Manitoba. 2014.
27. Venglat P, Xiang D, Qiu S, Stone SL, Tibiche C, Cram D, et al. Gene expression analysis of flax seed development. *BMC Plant Biol.* 2011; 11:74. <https://doi.org/10.1186/1471-2229-11-74> PMID: 21529361

28. Cutler AJ, Sternberg M, Conn EE. Properties of a microsomal enzyme system from *Linum usitatissimum* (linen flax) which oxidizes valine to acetone cyanohydrin and isoleucine to 2-methylbutanone cyanohydrin. *Arch Biochem Biophys.* 1985; 238:272–279. [https://doi.org/10.1016/0003-9861\(85\)90165-1](https://doi.org/10.1016/0003-9861(85)90165-1) PMID: 3985623
29. Ghose K, Selvaraj K, McCallum J, Kirby CW, Sweeney-Nixon M, Cloutier SJ, et al. Identification and functional characterization of a flax UDP-glycosyltransferase glucosylating secoisolariciresinol (SECO) into secoisolariciresinol monoglucoside (SMG) and diglucoside (SDG) *BMC Plant Biology* 2014; 14:82. <https://doi.org/10.1186/1471-2229-14-82> PMID: 24678929
30. Caputi L, Malnoy M, Goremykin V, Nikiforova S, Martens S. A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land. *Plant J* 2002; 69:1030–1042.
31. Koch BM, Sibbesen O, Halkier BA, Svendsen I, Møller BL. The primary sequence of cytochrome P450^{tyr}, the multifunctional N-hydroxylase catalyzing the conversion of L-tyrosine to p-hydroxyphenylacetaldehyde oxime in the biosynthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor* (L.) Moench. *Arch. Biochem. Biophys.* 1995; 323:177–186. <https://doi.org/10.1006/abbi.1995.0024> PMID: 7487064
32. Nielsen KA, Tattersall DB, Jones PR, Møller BL. Metabolon formation in dhurrin biosynthesis. *Phytochemistry* 2008; 69:88–98. <https://doi.org/10.1016/j.phytochem.2007.06.033> PMID: 17706731
33. Laursen T, Borch J, Knudsen C, Bavishi K, Torta F, Martens HJ, et al. Characterization of a dynamic metabolon producing the defense compound dhurrin in sorghum. *Science* 2016; 354:890–893. <https://doi.org/10.1126/science.aag2347> PMID: 27856908