

# Identification and Characterization of Abiotic Stress Induced novel UDP-Glucosyltransferase (UGT72L11) Gene from *Glycyrrhiza glabra* L.

**Shahnawaz Hussain**

CSIR-Indian Institute of Integrative Medicine: Council of Scientific & Industrial Research Indian Institute of Integrative Medicine

**Bhawna Verma**

CSIR-Indian Institute of Integrative Medicine: Council of Scientific & Industrial Research Indian Institute of Integrative Medicine

**Malik Muzafar Manzoor**

CSIR-Indian Institute of Integrative Medicine: Council of Scientific & Industrial Research Indian Institute of Integrative Medicine

**Pooja Goyal**

CSIR-Indian Institute of Integrative Medicine: Council of Scientific & Industrial Research Indian Institute of Integrative Medicine

**Ritu Devi**

CSIR-Indian Institute of Integrative Medicine: Council of Scientific & Industrial Research Indian Institute of Integrative Medicine

**Ajai Prakash Gupta**

CSIR-Indian Institute of Integrative Medicine: Council of Scientific & Industrial Research Indian Institute of Integrative Medicine

**Manoj kumar Dhar**

University of Jammu

**Fariha chowdhary**

CSIR-Indian Institute of Integrative Medicine: Council of Scientific & Industrial Research Indian Institute of Integrative Medicine

**Suphla Gupta**

`sgupta@iiim.res.in`

CSIR-Indian Institute of Integrative Medicine: Council of Scientific & Industrial Research Indian Institute of Integrative Medicine <https://orcid.org/0000-0001-6688-5575>

**Keywords:** ABA, enzymatic assay, Flavonoid, Glycyrrhiza glabra, in-vitro stress, UDP glycosyltransferase

**Posted Date:** March 19th, 2024

**DOI:** <https://doi.org/10.21203/rs.3.rs-3981251/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

The present study reports a unique broad spectrum UDP-glycosyltransferase from *Glycyrrhiza glabra* involved in multiple stress responses and abscisic acid mediated glycosylation. The identified *UGT72L11* gene was cytoplasmic with ORF of 1425 bp encoding a 52.2 kDa protein of 474 amino acids. Phylogenetic analysis revealed maximum homology (73.3%) with epicatechin 3-glucosyltransferase (ACC38470) from *Medicago truncatula* exhibiting sequence uniqueness. The gene was differentially expressed in shoot tissues and significantly upregulated in abscisic acid treatment (122.3 folds) and under cold stress (36 folds) *in planta*. *In-silico* Structure-Activity-Relationship revealed GLU279, ARG386, PRO380 and TRP379 residues being involved in receptor-ligand interactions. The UGT72L11 protein was optimal between 10°C to 30°C preferring quercetin-UDPGlc ( $K_m$  0.23) over kaempferol-UDPGlc ( $K_m$  0.47). The purified recombinant protein showed multi-substrate O-glycosylation towards various classes of aglycones, abscisic acid, and also displayed C-glycosylation with colchicine as a foundation for the future medicinal applications.

## 1. Introduction

Plant UDP dependent glycosyltransferases (UGTs) belongs to a large multi-gene superfamily of glucosyltransferases (GTs) that catalyses the transfer of UDP-activated sugars to diverse set of aglycons forming corresponding glycol-conjugates [1] as observed in secondary metabolites, hormones [2], xenobiotics [3] and several other small molecules. The glycosylation of secondary metabolites effectively regulate their properties and functions imparting altered characteristics [4]. Hence the chemical modification plays critical role in the response to plant growth and adaptation to biotic and abiotic stresses. Studies have highlighted role of plant UGTs in stress management. The rice *UGT85E1* has been demonstrated to mediate plant response to drought and oxidative stresses [5]. Similarly, ectopic over expression of *UGT76E11* from *Arabidopsis thaliana* showed enhanced tolerance to salinity and drought [6]. On the other hand, silencing of *UGT75C1* from tomato played crucial role in ABA-mediated fruit ripening, seed germination, and drought responses [7]. Higher accumulation of glycosylated specialized metabolites modified by UGTs have shown to be involved in drought and salt stress tolerance and higher oxidative stress management in plants (Li et al., 2018c, 2018d; Gharibi et al., 2019). Chill induced expression of *OsUGT90A1* from *Oryza sativa* imparted tolerance to low temperature through cell membrane integrity [8]. Glycosyltransferases have shown specialized metabolites modulations. The UGTs like *CrUGT87A*, identified from *Carex rigescens* exhibited increased flavonoid accumulation and salt stress resistance [9] and *CsUGT78A15* from *Camellia sinensis* was shown to be involved in biosynthesis of eugenol glucoside under cold stress [10] as was also demonstrated in *A. thaliana AtUGT87A2* associated with salinity, osmotic stress, drought and ABA [11]. Multiple stress inducible command further insight into the response mechanism in plants to understand the interplay between the UGT catalyzed glycosylation under the influence of stress, and specialized metabolite accumulation. The crystal structure of the di-C-glycosyltransferase from *G. glabra* catalyzing a two-step di-C-glycosylation of flopropione-containing substrates revealed the role of hydrogen-bond interactions of sugar hydroxyl

groups in sugar donor selectivity and space-chemistry in di-C-glycosylation capability [12]. This finding has immense application in developing efficient biocatalysts to synthesize C-glycosides of medicinal potential. Understanding of the intricate interplay between UGTs in plants is needed to unlock the full potential of UGTs in plant adaptive behaviour under climate change scenario and also as a prospective catalytic tool.

In recent years putative UGT genes have been identified from several plant species, but their functional characterization is equitably small, coming predominantly from studies conducted on *A. thaliana* [13]. Further, only a few members of the UGT72 subfamily have been characterized from Fabaceae as in the seed coat-specific UGT72L1 from *M. truncatula* [14] and UGT72L6 from *Lotus japonicas* (KT895087). In the Glycyrrhiza genus, literature reports few UGTs identified and characterized from *G. uralensis* species for terpenoids [15], saponin [16] and flavonoid [17] biosyntheses. A recent study on metabolic profiling in *G. uralensis* identified few key differentially expressed UGT genes co-regulating flavonoid and saponin biosyntheses in licorice under salt stress [18].

The present study mined a novel *UGT* gene from the transcriptome resource of *G. glabra* species having wide-substrate acceptability and involvement in multiple stresses. The purified recombinant protein was enzymatically characterized and *in-planta* assessed. Additionally, the isolated recombinant protein was studied for various kinetic parameters and *in-vitro* evaluated for substrate promiscuity utilizing diverse class of aglycons.

## 2. Materials and methods

### 2.1 Plant material

Plant material was collected from the experimental farm of the Council for Scientific and Industrial Research-Indian Institute of Integrative Medicine Jammu (32.73°N and 74.87°E) for RNA extraction and as explants for tissue culture. The regenerated plants were used throughout the experiments for various treatments. All tissues were frozen in liquid nitrogen (N) and stored at – 80°C until required.

### 2.2 Chemicals

Reagents used in the study were purchased from Takara Bio-USA, Promega, Thermofisher and ThermoScientific. Substrates and Isopropyl β-D-1-thiogalactopyranoside (IPTG) for enzyme assay were purchased from Sigma-Aldrich (Oakville, CA, USA). All the chemicals used in this study were of molecular/analytical/HPLC/MS grade.

### 2.3 Retrieval of UGTs from the transcriptomic data

Transcriptome data (PRJNA664636) generated through Next Generation Sequencing (NGS) of *G. glabra* plant was used as resource for the identification of UGT gene from the species using the Hidden Markov Model (HMM) profile (Pfam family: PF00201). The local similarity search (tBLASTN) was performed for the mining of *GgUGTs* using the BioEdit with an E-value cut off of  $10^{-3}$ . The obtained contigs were

further filtered by ORF finder (<https://www.ncbi.nlm.nih.gov/orffinder/>) tool to select full-length sequences and translated by ExPasy (<https://web.expasy.org/>) tool. The molecular weight (MW), theoretical isoelectric point (pI), instability index, aliphatic index, grand average of hydropathicity (GRAVY) were predicted *via* the ProtParam (<http://web.expasy.org/protparam/>). Additionally, sub-cellular localisation prediction was performed by an advanced protein sub-cellular localisation prediction tool CELLO2GO (<http://cello.life.nctu.edu.tw/cello2go/>).

## 2.4 Sequence alignment and phylogenetic analysis

The protein sequence of the identified UGTs were aligned with the sequences of the functionally characterized proteins from other plant species of the same subfamily, and proteins having mono- and disaccharide glycoside forming activities [19]. Multiple sequence alignment (MSA) of the deduced amino acid sequences was carried out using Clustal omega [20] and DNAMAN(<https://dnaman.software.informer.com/7.0/>). To construct the phylogenetic tree, amino acid sequences were aligned using the BLOSUM62 matrix with the ClustalW algorithm-based AlignX module from Mega MEGA (Ver7.0) to generate a neighbour joining tree with bootstrapping (1000 replicates) analysis and handling gaps with pair wise deletion. The mono/di-glucoside-based categorization was done by the Geneious alignment tool (<https://www.geneious.com/>) to construct phylogenetic tree using neighbour joining method in default parameters of the Geneious program. The conserved motifs of the UGTs were predicted using the Multiple Expectation Maximization for Motif Elicitation (MEME:<http://meme-suite.org/tools/meme>) tool having maximum ten motifs with minimum 6 and maximum 50 motif width.

## 2.5 In-silico 3D model construction

The three-dimensional protein structure of UGT72L11 protein was predicted by I-TASSER (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) and validated by Ramachandran plot. Ligand binding site was predicted using 3DLigandSite [21]. The ligand structure of kaempferol was downloaded from PubChem Compound Database (National Centre for Biotechnology Information; <https://pubchem.ncbi.nlm.nih.gov/>). For primary molecular docking, visualization and modifying receptor and ligand structures the DockThor-VS web online server (<https://dockthor.lncc.br/v2/>) was used (Discovery Studio Biovia 2017; Dassault Systèmes, San Diego, California, USA). Discovery Studio Biovia 2017 was also used for post-docking analyses, like prediction of the size and location of binding site, hydrogen-bond interactions, hydrophobic interactions, and bonding distances.

## 2.6 Cellular localisation:

Sub-cellular localisation of the gene was performed in onion cells using *Agrobacterium* mediated transformation having 35S:*UGT72L11*-GFP construct in PCAMBIA1302 vector following the published protocol [22]. Briefly, healthy and fresh onion scales (1–1.5×1 cm) were placed on a petridish with re suspension of *Agrobacterium* solution (OD600) having sucrose, acetosyringone and Silwet-for 24 hours at 28°C. Subsequently, onion scales were co-cultivated with *Agrobacterium* containing ½ strength Murashige and Skoog basal medium. GFP imaging of transient expression was visualised under

confocal fluorescence microscope (Olympus FLUOVIEW FV1000). For GFP detection, the excitation source was an argon-ion laser at 488 nm, and emission was observed between 510 and 530 nm.

## 2.7 RNA isolation, gene and promoter cloning

Total RNA having purity between 1.9–2.1 (A260/280) was isolated using Pure Link RNA Isolation kit (ThermoFisher). Extracted RNA (2 µg) was reverse-transcribed to cDNA using primeScriptIII First-Strand cDNA kit (ThermoFisher). The full-length amplification of the *UGT72L11* gene was performed using gene specific primers (Table S1). The amplified gene was cloned into pJET1.2 (ThermoScientific, USA) cloning vector, confirmed by Sanger sequencing and subsequently submitted to NCBI Data Bank. The corresponding amino acid sequence encoding the candidate UGT was protein-BLAST searched for homology using NCBI search tools [23]. For promoter cloning, the 5' upstream region of the identified UGT was cloned using Genome Walker kit (TAKARA, Japan) following the instruction manual. The amplified promoter region was scanned for the presence of various cis-regulatory elements using Plant Cis-acting Regulatory DNA Elements database (PLACE, <http://www.dna.affrc.go.jp/PLACE/>).

## 2.8 Real-time expression analysis

### 2.7.1 Treatments

The *in-vitro* regenerated plants grown under three different environmental conditions namely, tissue culture, hardening unit and field conditions were assessed for the expression of the identified gene. The *G. glabra* plants were also subjected to one biotic stress, eight abiotic stress treatments and three hormone based elicitations. The treatments were essentially performed following the standardized protocol published earlier [24, 25]. Each treatment along with control was given to a set of three *in-vitro* plants (2 months old) for different time durations. Abiotic treatments like dark and carbon starvation were subjected for 48 hrs; cold (4°C) for 64 hrs; wounding for 8 hrs; UV-C for 30 min, salinity for 128 hrs and drought for 4 days. For the senescence treatment, green and yellow leaves of the *in-vitro* plants were used. For biotic elicitation 3 pathogenic isolates of *Colletotrichum* sp. and one *Calonectria* sp. were employed. The hormone treatment was subjected for various time duration; NAA was given for 2 hrs, GA<sub>3</sub> for 6 hrs and ABA for 16 hrs to the *in-vitro* grown plants. RNA extraction and cDNA synthesis from the tissues of the control and treated plants, and plants grown under three different conditions (*in-vitro*, field and glass house) were performed as mentioned earlier.

### 2.8.2 Expression Analysis

The expression of *UGT72L11* transcripts in the aerial tissues of *G. glabra* plant were examined by quantitative reverse-transcription PCR following the thermal profile published earlier using *β-actin* gene as an internal control [26]. The primers for quantitative real-time (qRT)-PCR are listed in Table S1. qRT-PCR was performed on StepOne™ (Applied Biosystems) using SYBR premix ExTaq II (Clontech) following instructions of the manufacturer. The relative expression level was calculated by the 2<sup>-ΔΔct</sup> method [27]. For each gene, three independent experiments were performed, and three technical replicates were

analyzed for each sample [26]. Statistical analysis was performed using Graph Pad prism software/EXCEL sheet.

## 2.9 Heterologous expression and protein purification

The full-length gene in pJET1.2 (ThermoScientific, USA) cloning vector was amplified using unique *Nco*I and *Xho*I site specific forward and reverse primers, respectively. The amplified gene was digested with respective restriction enzymes and sub-cloned into previously digested pET28a(+) vector with the same set of enzymes. The resultant construct was transformed into *Escherichia coli* BL21 (DE3) cells and the transformed colonies were screened on the Kanamycin (50mg/ml) supplemented plates and confirmed by colony PCR. PCR amplification was performed using gene specific forward/reverse primers (supplementary table1) and T7 promoter/terminator primer combination using the plasmid extracted from the positive colony for the conformity of the transformation. For the recombinant protein expression and purification, the confirmed transformant was inoculated into 5 ml Luria Bertani (LB) containing appropriate antibiotic and incubated at 37°C overnight. Subsequently, the primary inoculum (1%) was added to 200 ml fresh LB media and incubated similarly, until its OD<sub>600</sub> reached a value of 0.6. The cloned gene was induced with the optimized conditions of IPTG concentration (0.8 mM), incubating temperature (22°C) and under shaking (180 rpm) for 20 hrs. The recombinant *E. coli* cells were harvested and the recombinant protein was extracted by cell lysis followed by centrifugation (6500 g for 10 min at 4°C). The crude protein (20 µl) was initially checked on SDS-PAGE (10%) to confirm protein induction. Subsequently the recombinant protein was purified using TALON single step column (Takara, Japan) essentially following the manufacturer's protocol. The purified recombinant protein (20 µl) was confirmed on SDS-PAGE gel (10%) and quantified using Bradford's method[28].

## 2.10 Enzyme Kinetics

Prior to the *in-vitro* assay using aglycon compound library, optimization of the reaction conditions for the recombinant protein was performed. The purified protein was optimized for parameters like incubation time (15–90 min), temperature (4°C–70°C) and pH in citrate (pH 3–6), phosphate (6–7.5) and Tris-HCl (7.5–10.5) buffers at a constant protein concentration (2.5 µg). The sugar acceptability using kaempferol and quercetin as substrates was assessed essentially following the UDP Glo™ assay protocol (Promega, USA) which detects UDP after UDP-sugar hydrolysis/ transfer by converting UDP to light in luciferase type reaction (measured in relative units). A standard curve of UDP concentration (0–25 µM) was generated, and linear range of detection was determined, where the luminescence is directly proportional to the UDP concentration. Following the protocol published earlier [29], each sugar nucleotide hydrolysis reaction was combined with the detection reagent in equal ratio and allowed to incubate at room temperature. After 1h of incubation, the luminescence was measured using FLUOstar Omega microplate reader (BMG Labtech). The amount of UDP released during the glycosylation reaction was recorded and the relative catalytic activity was calculated. Also, four substrates (kaempferol and quercetin) concentrations (3, 5, 7 & 9 µM) were examined and the released UDP amounts were quantified by relative luminescence for the enzyme activity. The enzyme kinetics parameter namely, Km, Vmax, kcat and kcat/Km were calculated for the two substrates.

## 2.11 In vitro enzyme assay

Enzyme reaction (200 µl) was performed under the *in-vitro* conditions using the optimized parameters of citrate buffer (50 mM) pH 6.0, purified protein (5µg), UDP-Glc (5 mM) and substrate (0.9 mM) at 30°C incubation temperature for 15 min. The reaction was quenched by the addition of ethyl acetate (100 µl) thrice and partitioned with water. The organic part was concentrated and re-suspended in methanol for UP-HPLC/LCMS analysis. After successful glycosylation of kaempferol and quercetin, a panel of sixteen natural products belonging to various classes of compounds (flavones, terpenes, alkaloids, phenolics, vitamins, Phyto-hormone and carboxylic acid) and drug molecules were tested under similar conditions.

## 2.12 LC-MS Analysis

Experiments were performed on an Agilent 1260 Infinity (Agilent, USA) HPLC system coupled with an Agilent 6410 (Agilent Technologies, USA) triple quadrupole MS/MS instrument equipped with an ESI ion source in the positive ion mode. Chromatographic separation was performed on a Chromolith high resolution RP18e column(100 x 4.6 mm) maintained at 300° C. Mobile phases of 0.1% (v/v) formic acid in water (eluent A) and acetonitrile (eluent B) were used at a flow rate of 600 µl/min to analyze each sample having injection volume of 10µl. A gradient programme used was: 0–7 min, 50%B followed by increment of 10% in eluent B after 8–12 min, 12–25 min, 25–30 min reaching upto80% for30–45 min and returning to 50% for 45–50 min. Data handling was performed using Mass Hunter workstation. The standards (kaempferol/quercetin/3-O-glycosylated kaempferol and quercetin) were analyzed under the optimized conditions before analysing the samples.

## 3. Results

### 3.1 Sequence and phylogenetic analysis

The selected full-length gene encoding a UGT gene was mined from the *in-house* library (PRJNA664636) of *G. glabra* plant and submitted to the NCBI database (MN163014.1) bearing 52.16 kDa protein (QHW04706.1). The full-length *GgUGT72L11* ORF was sequenced to be 1425bp encoding a cytoplasmic protein of 474 amino acid having instability index of 40.62, aliphatic index of 99.75, GRAVY value of 0.032 and a theoretical pI of 6.61. The MEME based motif analysis showed ten highly conserved fifty sequence length motifs (including the signature Plant Secondary Product Glycosyltransferase motif) common to all the studied protein homologs (Fig. 1). The MSA based homology of the UDPGT specific region ranged between 72.61% 85.90% to (supplementary Fig. 1). A 199 aa long GTB region (from 262 to 461 aa) was found to be highly variable (from 305 to 321aa); however, a 69 aa long region (from amino acid 322 to 391) in this region was highly conserved including the 44 aa long PSPG box (aa344 to aa388) at the C-terminal. Further, the protein-protein blast of the 58 aa long stretch was found to be highly specific to the proteins identified from Fabaceae family which was endorsed by the results of the MSA (Supplementary Fig. 1). Further, to predict the putative function of UGT72L11, twenty five functionally annotated UGT72 protein sequences reported from various plant species, were aligned

(Fig. 1). The phylogeny classified the identified protein akin to epicatechin 3-glucosyltransferase (UGT72L1) from *M. truncatula* and flavonol 3-O-glucosyltransferases from *Glycine max* (UGT72Z3), *G. uralensis* (UGT72X5) and *G. max.* (UGT72X4) plants displaying affinity towards flavonoids.

Additionally, phylogenetic tree (Fig. 2) based on thirty four proteins sequences having mono- and di-glycosylation propensity revealed the identified protein to be mono-glycosylated forming a distinct cluster with AtUGT71C1 and VvGT7A (XP\_002276546) corroborated by mono-glucoside characteristic GSS consensus motif (563–566) as reported by Huang et al.[19]. Contrary to this finding, the proline amino acid at position 95 (Supplementary Fig. 2) which is the characteristic feature of disaccharides was also present in the GgUGT72L11 protein further highlighting the sequence uniqueness of the identified protein.

## 3.2 Promoter Analysis

To further investigate the role of the identified gene, upstream region (1174 bp) of the selected gene was isolated by genome walking. *In silico* analysis of the promoter sequence revealed putative *cis*-acting regulatory elements and transcription factor-binding sites. The results demonstrated that the mapped *cis*-acting elements included promoter core elements (TATA-box and CAAT-box), phytochrome activated light-responsive element (5'GATAAGR3'), abscisic acid signalling/regulated gene expression motif (5'CACGTG/CANNTG3'). Stress responsive elements including dehydration (5'ACGT3'), cold induced (5'CAANTG3') and ABRE-Related Sequences (MACGYGB) were also found. The promoter region also had carbon (5'AAAG3') and sugar metabolism (5'TGACT3') motifs, pathogenesis related elements (5'YTGTCWC3') and wound (5'TGACY3') perceptive elements (supplementary Fig. 3).

## 3.33-D Model Prediction

In plants, the characteristic 44 aa long consensus sequence of the PSPG domain plays a crucial role during glycosylation [30]. The 10 aa highly conserved C-terminal domain binds to the donor sugar molecule, while the N-terminal is highly variable acceptor binding site potentially for accommodating wide variety of aglycons. Ramachandran plot analysis of protein under study, (supplementary Fig. 4a,b) showed maximum number of amino acids in favored region (97.40%) and allowed region (1.94%). A small fraction (0.65%) was found in the outlier region (supplementary Fig. 4b) also. The glycosylation can only be catalyzed if the acceptor molecule is positioned correctly in the substrate pocket such that the functional group of the sugar acceptor is in close proximity to the first carbon of the sugar for the formation of glycosidic bond [31]. Docking simulations of candidate ligand kaempferol with receptor UGT72L11 showed good binding affinity and stability of the complex as indicated by binding affinity of kaempferol (-6.140 kcal/mol), total energy (-31.247 kcal/mol), vdW energy (-0.438 kcal/mol) and electrostatic energy (-16.390 kcal/mol).

## 3.4 Subcellular localisation, recombinant Protein Expression and Purification

Sub-cellular localization of UGT72L11 protein was found to be predominantly in plasma membrane as demonstrated by onion cells transiently transformed with 35S:ugt72l11-GFP construct (Fig. 4). The recombinant protein was expressed in *E. coli* (BL-21 cells) and optimized for optimal induction. The protein was found to be optimally induced (Fig. 5) in 20 hr under 0.8 mM IPTG concentration at 20°C. The recombinant protein with C-terminal HIS tag was column purified and employed for enzyme kinetic studies and *in-vitro* assays.

### 3.5 Enzyme Kinetics

The purified recombinant protein activity was assessed in three different pH -acidic (citrate), neutral (phosphate) and basic (Tris HCl) buffer systems. The protein was found to be active in all the three buffers studied. However, optimum enzyme activity was found in citrate buffer (pH6.0) (Fig. 6a). Further, the time duration of 15 min to 90 min assessed for the enzyme activity showed its optimal activity at 15 min (Fig. 6b). The effect of temperature studied between 4°C to 70°C, revealed protein to be optimally active between 10°C to 30°C (Fig. 6c). The optimized parameters were further used for the enzyme kinetic studies. The enzyme-substrate relationship was evaluated with two substrates (quercetin and kaempferol) at four concentrations (3, 5, 7 & 9 µM) employing UDP-glucose as the sugar donor (Fig. 6d). The kinetic properties were determined for the selected substrates using previously optimized concentrations of 2.5µg protein and 15 µM UDP sugar. The reaction was kept for 15-min at 30°C. The Michaelis constant ( $K_m$ ), which signifies the concentration of the substrate at which half of the enzyme is saturated, was found to be more for quercetin-UDPGlc (0.23) than kaempferol-UDPGlc (0.47) and the catalytic number ( $K_{cat}$ )/ $K_m$  were calculated to be 0.66 and 0.44 mm<sup>-1</sup> s<sup>-1</sup>, respectively reflecting its preference for quercetin over kaempferol.

### 3.6 In-vitro enzyme activity and product identification

The substrate affinity of the purified recombinant protein was established using *in-vitro* enzyme assay employing various classes of substrates as acceptors and UDP-glucose as the sugar donor following the optimized conditions of 30°C for 15 min. The protein was initially assessed with quercetin and kaempferol as the two model substrates as they harbour several hydroxyl groups. The *in-vitro* enzyme activity and the position of glycosylation were confirmed by their respective 3-O-glucoside standards. The retention time (RT) corresponding to the substrates and respective glycosylated products were compared and confirmed by peaks and mass spectra. It was conclusively shown that the substrates, kaempferol (Fig. 7a) and quercetin (Fig. 7b) were glycosylated to quercetin-3-glucoside and kaempferol-3-glucoside, respectively.

### 3.7 In-vitro activity of GgUGT72L11 Enzyme (Substrate scope)

The substrates investigated in the present study belonged to flavones, terpenes, carboxylic acids, vitamins, alkaloid, phenol classes of compounds and few commercial drugs (Fig. 8). Flavones, such as quercetin and catechin hydrate having polyhydroxyl groups, when subjected to glycosylation under optimized conditions showed the formation of respective glucosides as confirmed by LCMS (supplementary Data 1; Table 1). Similarly, diosgenin having monohydroxy group and andrographolide with carboxylic groups were catalyzed for the addition of glucose moiety to form respective glycosides by the purified protein (supplementary Data 1). Isovanillin, (a phenolic), dihydrozeatin & abscisic acid (phytohormones) were also seen to form their respective glycosides catalyzed by GgUGT72L11. The alkaloidal substrates studied were colchicine, monocrotaline and quinine hydrochloride. The identified GgUGT72L11 catalyzed glycosylation in colchicine whereas monocrotaline and quinine hydrochloride underwent hydroxyl glycosylation forming respective glycosides. Structurally, colchicine possess no hydroxyl and/or carboxyl groups, contrary to monocrotaline and quinine hydrochloride that had -OH groups

Table 1

List of glycosylated compounds (flavanols, terpenes, carboxylic acids, vitamins, alkaloid, phenolics and commercial drugs) with molecular formulae observed, and calculated molecular weight of the glycosides

| S.no | Product                         | Mass calculated for                    | Calculated molecular weight | Observed molecular weight |
|------|---------------------------------|----------------------------------------|-----------------------------|---------------------------|
| 1    | Kaempferol glucoside            | (M-H) ( $C_{21}H_{19}O_{11}$ )         | 448.10                      | 448.02                    |
| 2    | Quercetin glucoside             | (M + H)<br>( $C_{21}H_{21}O_{12}$ )    | 465.10                      | 464.65                    |
| 3    | Catechin hydrate glucoside      | (M-H) ( $C_{21}H_{25}O_{12}$ )         | 469.14                      | 469.83                    |
| 4    | Diosgenin glucoside             | (M + H)<br>( $C_{33}H_{53}O_8$ )       | 577.37                      | 577.35                    |
| 5    | Andrographolide glucoside       | (M-H) ( $C_{26}H_{39}O_{10}$ )         | 513.26                      | 514.88                    |
| 6    | Citric acid glucoside           | (M-H) ( $C_{12}H_{17}O_{12}$ )         | 353.07                      | 353.26                    |
| 7    | Nicotinic acid glucoside        | (M-H)<br>( $C_{12}H_{14}NO_7$ )        | 284.08                      | 285,25                    |
| 8    | Riboflavin glucoside            | (M-H)<br>( $C_{23}H_{29}N_4O_{11}$ )   | 537.18                      | 537.32                    |
| 9    | Iso vanillin glucoside          | (M-H) ( $C_{14}H_{17}O_8$ )            | 313.09                      | 313.51                    |
| 10   | Dihydro zeatin glucoside        | (M + H)<br>( $C_{16}H_{26}N_5O_6$ )    | 384.19                      | 383.8                     |
| 11   | Absciscic acid                  | (M-H)<br>( $C_{20}H_{29}O_9$ )         | 425.45                      | 425.35                    |
| 12   | Colchicine glucoside            | (M + Na)<br>$C_{28}H_{35}NNaO_{11}$ )  | 585.21                      | 585.67                    |
| 13   | Quinine hydrochloride glucoside | (M-H)<br>( $C_{26}H_{34}ClN_2O_7$ )    | 521.21                      | 520.9                     |
| 14   | Monocrotaline glucoside         | (M + H)<br>( $C_{22}H_{34}NO_{11}$ )   | 488.21                      | 488.18/486.6              |
| 15   | Podophyllotoxin glucoside       | (M + H)<br>( $C_{28}H_{33}O_{13}$ )    | 577.19                      | 576.94                    |
| 16   | Sofosbuvir glucoside            | (M-H)<br>( $C_{28}H_{38}FN_3O_{14}P$ ) | 690.21                      | 690.42                    |
| 17   | Ephedrine HCL glucoside         | (M-H)<br>( $C_{16}H_{25}ClNO_6$ )      | 362.14                      | 361.99                    |

| S.no | Product                         | Mass calculated for                                                          | Calculated molecular weight | Observed molecular weight |
|------|---------------------------------|------------------------------------------------------------------------------|-----------------------------|---------------------------|
| 13   | Quinine hydrochloride glucoside | (M-H)<br>(C <sub>26</sub> H <sub>34</sub> ClN <sub>2</sub> O <sub>7</sub> )  | 521.21                      | 520.9                     |
| 14   | Monocrotaline glucoside         | (M + H)<br>(C <sub>22</sub> H <sub>34</sub> NO <sub>11</sub> )               | 488.21                      | 488.18/486.6              |
| 15   | Podophyllotoxin glucoside       | (M + H)<br>(C <sub>28</sub> H <sub>33</sub> O <sub>13</sub> )                | 577.19                      | 576.94                    |
| 16   | Sofosbuvir glucoside            | (M-H)<br>(C <sub>28</sub> H <sub>38</sub> FN <sub>3</sub> O <sub>14</sub> P) | 690.21                      | 690.42                    |
| 17   | Ephedrine HCL glucoside         | (M-H)<br>(C <sub>16</sub> H <sub>25</sub> ClNO <sub>6</sub> )                | 362.14                      | 361.99                    |

## 3.8 Expression Dynamics of UGT72L11

### 3.8.1 Growth Conditions

qRT-PCR analysis was performed to analyze the expression profile and levels of the identified gene in different tissues. The real-time quantitative analysis of the transcripts was performed using the RNA extracted from the aerial and underground tissues on the plants grown under three different conditions namely, tissue culture, hardening unit and the experimental farms (Fig. 9a). The identified gene displayed preferential expression under the three surroundings. The *GgUGT11* was found to be significantly up-regulated (2.41 folds) in the roots of the field grown plants, while it was significantly down regulated in the roots of the glass house stationed plants as compared to the respective shoot tissues (Fig. 9a). The gene was found to be similarly expressed in the stem tissues under the three conditions.

### 3.8.2 Stress

Total RNA extracted from the aerial tissues of the *in-vitro* grown plants was subjected to nine stressors and assessed for transcript accumulation. Eight abiotic treatments (Fig. 9b) given to the plants were dark, carbon starvation, wounding, senescence, drought, salinity, cold and UV light. The *GgUGT72L11* gene was found to be repressed under the influence of carbon starvation (0.6 fold), salinity (0.9 folds) and dark (0.3 fold) conditions. The gene transcripts were found to be maximally up regulated under cold (36 folds) conditions followed by drought (9 folds), senescence (1.7 folds), and UV light (1.2 folds). Further, the biotic stress subjected through co-cultivation of *in-vitro* grown *G. glabra* plants (individually) with four plant pathogens (3 isolates of *Colletotrichum* sp. (P1-*Colletotrichum gloeosporioides*, P2- *Colletotrichum siamense sub sp.*, P3- *Colletotrichum siamense sub sp.* and P4- one isolate of *Calonectria cylindrospora*) showed increased transcript levels with P2 (2.86 folds), while P1 was found to be suppressed (Fig. 9c). The expression profile of the *GgUGT72L11* gene in the *in-vitro* plants subjected to auxin, abscisic acid and GA hormone treatments showed significant reverberation. The gene was found to be highly up regulated under ABA (400 folds) treatment at 16 hrs (Fig. 9d) and moderately up regulated in the

presence of auxin (2 folds) between 30 to 60 minutes. At the same time, it was seen to be down regulated (0.4 to 0.8) under GA treatment between 1 to 6 hrs in the aerial tissues of the plant as compared to the untreated control plants.

### 3.8.3 Determination of expression of ROS scavenging genes and cold responsive genes

The above results indicated higher accumulation of transcripts in the aerial tissues of the *in-vitro* grown plants under cold (4°C) and ABA stresses. Further, cold -specific COR genes (*COR47*, *ABI5*, *DREB1*) and Reactive Oxygen Species scavenging genes catalase (*CAT1*), super oxide dismutase (*SOD*), Glutathione Peroxidase (*GPX*) were studied to understand the molecular mechanism (Fig. 10A). The cold responsive genes were seen to be significantly up regulated-*COR 47* (22.3 folds), *DREB1* (13.8 folds) and *ABI5* (10.2 folds) under cold stress. Also expression analysis of ROS scavenging enzymes and cold responsive genes was carried out in ABA treated plants (Fig. 10B). Both the ROS scavenging genes -*CAT1* ( 5.97 folds), *SOD* (18.6 folds) and *GPX* (16 folds), and cold responsive genes-*DREB1* (2.9 folds) and *ABI5* (14.6folds) were found to be up regulated in the plants subjected to ABA treatments suggesting the role of *GgUGT72L11* under multiple stress. However, ROS scavenging genes (*CAT1*, *SOD*, *GPX*) were observed to be uninfluenced under the cold stress suggesting no involvement of ROS scavenging enzymes under cold stress.

## Discussion

The pBLAST search (NCBI) displayed maximum homology (73.3%) homology with UGT72L1 (ACC38470) encoding epicatechin 3-glucosyltransferase from *M. truncatula*. The CAZy database divulged information on only two proteins from *G. uralensis* species (NCBI accession No. MK341791 and MK341793) which were involved in flavonoid biosynthesis [17]. The identified protein from *G. glabra* seems to be unique as it showed less than 40% homology with the two proteins reported from the *G. uralensis*. The uniqueness of the protein was further enhanced by having both mono- and disaccharide glycoside forming activities. The protein sequence possessed mono-glucoside characteristic GSS consensus motif (563–566), and also had proline amino acid at position 95 characteristic of disaccharides as reported by Huang et al.[19].

Understanding the glycosylation mechanism and its physiological significance in plants hold importance not only *in-planta* for channelizing metabolic pathways for producing the desired compound [32]; but also for valuable glycosides in synthetic medical biology. In our study, the *in-silico* mapping on receptor (UGT72L11) and ligand (kaempferol) showed that kaempferol formed five hydrogen bond interactions with UGT72L11 protein specifically with four amino acid residues GLU279, ARG386, PRO380 and TRP379 (Fig. 3). On the other hand, in the well characterized structure of AtUGT72B1 protein the substrate was demonstrated to be enclosed in six hydrophobic residues (I-86, L-118, F-119, F-148, L-183, and L-197) to form catalytic site roofed by E83 covering the site[33]. Further, the ligand and receptor interaction was also stabilized by pi donor hydrogen bond and pi alkyl interactions involving SER277 and LEU409 amino acid residues respectively. Protein-ligand binding occurs spontaneously when the binding affinity/free

energy change is negative [34] as these complexes are stable which is an essential characteristic of effective ligands. The present study further experimented to confirm the results of docking experiment with the *in-vitro* enzyme activity results using the substrates. *In vitro* assay using the purified protein demonstrated multi-substrate acceptability with wide range of aglycons and medicinally important drugs like ephedrine hydrochloride, podophyllotoxin, and sofosbuvir (Table 1 and supplementary Data1). The GgUGT72L11 protein introduced glucose moiety to the mono hydroxyl group in the studied compounds forming their respective mono-glycosylated products as confirmed by LCMS analyses (supplementary Data 1). Previous *in-vitro* enzymatic studies in thirty reported recombinant UGT72 family proteins, across 12 different angiosperm species, have shown acceptance to varied substrates including flavonoids, monolignols, and their precursors/derivatives [35]. UGTs can glycosylate at *O*-position (3, 5, 7, 3',4', or 5'), and/or *C*-positions as seen in rice, maize and buckwheat (Du et al., 2010; Falcone Ferreyra et al., 2013; Nagatomo et al., 2014). Evidences suggest members of UGT72 family possess broad spectrum substrate recognition predominantly in flavanols with different glycosylation patterns forming 3-*O*-monoglycosides as in *L. japonicus* (LjUGT72AH1, LjUGT72Z2, LjUGT72V3) [36], *Glycine max* (GmUGT72X4, GmUGT72Z3) [37] and *M. truncatula* (MtUGT72L1) [14], to multi-site glycosylation at 3, 4, and 7 positions by *Clonorchis sinensis* CsUGT72A1 [38] and *Hieracium pilosella* HpUGT72B11 [39].

The identified GgUGT72L11 protein also catalyzed *C*-glycosylation in colchicine. The *C*-glycosylation pattern and its extent have a crucial role in forming potent drug-like molecule [12]. The conjugates of *C*-glycosides exhibit unique characteristics as these bonds are resistant to acid hydrolysis and glycosidase cleavage [40]. The UGTs catalyzing *C*-glycosylation has been reported from several plants including *Fagopyrum esculentum* (FeCGTa and FeCGTb) [41], *Glycine max* (UGT708D1), *Dendrobium catenatum* [40], *Gentiana triflora* [42] and Citrus plants (UGT708G1, UGT708G2) [43]. Literature cites singular report of *C*-glycosyltransferase (GgCGT) from *G. glabra* catalysing two-step *C*-glycosylation of the drug flopropione [12].

The presence of light, hormonal and stress responsive motifs indicated involvement of gene in complex hormone regulatory network pertaining to stress and secondary metabolite biosyntheses. Overall, the *cis*-element analysis suggested gene responsive to different kinds of stresses. This was explored further to assess the involvement of the identified UGT under the influence of various stresses. The identified flavonoid glycosyl transferase displayed high expression under cold, senescence, drought and abscisic acid stress. Recent comprehensive review [44] highlighted buffering effects of UGTs in multiple biotic and abiotic stresses. A series of protective mechanisms are triggered when plants sense stress [45] underpinning the overlapping but not completely redundant biological functions of UGTs in mediating developmental and stress responses.

In the current study, *GgUGT72L11* gene was shown to be significantly up regulated under cold stress along with the downstream chill-induced *COR* and *DREB* genes. However, ROS scavenging genes were down regulated under similar conditions. Contrary to this studies have highlighted up regulation of ROS-scavenging genes such as superoxide dismutase (*SOD*), catalase (*CAT 1*) and Glutathione Peroxidase (*GPX*) in detoxification particularly under stress [46]. One probable reason for down regulation could be

the formation of glycosylated flavonoids catalyzed by the identified *flavonol glycosyl transferase* (*GgUGT72L11*) imparting hydrophilic and cartable [47] properties to counter the oxidative burst more efficiently under stressed conditions [45]. Our findings also revealed the gene was significantly induced when subjected to ABA treatment and catalyses ABA to respective glycosylated conjugate (supplementary data) under *in-vitro* conditions. Hence, the ABA homeostasis was maintained thereby inhibiting ROS accumulation. Corroborating our findings, few studies have confirmed strong induction of UGTs under the cold stress simultaneously participating in flavonol biosyntheses [45] ,[48]. Studies have also confirmed involvement of several UGTs (AtUGT79B2, AtUGT79B3 and CsUGT78A14) in glycosylation of ABA regulating the dynamic state of cellular ABA/ABA-GE levels (ABA homeostasis).

## Conclusion

Studies targeting association of flavonoids under cold stress and ROS scavenging ability will provide novel insights into the biological role of the UGTs and flavonoids in plant defence and stress biology. Further, investigation into the underlying molecular processes involving ABA homeostasis and ABA-mediated stresses (cold and drought in particular) will throw light on the interplay between stressors, glycosylated metabolites and signalling mechanisms for multiple stress tolerance. The recognition of a simple but unique glycosylation mechanism in the identified UGT offers physiological advantage to the plant and could be the foundation for numerous applications in medical biology expanding the magnitude to manipulate the biological and physicochemical properties of molecules. Further investigation of the underlying downstream processes will elucidate the regulatory mechanism of *UGT72L11* in plant stress tolerance.

## Declarations

### Contribution:

Malik Muzafar Manzoor, Shahnawaz Husain, Bhawna Verma, Suphla Gupta - Conceptualization, Malik Muzafar Manzoor, Shahnawaz Husain, Bhawna Verma ,Pooja Goyal, Ritu Devi, 4Fariha chowdhary- Methodology :, Ajai P Gupta ,Pooja Goyal, Bhawna Verma, Suphla Gupta, Bhawna Verma - Writing- Original draft preparation. Suphla Gupta,: Visualization, Investigation and Supervision.: Ajai P Gupta: Software, Validation.: Suphla Gupta, Bhawna Verma- Writing- Reviewing and Editing.

### Funding and Acknowledgements:

Authors acknowledge the grant (SERB/SB/SO/PS/90/2013) from Science and Engineering Research Board, India for funding the study. MM and SH acknowledges Council for Scientific & Industrial Research for Fellowships; PG and BV acknowledges University Grant Commission for Fellowship; SH & BV acknowledges Council for Scientific & Industrial Research and Academy of Scientific and Innovative Research, CSIR-HRDC Campus, Sector-19, Ghaziabad, U.P, India is also acknowledged. All the authors acknowledge director, CSIR-IIIM for providing all the necessary facilities.

**Data Statement:** The article's supporting data and materials are in the public domain and can be accessed at NCBI SRA Data (PRJNA664636).

### **Declaration of competing interest**

All the authors declare no competing interest.

### **Declaration of generative AI in scientific writing**

We have not used any generative AI in scientific writing.

## **References**

1. Bock KW (2016) The UDP-glycosyltransferase (UGT) superfamily expressed in humans, insects and plants: Animal-plant arms-race and co-evolution. *Biochem Pharmacol* 99:11–17
2. Gilbert MK, Bland JM, Shockey JM, Cao H, Hinchliffe DJ, Fang DD, Naoumkina M (2013) A transcript profiling approach reveals an abscisic acid-specific glycosyltransferase (UGT73C14) induced in developing fiber of Ligon lintless-2 mutant of cotton (*Gossypium hirsutum* L). *PLoS ONE* 8(9):e75268
3. Zhao Y, Chow TF, Puckrin RS, Alfred SE, Korir AK, Larive CK, Cutler SR (2007) Chemical genetic interrogation of natural variation uncovers a molecule that is glycoactivated. *Nat Chem Biol* 3(11):716–721
4. Jarrar Y, Lee S-J (2021) The Functionality of UDP-Glucuronosyltransferase Genetic Variants and their Association with Drug Responses and Human Diseases. *J Pers Med* 11(6):554
5. Zheng S, Luo J, Zhu X, Gao X, Hua H, Cui J (2022) Transcriptomic analysis of salivary gland and proteomic analysis of oral secretion in *Helicoverpa armigera* under cotton plant leaves, gossypol, and tannin stresses. *Genomics* 114(2):110267
6. Li Q, Yu H-M, Meng X-F, Lin J-S, Li Y-J, Hou B-K (2018) Ectopic expression of glycosyltransferase UGT76E11 increases flavonoid accumulation and enhances abiotic stress tolerance in *Arabidopsis*. *Plant Biol* 20(1):10–19
7. Sun Y, Ji K, Liang B, Du Y, Jiang L, Wang J, Kai W, Zhang Y, Zhai X, Chen P et al (2017) Suppressing ABA uridine diphosphate glucosyltransferase (SlUGT75C1) alters fruit ripening and the stress response in tomato. *Plant J* 91(4):574–589
8. Shi Y, Phan H, Liu Y, Cao S, Zhang Z, Chu C, Schläppi MR (2020) Glycosyltransferase OsUGT90A1 helps protect the plasma membrane during chilling stress in rice. *J Exp Bot* 71(9):2723–2739
9. Zhang K, Sun Y, Li M, Long R (2021) CrUGT87A1, a UDP-sugar glycosyltransferases (UGTs) gene from *Carex rigescens*, increases salt tolerance by accumulating flavonoids for antioxidation in *Arabidopsis thaliana*. *Plant Physiol Biochem* 159:28–36
10. Zhao M, Cai B, Jin J, Zhang N, Jing T, Wang J, Pan Y, Zhou Z, Zhao Y, Feng Y et al (2020) Cold Stress-induced Glucosyltransferase CsUGT78A15 is Involved in the Formation of Eugenol Glucoside in

*Camellia sinensis*. *Hortic Plant J* 6(6):439–449

11. Li P, Li Y-j, Wang B, Yu H-m, Li Q, Hou B (2017) -k: The Arabidopsis UGT87A2, a stress-inducible family 1 glycosyltransferase, is involved in the plant adaptation to abiotic stresses. *Physiol Plant* 159(4):416–432
12. Zhang M, Li FD, Li K, Wang ZL, Wang YX, He JB, Su HF, Zhang ZY, Chi CB, Shi XM et al (2020) Functional Characterization and Structural Basis of an Efficient Di-C-glycosyltransferase from *Glycyrrhiza glabra*. *J Am Chem Soc* 142(7):3506–3512
13. Caputi L, Malnoy M, Goremykin V, Nikiforova S, Martens S (2012) 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* 69(6):1030–1042
14. Pang Y, Cheng X, Huhman DV, Ma J, Peel GJ, Yonekura-Sakakibara K, Saito K, Shen G, Sumner LW, Tang Y et al (2013) Medicago glucosyltransferase UGT72L1: potential roles in proanthocyanidin biosynthesis. *Planta* 238(1):139–154
15. Nomura Y, Seki H, Suzuki T, Ohyama K, Mizutani M, Kaku T, Tamura K, Ono E, Horikawa M, Sudo H et al (2019) Functional specialization of UDP-glycosyltransferase 73P12 in licorice to produce a sweet triterpenoid saponin, glycyrrhizin. *Plant J* 99(6):1127–1143
16. Chung SY, Seki H, Fujisawa Y, Shimoda Y, Hiraga S, Nomura Y, Saito K, Ishimoto M, Muranaka T (2020) A cellulose synthase-derived enzyme catalyses 3-O-glucuronosylation in saponin biosynthesis. *Nat Commun* 11(1):5664
17. Jiang D, Li P, Yin Y, Ren G, Liu C (2021) Molecular cloning and functional characterization of UGTs from *Glycyrrhiza uralensis* flavonoid pathway. *Int J Biol Macromol* 192:1108–1116
18. Wang C, Chen L, Cai Z, Chen C, Liu Z, Liu S, Zou L, Tan M, Chen J, Liu X et al (2021) Metabolite Profiling and Transcriptome Analysis Explains Difference in Accumulation of Bioactive Constituents in Licorice (*Glycyrrhiza uralensis*) Under Salt Stress. *Front Plant Sci* 12:727882
19. Huang F-C, Giri A, Daniilidis M, Sun G, Härtl K, Hoffmann T, Schwab W (2018) Structural and Functional Analysis of UGT92G6 Suggests an Evolutionary Link Between Mono- and Disaccharide Glycoside-Forming Transferases. *Plant Cell Physiol* 59(4):862–875
20. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R (2013) Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res* 41(W1):W597–W600
21. Wass MN, Kelley LA, Sternberg MJ 3DLigandSite: predicting ligand-binding sites using similar structures. *Nucleic Acids Res* 2010:gkq406
22. Sun W, Cao Z, Li Y, Zhao Y, Zhang H (2007) A simple and effective method for protein subcellular localization using *Agrobacterium*-mediated transformation of onion epidermal cells. *Biologia* 62(5):529–532
23. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410
24. Manzoor MM, Goyal P, Pandotra P, Dar MS, Dar MJ, Misra P, Gupta AP, Vishwakarma RA, Ahuja A, Dhar MK Transcriptome-wide identification of squalene epoxidase genes from *Glycyrrhiza glabra* L.:

- expression analysis and heterologous expression of GgSQE1 suggest important role in terpenoid biosynthesis. *Protoplasma* 2021:1–17
25. Goyal P, Muzafar Manzoor M, Gupta AP, Pandotra P, Gupta S (2022) : Molecular Dissection of Genes and Promoters involved in Glycyrrhizin Biosynthesis Revealed Phytohormone Induced Modulation in *Glycyrrhiza glabra* L. *Gene* :146682
  26. Li Y, Liang X, Zhou X, Wu Z, Yuan L, Wang Y, Li Y (2020) : Selection of Reference Genes for qRT-PCR Analysis in Medicinal Plant *Glycyrrhiza* under Abiotic Stresses and Hormonal Treatments. *Plants* (Basel Switzerland) 9(11)
  27. Schmittgen TD, Zakrajsek BA, Mills AG, Gorn V, Singer MJ, Reed MW (2000) Quantitative Reverse Transcription-Polymerase Chain Reaction to Study mRNA Decay: Comparison of Endpoint and Real-Time Methods. *Anal Biochem* 285(2):194–204
  28. Bradford M A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 7: 248–254. *Find this article online* 1976.
  29. McGraphery K, Schwab W (2020) : Comparative Analysis of High-Throughput Assays of Family-1 Plant Glycosyltransferases. *Int J Mol Sci* 21(6)
  30. Louveau T, Leitao C, Green S, Hamiaux C, van der Rest B, Dechy-Cabaret O, Atkinson RG, Chervin C (2011) Predicting the substrate specificity of a glycosyltransferase implicated in the production of phenolic volatiles in tomato fruit. *FEBS J* 278(2):390–400
  31. Breton C, Šnajdrová L, Jeanneau C, Koča J, Imberty A (2005) Structures and mechanisms of glycosyltransferases. *Glycobiology* 16(2):29R–37R
  32. Guo H, Wang H, Huo YX (2020) Engineering Critical Enzymes and Pathways for Improved Triterpenoid Biosynthesis in Yeast. *ACS Synth Biol* 9(9):2214–2227
  33. Brazier-Hicks M, Offen WA, Gershater MC, Revett TJ, Lim E-K, Bowles DJ, Davies GJ, Edwards R (2007) : Characterization and engineering of the bifunctional N-and O-glucosyltransferase involved in xenobiotic metabolism in plants. *Proceedings of the National Academy of Sciences* 104(51):20238–20243
  34. Afriza D, Suriyah W, Ichwan S (2018) : In silico analysis of molecular interactions between the anti-apoptotic protein survivin and dentatin, nordentatin, and quercetin. In: *Journal of Physics: Conference Series* : IOP Publishing; 2018: 032001
  35. Speeckaert N, El Jaziri M, Baucher M, Behr M (2022) : UGT72, a Major Glycosyltransferase Family for Flavonoid and Monolignol Homeostasis in Plants. *Biology* (Basel) 11(3)
  36. Yin Q, Shen G, Chang Z, Tang Y, Gao H, Pang Y (2016) Involvement of three putative glucosyltransferases from the UGT72 family in flavonol glucoside/rhamnoside biosynthesis in *Lotus japonicus* seeds. *J Exp Bot* 68(3):597–612
  37. Noguchi A, Saito A, Homma Y, Nakao M, Sasaki N, Nishino T, Takahashi S, Nakayama T (2007) A UDP-glucose:isoflavone 7-O-glucosyltransferase from the roots of soybean (*glycine max*) seedlings. Purification, gene cloning, phylogenetics, and an implication for an alternative strategy of enzyme catalysis. *J Biol Chem* 282(32):23581–23590

38. Zhao X, Dai X, Gao L, Guo L, Zhuang J, Liu Y, Ma X, Wang R, Xia T, Wang Y (2017) Functional Analysis of an Uridine Diphosphate Glycosyltransferase Involved in the Biosynthesis of Polyphenolic Glucoside in Tea Plants (*Camellia sinensis*). *J Agric Food Chem* 65(50):10993–11001
39. Speeckaert N, Adamou NM, Hassane HA, Baldacci-Cresp F, Mol A, Goeminne G, Boerjan W, Duez P, Hawkins S, Neutelings G et al (2020) : Characterization of the UDP-glycosyltransferase UGT72 Family in Poplar and Identification of Genes Involved in the Glycosylation of Monolignols. *Int J Mol Sci* 21(14)
40. Ren Z, Ji X, Jiao Z, Luo Y, Zhang G-Q, Tao S, Lei Z, Zhang J, Wang Y, Liu Z-J et al (2020) Functional analysis of a novel C-glycosyltransferase in the orchid *Dendrobium catenatum*. *Hortic Res* 7(1):111
41. Nagatomo Y, Usui S, Ito T, Kato A, Shimosaka M, Taguchi G (2014) Purification, molecular cloning and functional characterization of flavonoid C-glucosyltransferases from *Fagopyrum esculentum* M. (buckwheat) cotyledon. *Plant J* 80(3):437–448
42. Sasaki N, Nishizaki Y, Yamada E, Tatsuzawa F, Nakatsuka T, Takahashi H, Nishihara M (2015) Identification of the glucosyltransferase that mediates direct flavone C-glucosylation in *Gentiana triflora*. *FEBS Lett* 589(1):182–187
43. Ito T, Fujimoto S, Suito F, Shimosaka M, Taguchi G (2017) C-Glycosyltransferases catalyzing the formation of di-C-glucosyl flavonoids in citrus plants. *Plant J* 91(2):187–198
44. Zhang W, Wang S, Yang J, Kang C, Huang L, Guo L (2022) Glycosylation of plant secondary metabolites: Regulating from chaos to harmony. *Environ Exp Bot* 194:104703
45. Zhao M, Jin J, Gao T, Zhang N, Jing T, Wang J, Ban Q, Schwab W, Song C (2019) : Glucosyltransferase CsUGT78A14 Regulates Flavonols Accumulation and Reactive Oxygen Species Scavenging in Response to Cold Stress in *Camellia sinensis*. *Front Plant Sci* 10
46. Suzuki N (2005) Reactive oxygen species and temperature stresses: A delicate balance between signaling and destruction. *Physiol Plant* 126(1):45–51
47. Song C, Gu L, Liu J, Zhao S, Hong X, Schulenburg K, Schwab W (2015) Functional Characterization and Substrate Promiscuity of UGT71 Glycosyltransferases from Strawberry (*Fragaria x ananassa*). *Plant Cell Physiol* 56(12):2478–2493
48. Chen X (2023) : BcMYB111 Responds to BcCBF2 and Induces Flavonol Biosynthesis to Enhance Tolerance under Cold Stress in Non-Heading Chinese Cabbage. *Int J Mol Sci* 24(Issue 10).

## Figures

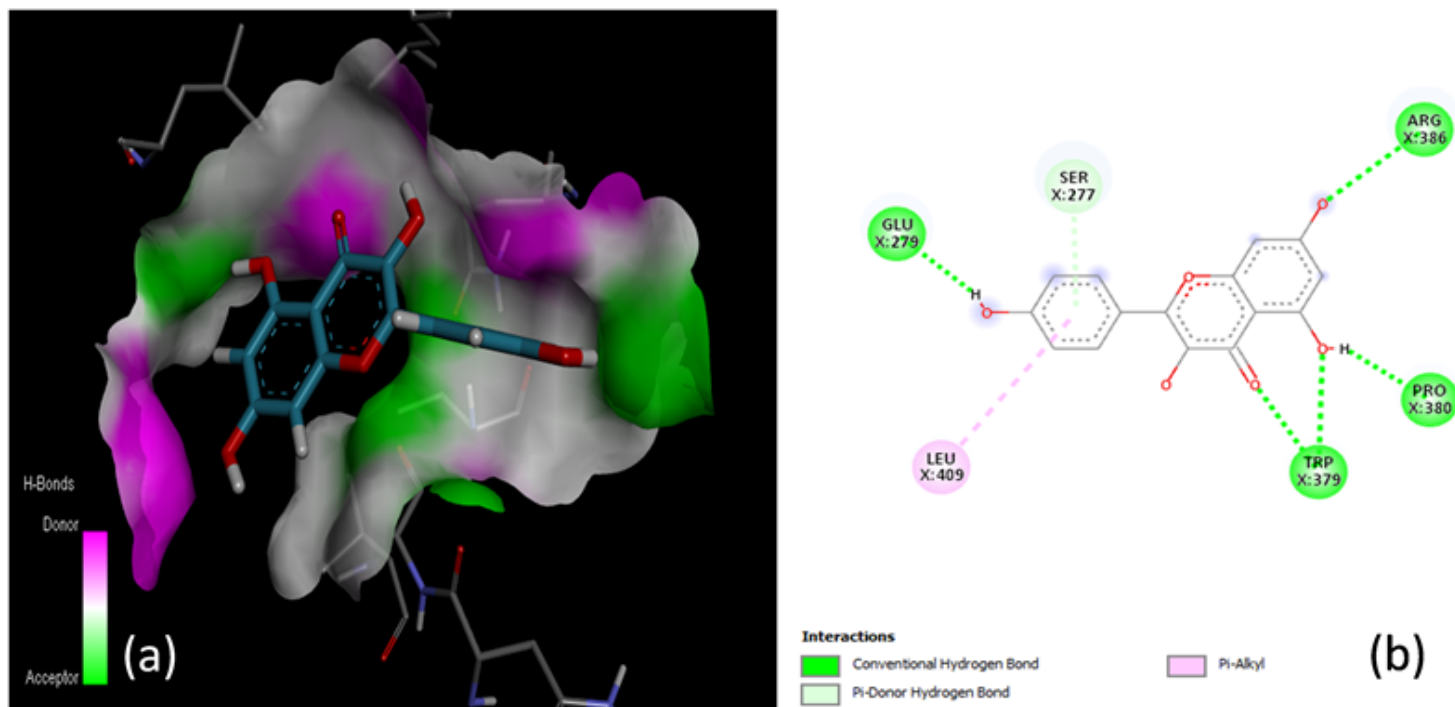


**Figure 1**

Phylogenetic tree and MEME of deduced amino acid sequence of GgUGT72L11 (pink color) along with 25 UGT proteins of UGT72 subfamily identified from other plant species (green color) displaying homology with UGT72L1 (ACC38470) epicatechin 3-glucosyltransferase from Medicago truncatula. GenBank accession numbers for the sequences are shown in parentheses: GgUGT72L11 (QHW04706.1) Hp\_UGT72B11(ACB56923.1); UGT728B2 (AAL06646.2); UGT727A5(BAD52007.1); UGT725D3 (BAG32255.1); UGT724B2 (AAK53551.1); UGT723F9 (ACM66950.1); UGT723A3 (BAD06874.1); UGT721B8 (Q8S3B6); UGT720D4 (ABA18631.1); UGT720A5 (CAA50376.1); UGT72Z3 (XP\_003532192.3); UGT72X5 (QDM38910.1); UGT72X4 (XP\_003532193.1); UGT72L1 (ACC38470); UGT72G15 (NpUGT6); UGT72G4 (MK644229.1); UGT72E8 (liUGT1); UGT72E2 (NP\_201470.1); UGT72C1 (CAB16822.1); UGT72BD1 (QHB92369.1); UGT72B42 (QDM38905.1); UGT72B29 (5NLM); UGT72B17 (6JEM); UGT72B14 (ACD87062.1); UGT72B1 (2VCE); Pt\_UGT72B29(AUO17143.1).

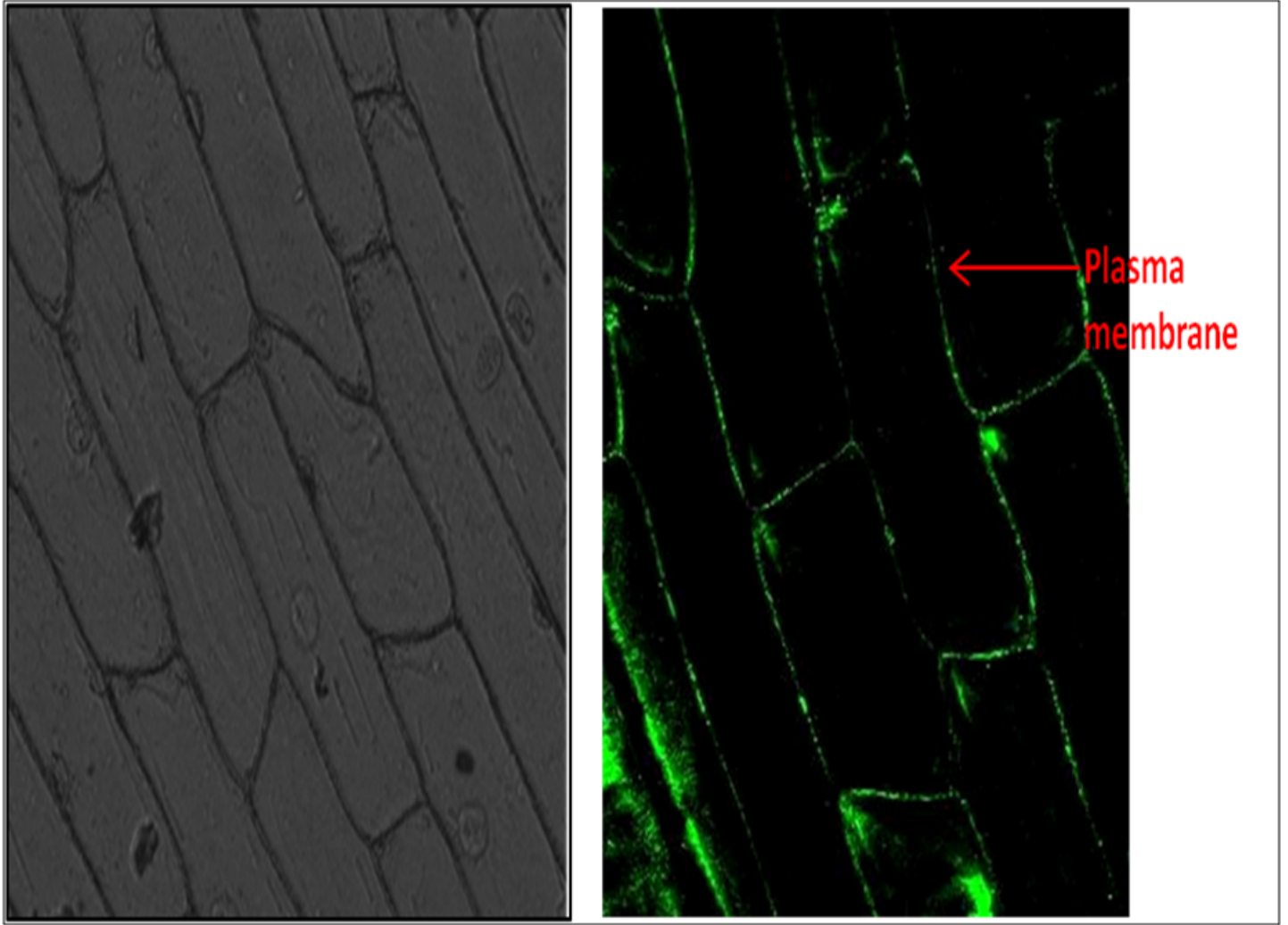


Cm, *Citrus maxima*; Cr, *Catharanthus roseus*; Cs, *Camellia sinensis*; Csi, *Citrus sinensis*; Gm, *Glycinemax*, Ip, *Ipomoea purpurea*; Pg, *Panax ginseng*; Ph, *Petunia hybrida*; Si, *Sesamum indicum*; Sl, *Solanum lycopersicum*; Vv, *Vitis vinifera*. The bar represents 0.10 amino acid substitutions per site.



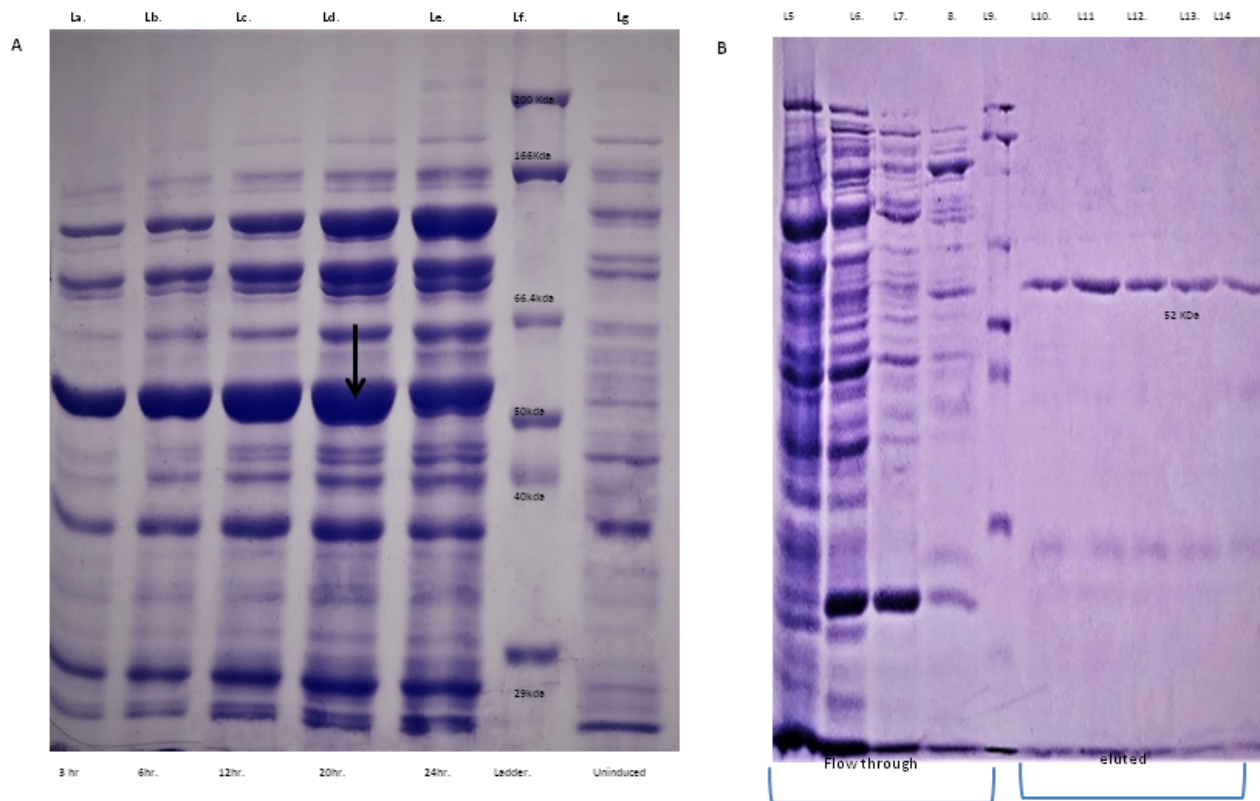
**Figure 3**

Binding conformation and interaction of substrate (kaempferol) with receptor protein (GgUGT72L11). (a) Ribbon 3D model showing binding orientation of ligand (Kaempferol) in the active site of receptor (GgUGT72L11); (b) Showing interactions of ligand (kaempferol) with amino acid residues of ligand GgUGT72L11.



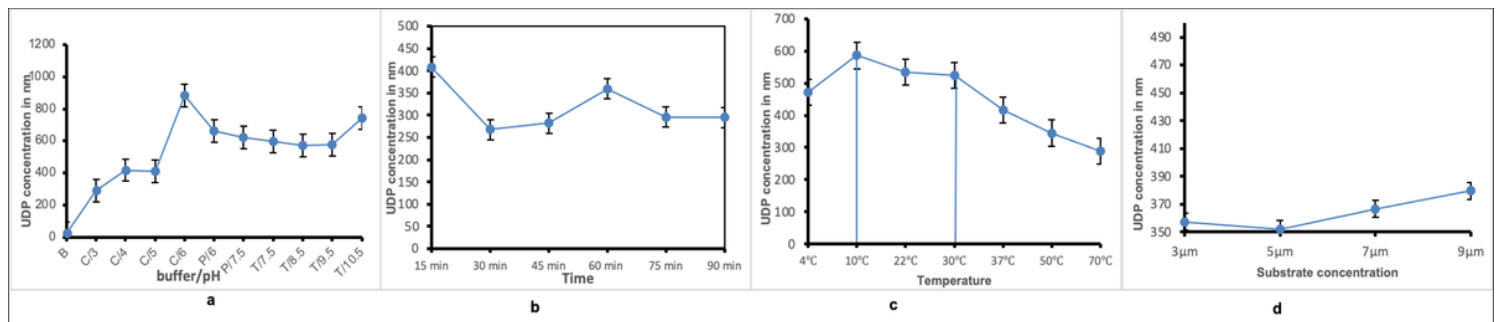
**Figure 4**

Subcellular localization of transiently expressed UGT72L11 in onion epidermal cells using GFP tagging . The transformed onion cells were obtained after introduction of 35S:ugt72l11-GFP through the *Agrobacterium*- mediated transformation. (a) Images of cells in white light (b) in fluorescence photographed by confocal microscopy.



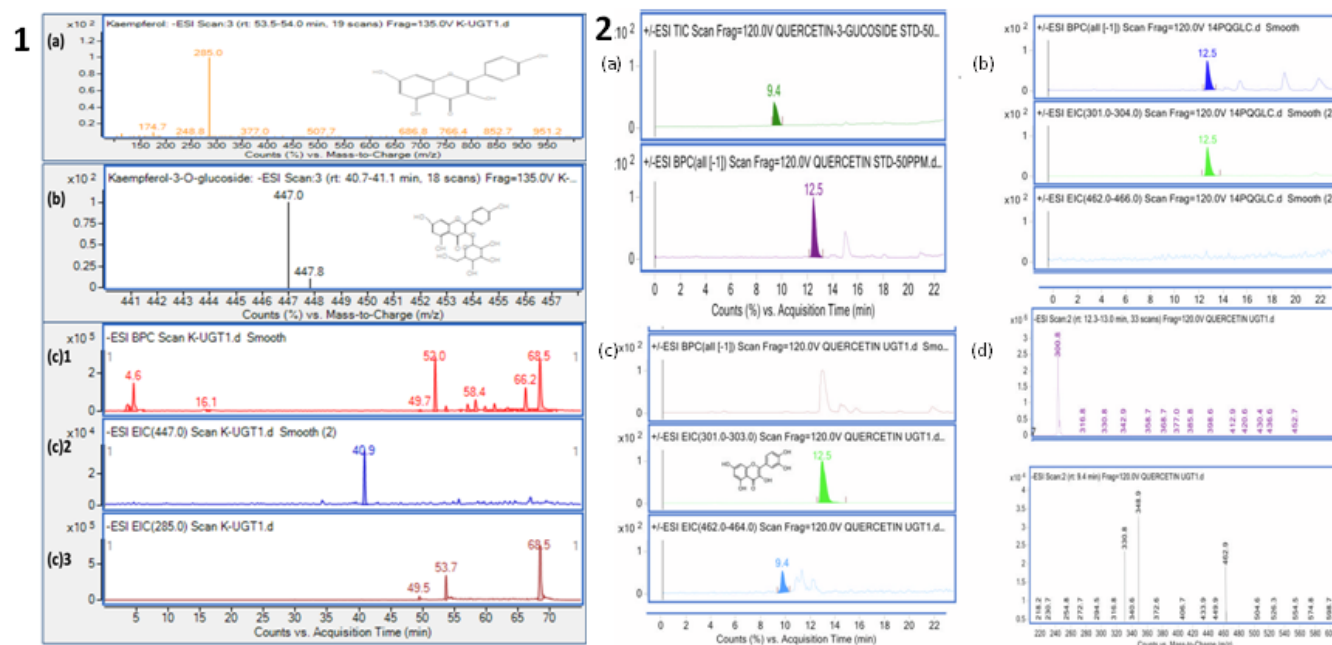
**Figure 5**

(A): SDS-PAGE electrophoresis of prokaryotic expression of pET28a-UGT72L11. Isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) of the gene in various timeline. Lane a-3hr; Lane b-6 hr; Lane c-12hr; Lane d-20hr; Lane e-24hr; Lane f- protein marker; Lane g- control without isopropyl- $\beta$ -D-thiogalactopyranoside; arrow indicates maximum induction of the fused protein at 20 hrs; (5B) SDS-PAGE electrophoresis of prokaryotic expression of eluted pET28a-UGT72L11; Lane5-8,flow through; lane10-lane14- eluted protein having approximately 52kDa of molecular weight



**Figure 6**

Optimization of various parameters (X axis) with respect to UDP concentration (nm) using purified recombinant GgUGT11 protein; (a) buffers/pH -C/3, C/4, C/5, C/6 in citrate buffer; P/6, P/7.5 in phosphate buffer; and T/7.5, T/8.5, T/9.5 and T/10.5 in tris-HCl buffer; (b) various time intervals (in minutes) 15,30,45,60 & 90; (c) different temperature (°C) at 4,10, 22, 30, 37, 50 & 70; and (d) kaempferol substrate (in  $\mu\text{M}$ ) at 3,5,7,9 concentrations.

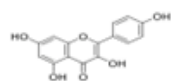


**Figure 7**

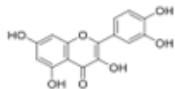
(1) HPLC/LCMS data of (a) standard kaempferol evidence peak with molecular weight 285 and RT value 53.4; (b) standard kaempferol 3-O-glucoside evidence peak with molecular weight 447 and retention time value 40.7; (c1) UGT72L11 enzymatic reaction peaks (c2): extraction of peak from (c1) at molecular weight 447 confirmed peak with retention time value 40.9 of kaempferol 3-O-glucoside; (c3) extraction of peak from (c1) at molecular weight 285 confirmed peak with retention time value 53.7 of kaempferol as in standard. (2) HPLC LC-MS analysis of UGT72L11 enzymatic reaction using quercetin as a substrate.

(a) total ion chromatograph (TIC) of the standards in positive/negative-ion mode showing peaks at RT of 12.5 and 9.4 mins for quercetin and quercetin 3-O-glucoside respectively; (b) control reactions (without protein) showing the base peak intensity (BPI) chromatogram in positive/negative-ion mode at RT of 12.5 min; (c) *in-vitro* enzymatic reaction with protein at the base peak intensity in positive/negative-ion mode showing peaks at retention times of 12.5 and 9.4 min; (d) peaks in the extracted ion chromatogram at a mass/charge ratio ( $m/z$ ) of 300.8 $[M-H]^-$  and 462.9  $[M-H]^-$  corresponding to quercetin and quercetin 3-O-glucoside, respectively

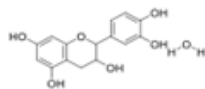
#### Flavone



kaempferol

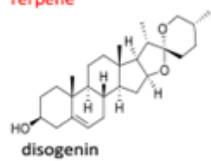


quercetin

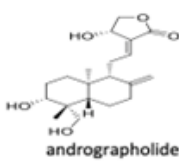


catechin hydrate

#### Terpene

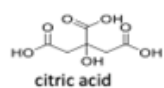


disogenin

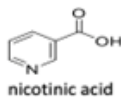


andrographolide

#### Carboxylic acid

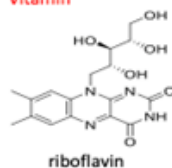


citric acid



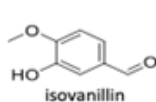
nicotinic acid

#### Vitamin



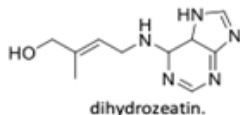
riboflavin

#### Phenolic

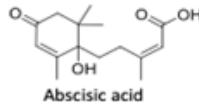


isovanillin

#### Hormone

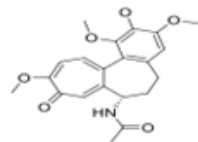


dihydrozeatin.

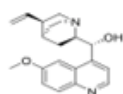


Absciscic acid

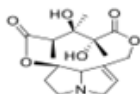
#### Alkaloid



Colchicine

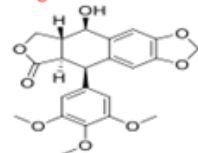


quinine hydrochloride

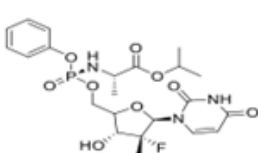


monocrotaline

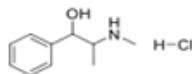
#### Drugs



Podophyllotoxin



sofosbuvir



ephedrine hydrochloride

Figure 8

Chemical structures of substrates used in the study (flavanols, terpenes, carboxylic acids, vitamins, alkaloid, phenolics and commercial drugs) for glycosylation.

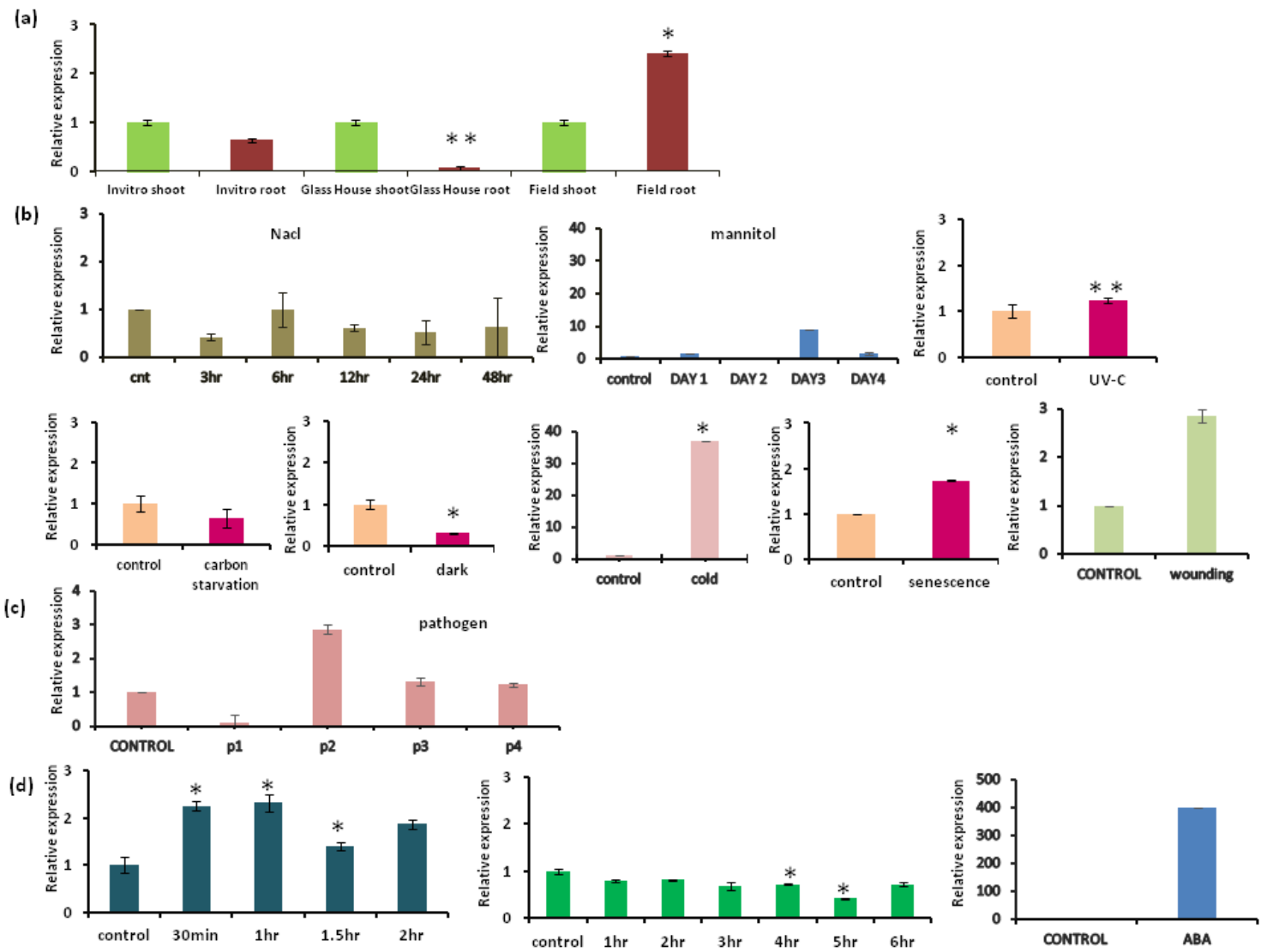
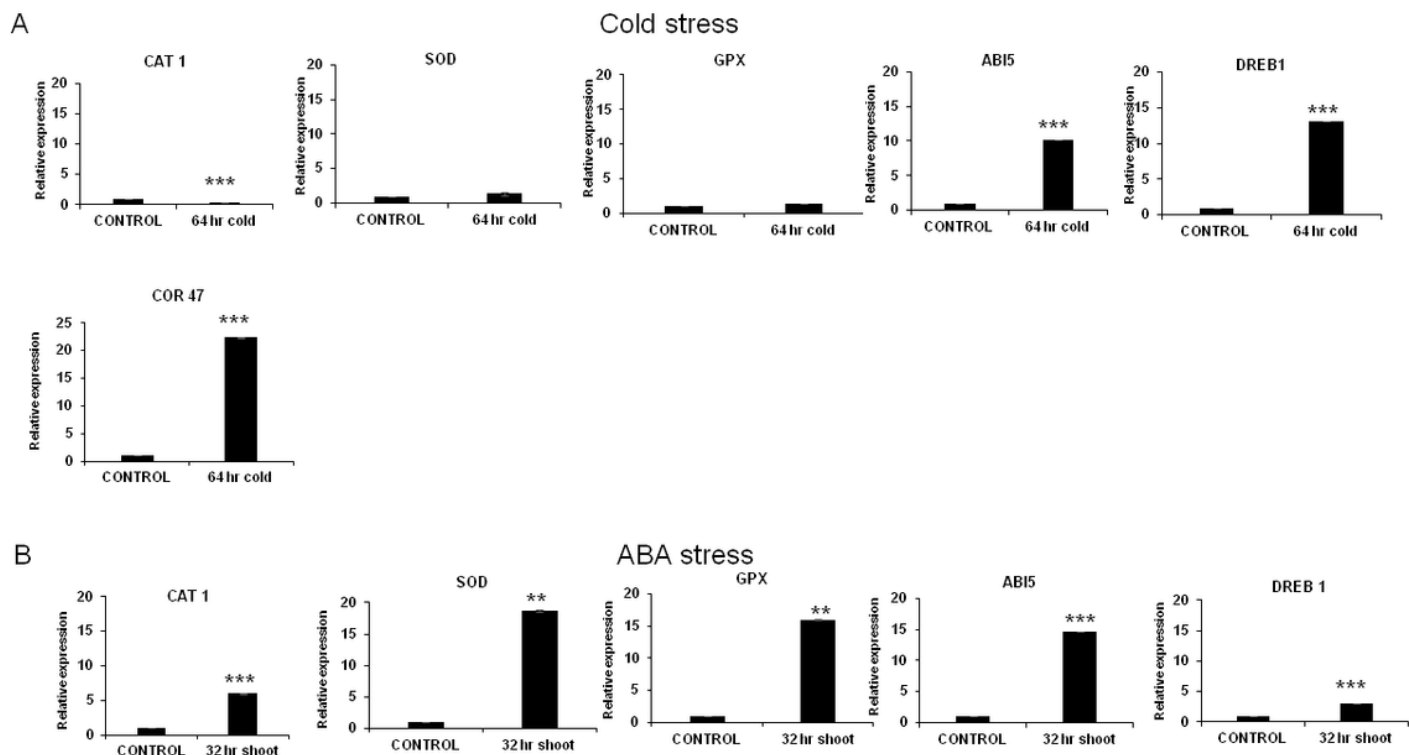


Figure 9

Quantitative expression level of *GgUGT72L11* by qRT-PCR. (a) tissue-specific expression under three conditions (tissue culture, hardening unit and field); (b) abiotic stress (NaCl , mannitol , UV, carbon starvation, dark, cold senescence and wounding); (c) pathogens(P1 - *Colletotrichum gloeosporioides* , P2- *Colletotrichum siamense.subsp*, P3-*Colletotrichum siamense sub sp.* and P4- *Calonectria cylindrospora*) (d) hormonal treatments (NAA, ABA & GA<sub>3</sub>). The y – axis indicates relative expression value and x-axis indicate the control and treated shoot tissue under different conditions/stress,. Three biological replicates were used to calculate error bars using standard deviation. Asterisks indicate the statistical significance of observed values when compared with control [(\*) p value  $\leq 0.05$ , (\*\*) p value  $\leq 0.01$ , (\*\*\*) p value  $\leq 0.001$ ] calculated by One -way ANNOVA analysis in Graph pad prism 8.0.



**Figure 10**

Quantitative expression levels of Reactive Oxygen Species scavenging genes catalase (*CAT1*), super oxide dismutase (*SOD*), Glutathione Peroxidase (*GPX*) and cold –responsive genes (*COR47*, *ABI5*, *DREB1*) by qRT-PCR. (A) 64hr cold treated shoot; (B) 32 hr ABA treated shoot. The y – axis indicates relative expression value and x-axis indicate the control and treated shoot tissue under stress. Three biological replicates were used to calculate error bars using standard deviation. Asterisks indicate the statistical significance of observed values [(\*) p value  $\leq 0.05$ , (\*\*) p value  $\leq 0.01$ , (\*\*\*) p value  $\leq 0.001$ ] calculated by One way ANNOVA analysis in Graph pad prism 8.0

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementaryfig1.pdf](#)
- [supplementaryfig2.pdf](#)
- [supplementaryfig3.tif](#)
- [Supplementaryfig.4.jpg](#)
- [SUPTABLE.docx](#)
- [SupportingData1.pdf](#)