In vivo Gloriosa superba and *Colchicum autumnale* multi-tissues transcriptome analysis for colchicine pathway and rhizome development candidate genes identification

A Thesis presented to

The Faculty of the Department of Engineering Technology

University of Houston

In Partial Fulfilment of the Requirements for the Degree

Master's in Engineering Technology

by

John Samuel Bass

December 2018

In vivo Gloriosa superba and *Colchicum autumnale* multi-tissues transcriptome analysis for colchicine pathway and rhizome development candidate genes identification

Mr. John Samuel Bass
Masters Student
Engineering Technology – Biotechnology track

Sivakumar Ganapathy Ph. D.
Assistant Professor and Committee Chair
Engineering Technology

Rupa Iyer, Ph. D.
Associate Professor
Engineering Technology

Patrick Cirino, Ph. D.
Associate Professor
Cullen College of Engineering

Wajiha Shireen, Ph. D.
Professor and Chair
Engineering Technology

George Zouridakis, Ph. D.
Associate Dean
College of Technology

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor Dr. Sivakumar Ganapathy for providing me with the necessary guidance and opportunity to discover a new founded love for biotechnology, as well as the many valuable discussions about this thesis and the possible future that I now can achieve. I would also like to thanks Brian Iken who gave me direction during critical stages. Also, I would like to thank my fellow classmate and friend, Santhanakrishnan Raghunandan. Next, thanks to Dr. Toni Kutchan who established the transcriptomes and Dr. David Gang who conceived of, designed, and coordinated the rhizome developmental genes verification. I would also like to thank my committee members for the opportunity and support, Drs. Sivakumar Ganapathy, Rupa Iyer, and Patrick Cirino. Lastly, I would like to deeply thank my family and my amazing girlfriend for the everlasting support and patience throughout this process.

ABSTRACT

Background: The continued emergence of side-effects caused by synthetic drugs underscores the need for plant-based drugs in human medicine. Medicinal rhizomatous crops are "the goldmine for modern drugs", and include such species as *Gloriosa superba* L., and *Colchicum autumnale* L., the producer of colchicine, a plant-based medicine. The natural isomer of bioactive colchicine is used to effectively treat major diseases such as cancer, cardiovascular disease, and gout. The medicinal properties of colchicine are well characterized, however, almost nothing is known about its biosynthetic mechanism and colchicine pathway has not been elucidated that are significant barriers in biomanufacturing of biomedicine. The comparative transcriptomes study of *G. superba* and *C. autumnale* can serve as sequence resource and synthetic biology toolbox components for identifying biomedicine pathway and rhizome development genes, which could aid colchicine pathway metabolic engineering or synthetic biotechnology to improve colchicine biomanufacturing.

Result: Predominantly colchicine synthesizing two monocots such as *G. superba* and *C. autumnale* transcriptomes were used to identify putative protein involved in the colchicine biosynthetic pathway and rhizome development along with transcription factors. Mining of the transcriptomes using Blast2GO, 20 and 29 candidate genes [3 and 1 candidate *N-methyltransferase (NMT); 10 and 16 candidate 3-O-methyltransferase (3-OMT); cytochrome P450s*, a class that could catalyze several steps in the pathway namely, 2 and 5 candidate *CYP96T1,* 1 and 4 candidate *CYP82E10*; 4 and 3 candidate *N-acetyltransferase (NAT)*] were identified in colchicine pathway for *G. superba* and *C. autumnale*, respectively. Similarly, 19 and 15 candidate rhizome developmental genes [2 and 1 candidate *GIGANTEA (GI),* 5 and 4 candidate *CONSTANS (CO),* 2 and 1 candidate *Phytochrome B (PHYB),* 2 and 5 candidate *Sucrose Synthase (SuSy),* 5 and 2 candidate *Flowering Locus T (FT), and* 3 and 2 candidate *REVOLUTA (REV)*] were identified in *G. superba* and *C. autumnale*, respectively. While 16 and 12 transcription factors in rhizome development and regulating secondary metabolic pathways in rhizomes [3 and 1 candidate *MADS-box,* 6 and 2 candidate *AP2-EREBP*, 2 and 2 candidate *bHLH*, 1 and 2 candidate *MYB*, 2 and 2 candidate *NAC*, and 2 and 3 *WRKY*] were screened in *G. superba* and *C. autumnale,* respectively. These genes could represent potential leads for metabolic engineering of *G. superba* or synthetic biotechnology of colchicine metabolism for enhanced colchicine and biorhizome biomass in biomanufacturing.

Conclusion: The study of *G. superba* and *C. autumnale* genes predicated to encode colchicine pathway enzymes are highly significant for fundamental information on plant-based biomedicine biosynthesis, which could facilitate engineered production in biorhizomes, a potentially important area of synthetic biotechnology. Additionally, increasing our understanding of rhizome genomics could improve colchicine production in *G. superba*, and generate important knowledge that could be applied to many other medicinal plant species, and could allow engineered production of additional biomedicines in biorhizomes, a potentially important area of expansion for synthetic biotechnology to solve overarching biomanufacturing challenges.

INTRODUCTION

The growing use for ultrapure plant-based medicines to treat human disease often cannot be met due to a lack of feasible upstream biomanufacturing processes. For example, therapeutic colchicine alkaloid, a drug used to treat cancer, cardiovascular disease, and gout [1-25], is uniquely biosynthesized by the Colchicaceae family and extracted commercially from *Gloriosa superba* and *Colchicum autumnale*, an important rhizome crops and the prime pharmaceutical source of colchicine. Further, colchicine is an FDA approved drug [26]. The colchicine analog CH-35 was more effective at inhibiting the βIII isotype of breast cancer than taxol [27-30]. The results of clinical trials suggest that colchicine could prevent the recurrence of atrial fibrillation after cardiac surgery and renal diseases [31-34]. Moreover, colchicine is undergoing clinical trials to treat non-diabetic metabolic syndrome and diabetic nephropathy [35-37]. The best chemical synthesis method available generates only a 9.2% yield of >99% (-)-colchicine [38-39]. Therefore, primarily the commercially viable source remains extraction from the Ayurvedic medicinal plant *G. superba* [22, 40]. However, the annual production of pharmaceutical colchicine is low, and the source is limited [41-45]. Despite the availability of *G. superba* and *C. autumnale* sequence data from the Medicinal Plant Transcriptome (http://www.medplantrnaseq.org) and chloroplast genome, little is known about the colchicine biosynthetic pathway, and almost nothing is known about how colchicine production is regulated in the plant [46]. The efficiency of colchicine biosynthesis depends on the expression of gene circuits on other components within the biosynthetic network, and how those gene circuits are regulated. The situation with colchicine production is not unusual—as more plant-based medicinally important compounds are discovered, especially where high purity and large amounts are required, this situation will continue to be faced. To overcome the plant-based therapeutic colchicine production, an advanced non-dormant *in vitro* biorhizome

technology from the *G. superba* has been established [47]. Biorhizomes are non-transgenic, serve as asexual reproductive organs, an advanced biotechnological platform compared to root, and cell cultures due to their continuous and rapid colchicine production [48]. Nevertheless, the biochemical pathways and regulatory networks in the biorhizomes that control colchicine biosynthesis are yet to be characterized, leaving a significant barrier to improving colchicine biomanufacturing. Therefore, the first steps in building a synthetic biology toolbox for colchicine production include analysis of genes from the different Colchicaceae species in order to identify bottleneck or other regulatory steps that can be then be adjusted to enhance colchicine production in the biorhizomes.

The current understanding of colchicine biosynthesis in planta is based on radiolabeling studies [49-51] and the transformation of *O*-methylandrocymbine to demecolcine with microsomes prepared from immature *C. autumnale* seeds [52]. The phenylalanine amino acid precursor in the cinnamic acid pathway and trihydroxylated phenethylisoquinoline in the colchicine pathway have been studied [53-56]. However, research has not been performed at the molecular level to reveal a colchicine biosynthetic pathway in *G. superba* and *C. autumnale*. To fully understand how colchicine is biosynthesized in *Gloriosa* biorhizomes, the key genes and enzymes that control the colchicine pathway from the alkaloidal precursor trihydroxylated phenethylisoquinoline to colchicine must be identified. We constructed a full-length cDNA library using mRNA isolated from *G. superba* leaves that consists of 2,790 processed sequences, of which 1,379 were assembled into 292 contigs and 1,411 singletons [47]. The cDNA library contains gene families expected to be involved in the colchicine pathway, including *NMT*, *3-OMT, cytochrome P450s: CYP96T1, CYP82E10, and NAT*. Furthermore, we manually curated these gene families and identified

specific genes whose corresponding enzymes are excellent candidates for involvement in the colchicine pathway, from the alkaloid formation steps to later steps in the pathway (Figure 1), including: 1) an *NMT* enzyme that catalyzes the conversion of the trihydroxylated phenethylisoquinoline intermediate to (*S*)-autumnaline; 2) a *P450: CYP96T1* that catalyzes (*S*)-autumnaline to isoandrocymbine; 3) an *OMT* enzyme that catalyzes the transformation of isoandrocymbine to *3-O*-methylandrocymbine; 4) a P450 that catalyzes *O*-methylandrocymbine to demecolcine; 5) an additional *P450: CYP82E10* that catalyzes the conversion of demecolcine to deacetylcolchicine; and 6) an *NAT* enzyme that catalyzes the transfer of an acetyl group from acetyl-CoA to a deacetylcolchicine nitrogen group to yield colchicine [47].

What genes are expressed in rhizomes? Several recent investigations have identified rhizome-specific genes in different rhizomatous plants by directly comparing leaves, other tissues, and rhizomes, leading to the detection of genetic mechanisms responsible for controlling rhizome development, growth, and metabolism. Many genes exhibit significantly altered expression during rhizome development but genes associated with auxin hormone signaling appear to trigger rhizome induction [57-59]. The rhizome developmental gene *REV* is highly expressed in bamboo rhizome buds and plays an important role in meristem initiation [60]. In potato plants, calmodulin-binding protein plays a regulatory role in signal transduction for tuber formation [61]. As well, the *FT*, *CO*, and *GI* genes are involved in the transduction of photoperiodic signals, which may promote rhizome budding in potatoes [62-63]. There are 14 other important rhizome formation-related genes, including a *MADS-box* that could be involved in rhizome enlargement [64]. Genes encoding *PHYB, CO*, *GI*, *FT* and *SuSy* were identified in Lotus rhizomes, but their expression and regulation differed in the shoot and rhizome [65]. Transcription factor families such as *AP2-EREBP, bHLH,*

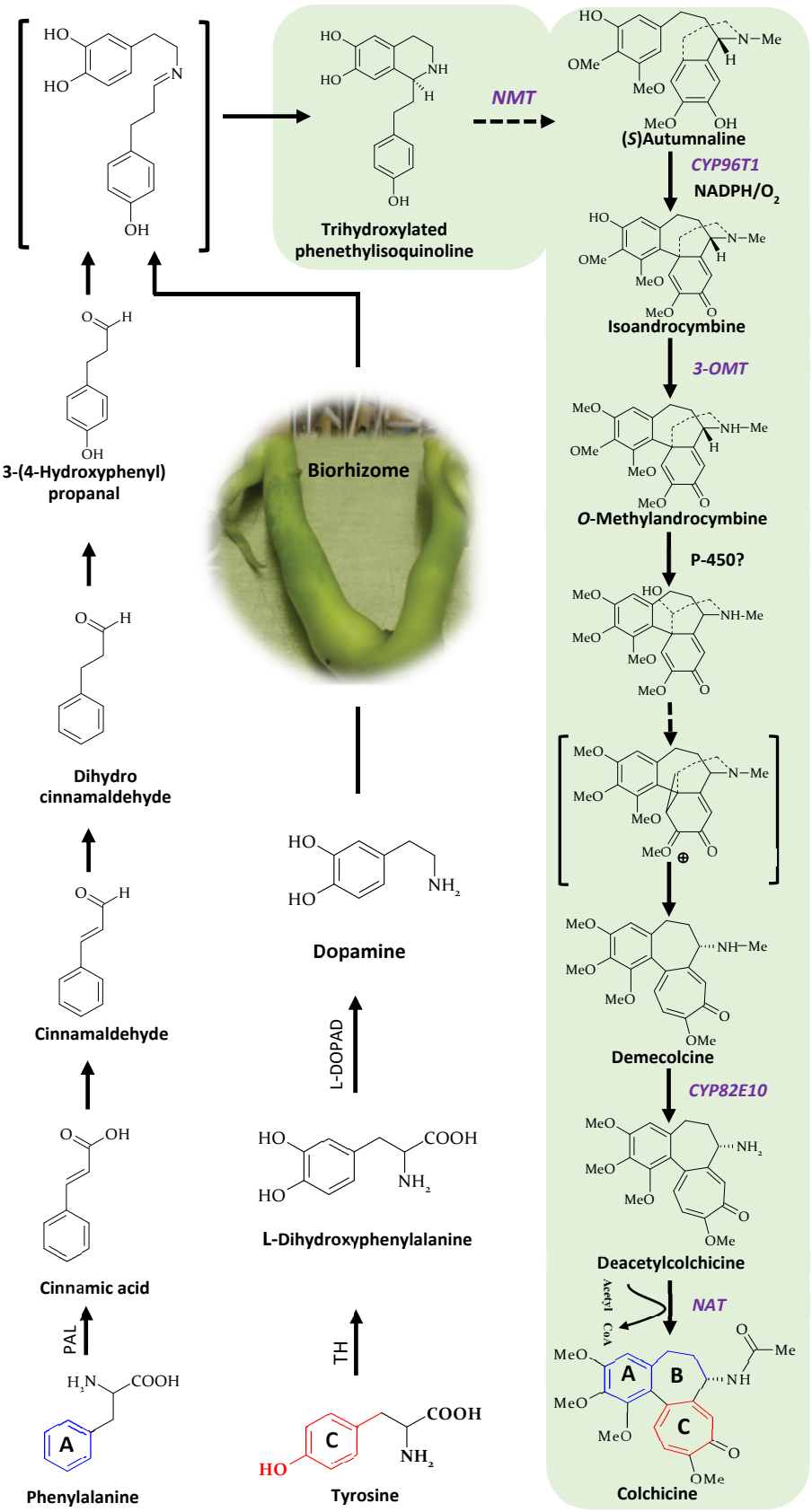

Figure 1. Proposed colchicine pathway in *Gloriosa superba*. TH: Tyrosine hydroxylase; L-DOPAD: L-Dihydroxyphenylalanine decarboxylase; PAL: Phenylalanine ammonia-lyase; NMT: *N*-methyltransferase; OMT: *O*-methyltransferase; NAT: *N*-acetyltransferase [47]

MYB, NAC, and *WRKY* reportedly were important to regulating secondary metabolic pathways in rhizomes [66]. Bioactive small molecules such as curcuminoid and candidate genes for gingerol synthesis are also highly expressed in rhizomes [67-71]. Notably, biosynthetic genes involved in benzylisoquinoline alkaloid formation were highly upregulated during bulb development in *Corydalis yanhusuo* [72]. In wild rice, microRNAs were differentially expressed in aerial shoots and rhizomes [73]. However, the exact roles that the corresponding genes might play in biorhizomes are not known.

Although significant progress has been made in understanding rhizome-specific mechanisms in plants, the mechanisms underlying the regulation of *Gloriosa* biorhizome growth, including the dormancy mechanism in field-grown *G. superba* and *C. autumnale* are not yet known. Similar to many other rhizomatous species, natural *G. superba* and *C. autumnale* rhizomes undergo a dormancy period in their normal growth cycle, but the biorhizomes do not go dormant. Important questions remain regarding the genes involved in biorhizome development and colchicine biosynthesis in *G. superba*. Extensive studies examined the phenotypic variation between plant species, but why dormancy-free *Gloriosa* species biorhizomes produce different levels of colchicine in the control bioreactor environment remains unclear. In addition, the dormancy-associated genes or transcription factors that may be over-expressed in biorhizomes and whether the core regulatory machinery that controls colchicine biosynthesis vary between species are not known. We hypothesize that *G. superba* and *C. autumnale* genes and gene networks are comparable, but that subtle differences in their regulation lead to changes in compound accumulation. Comparison of the transcriptomes of these species will fill the identified knowledge gaps. Ideally, several Colchicaceae species would be analyzed to identify their rhizome-specific

genes and colchicine metabolism. However, many species do not produce significant colchicine levels. For that reason, we selected *in vivo* multi-tissue *G. superba* and *C. autumnale*, a sample large enough to cover comparative data within the genus. This will enable the breadth and depth of genes expressed in the rhizome and generate a much better starting set of colchicine pathway candidate genes to test. The objective was to identify bottleneck and other regulatory mechanisms in colchicine production (*NMT, CYP96T1, 3-OMT, CYP82E10, and NAT)*, rhizome developmental *(GI, CO, PHYB, SuSy, FT,* and *REV)* along with its transcriptional candidate genes (*MADS-box AP2-EREBP, bHLH, MYB, NAC*, and *WRKY*) in *G. superba* and *C. autumnale* transcriptomes that can be then be targeted using metabolic engineering or synthetic biology approaches to enhance colchicine production in biomanufacturing (Figure 2).

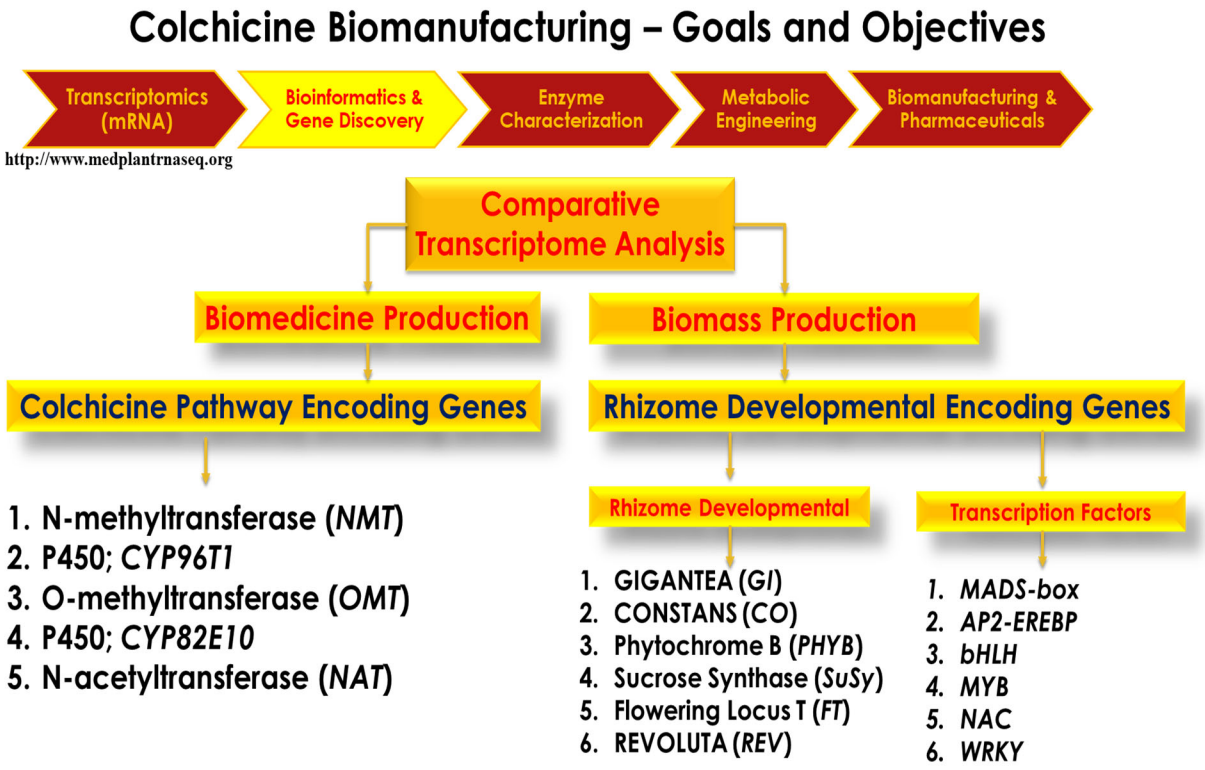

Figure 2. Research goals and objectives.

RESULTS AND DISCUSSION

Benchmarking universal single-copy orthologue (BUSCO) analysis: Three *in vivo* tissues (leaf, fruit, and rhizome) were previously used to generate a combined RNA-seq dataset of *G. superba* and *C. autumnale* (http://www.medplantrnaseq.org). The results of the analysis indicated that the N50 for the assembly was fairly long at 2,134, given that contigs of ≥ 100 (instead of ≥ 200) were included in the assembly and that the data were generated from 50 bp single-end reads. The average contig length was somewhat short, however, as the statistic was skewed due to the inclusion of several contigs that were shorter than 200. BUSCO analysis, which in this case examined the core eukaryotic genes in plants, indicated that the dataset was ~89% complete (64% of the genes detected were found as a single sequence in the assembly, and 25% had duplicated sequences). In comparison, a total transcriptome analysis that we conducted focused on only a single sample type (containing combined RNA samples from the 1st and 2nd Asian Citrus Psyllid instars, including six biological replicates and utilizing 150 bp reads), led to BUSCO analysis results suggesting that the dataset was 96% complete (13% single sequence and 83% duplicate sequences), with only 1.4% fragments of core genes and 2.6% missing core genes.

Annotation and comparative transcriptomes analysis: The *C. autumnale* and *G. superba* transcriptomes consist of 60,927 and 32,312 assembled multiple-tissue transcripts with 21,948 and 15,089 unigenes identified as having functions belonging to known plant-specific gene ontology (GO), respectively. Among 23,247 and 45,292 sequences were assigned and annotated as biological processes, 27,199 and 52,366 sequences as cellular components, and 22,760 and 44,942 sequences as molecular functions in *G. superba* and *C. autumnale*, respectively (Figure 3). The *G. superba* leaf tissue cDNA library was also annotated using the Blast2GO suite and contained a

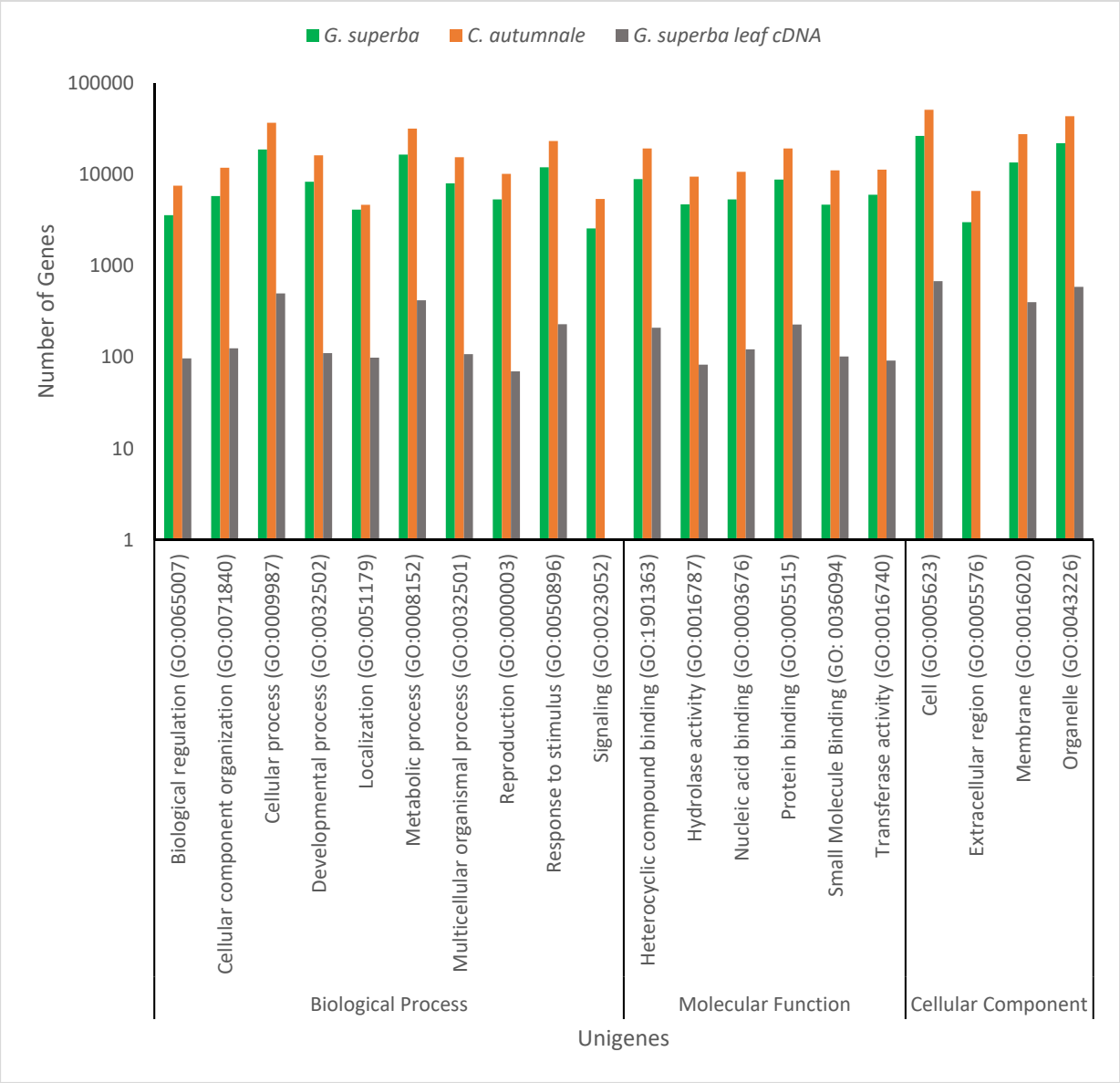

Figure 3. Gene ontology (GO) annotation of *Gloriosa superba* and *Colchicum autumnale* transcriptomes.

total of 1,703 unigenes with 588 sequences were assigned as biological processes, 700 sequences as cellular components, and 568 sequences as molecular functions (Figure 3). Figure 4 shows the percentages of GO terms assigned to the transcriptomes of *G. superba* and *C. autumnale*, and the cDNA library. Additional GO terms were used to identify Kyoto Encyclopedia of Genes and

Genomes (KEGG) pathway members, putative colchicine pathway, rhizome developmental genes along with it transcription factors. KEGG annotated 15184 of 57580 (26.4%) and 33505 of 100528 (33.3%) sequences in *G. superba* and *C. autumnale* transcriptomes, respectively (Figure 5). *G. superba* cDNA library was also annotated by BlastKOALA, which annotates smaller 66 of 848 (7.8%) sequences sets of proteins based on their role in specific pathways (Figure 5).

Figure 4. Percentage of gene ontology annotations for molecular function, biological process, and cellular components in *Gloriosa superba* and *Colchicum autumnale* transcriptome and *G. superba* leaf cDNA library.

The mining of *G. superba* and *C. autumnale* transcriptomes revealed a total of 1299 possible colchicine pathway candidate genes, which includes 647 sequences from *G. superba* and 652 from *C. autumnale*. From 647 candidate sequences in *G. superba*, 186 were *NMT*, 105 were *OMT*, 19 were *NAT*, and 337 were P450s; while from 652 candidate sequences in *C. autumnale*, 16 were

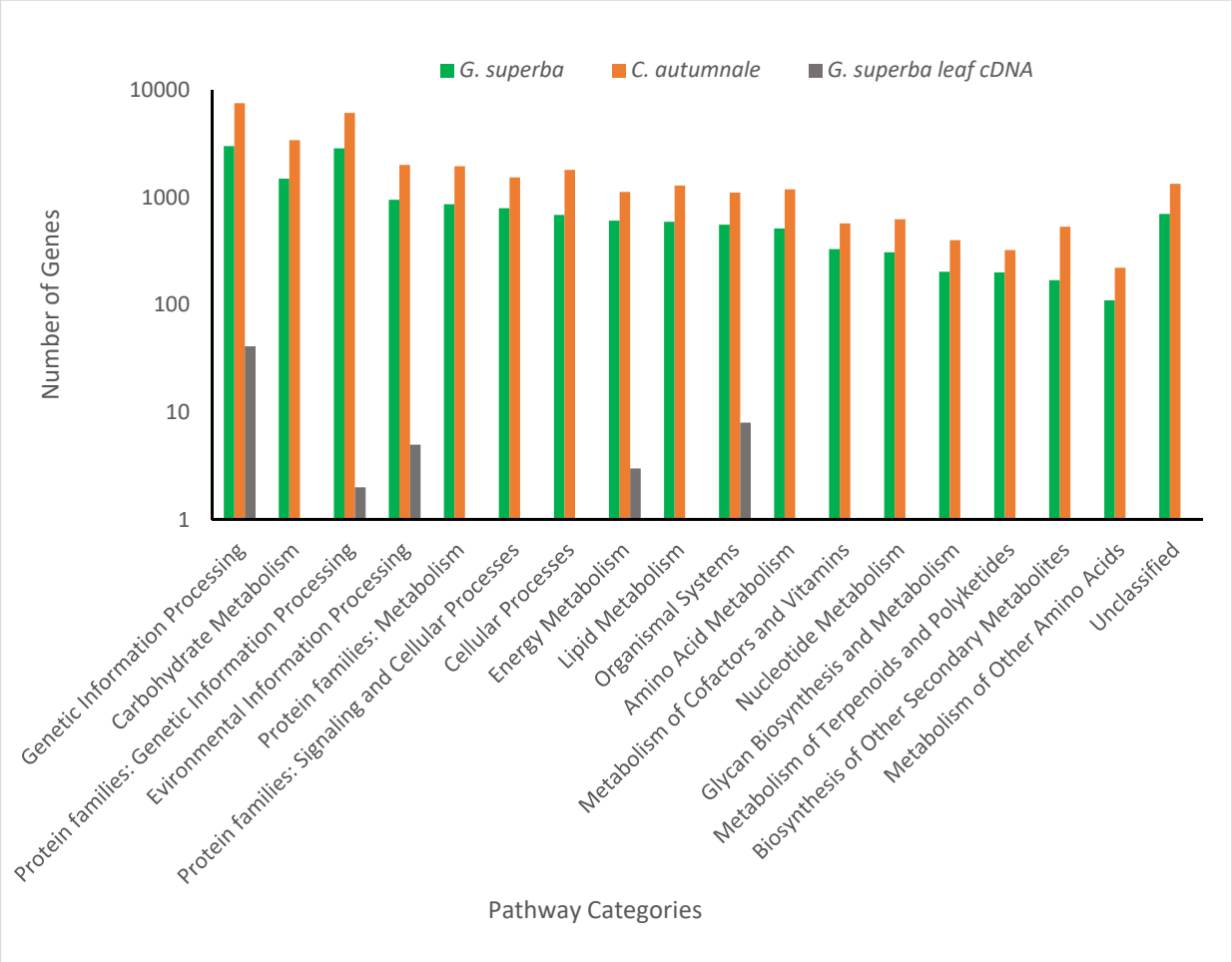

Figure 5. Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation of *Gloriosa superba* and *Colchicum autumnale* transcriptomes and *G. superba* leaf cDNA library predicted proteins.

NMT, 106 were *OMT*, 20 were *NAT*, and 510 were P450s (Figure 6). Further, these transcripts were narrowed down to colchicine biosynthetic pathway enzymes, which could be candidates for the genes that catalyze the reactions in the colchicine biosynthetic pathway. Moreover, a total of

339 rhizome developmental genes such as *GI, CO, PHYB, SuSy, FT, and REV* where also identified in *G. superba* and *C. autumnale* transcriptomes (Figure 7). Among 113 sequences from *G. superba,* 226 from *C. autumnale*. To identify these genes specific functions, a known rhizome developmental reference genes 119 sequences from *Nelumbo nucifera,* and 142 sequences from *Solanum tuberosum* were used to obtain the rhizome developmental candidate genes in *G. superba* and *C. autumnale*. In addition, a total of 1146 transcription factor sequences associated with rhizome development namely *MADS-box, AP2-EREBP, bHLH, MYB, NAC*, and *WRKY* were analyzed (Figure 8). This includes 481 sequences from *G. superba*, 665 from *C. autumnale*, 811 from *N. nucifera,* 1019 from *S. tuberosum* were selected for further refinement to obtain the transcription factor candidate genes.

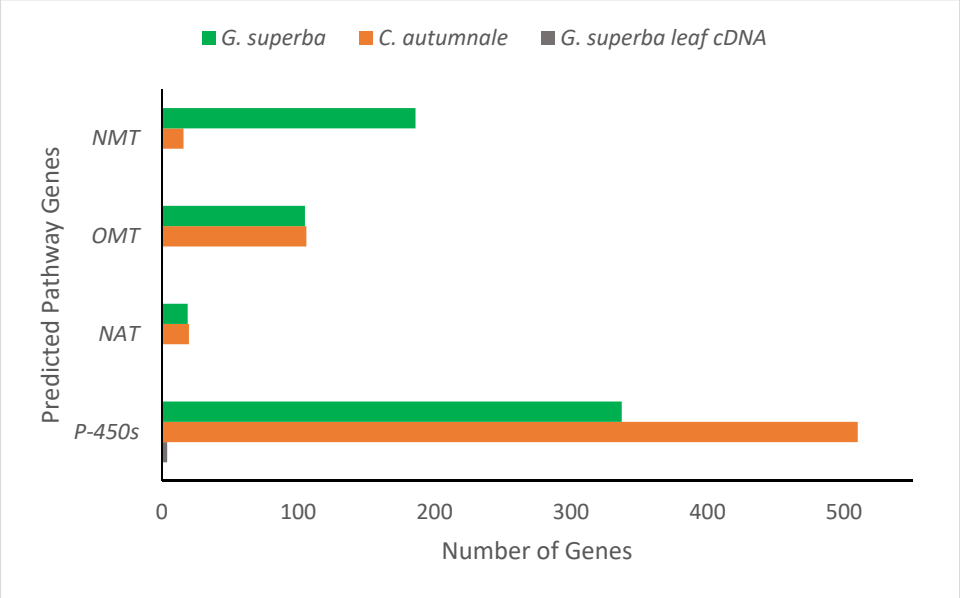

Figure 6. Total number of transcripts in colchicine pathway in *Gloriosa superba* and *Colchicum autumnale* transcriptomes and *G. superba* leaf cDNA library.

Why are colchicine pathway candidate NMT, P450s: CYP96T1 and CYP82E10*, 3-OMT, and NAT genes from the transcriptomes the best targets?* First, a cDNA contig consisting of three sequences

showed significant sequence similarity to *Ricinus communis NAT* and partial similarity to *Oryza sativa*. This contig, which contains a full-length transcript (cDNA *Gloriosa* 148), is a potential candidate for the enzyme that catalyzes the acetyl transfer from acetyl-CoA deacetylcolchicine to colchicine, the final step in the proposed colchicine pathway [52]. The putative *NAT* gene also showed complete homology to a full-length transcript in the *G. superba* transcriptome (*Gloriosa*-20120814_26859). Similarly, a serotonin *NAT 1* and *2* (*SNAT1* and *2)* reactions were also identified in melatonin biosynthesis in *O. sativa* [74-75]. The *SNAT1* transcript Q5KQI6.1 is also considered for possible alternative reference sequence, which homology to *G. superba* and *C. autumnale* transcriptomes (*Gloriosa*-20120814|70209_1 and *Colchicum*_20101112|6813) with an identity of 69%, and 77%, respectively (Table 1 and Figure 9).

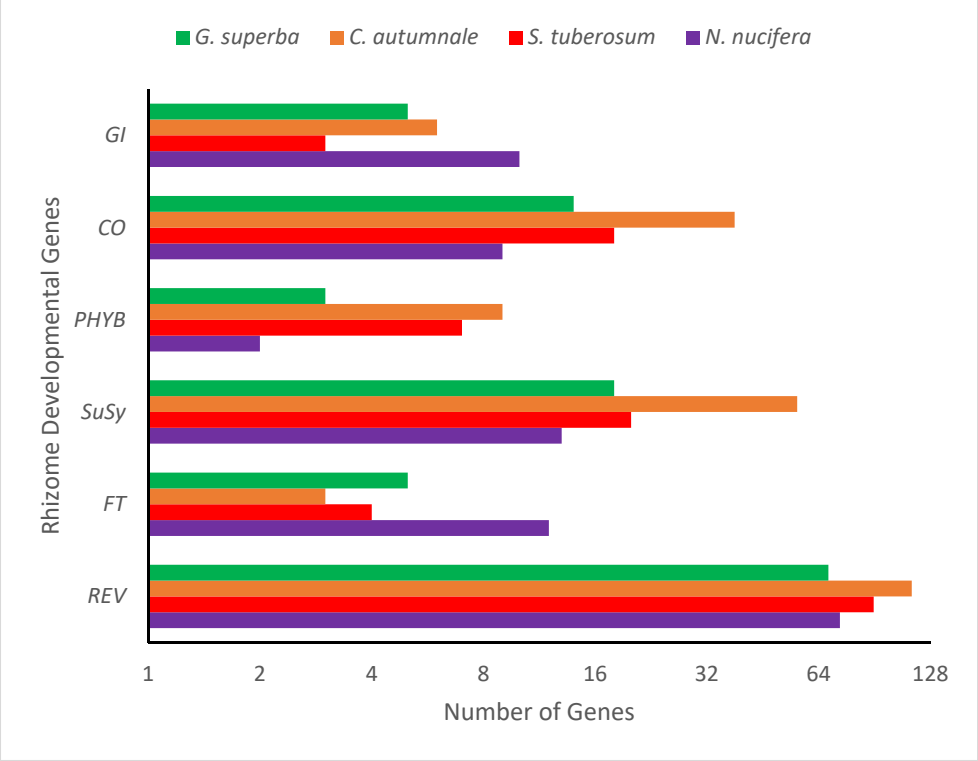

Figure 7. Total number of transcripts of rhizome developmental genes in *Gloriosa superba* and *Colchicum autumnale* transcriptomes and *G. superba* leaf cDNA library. *Solanum tuberosum* and *Nelumbo nucifera* were reference transcriptomes.

Second, the cDNA clone (*Gloriosa* 14D06) identified as a putative *NMT* is a partial clone consisting mainly of the 3' end of the cDNA. The putative *NMT* clone had high sequence homology to *Coptis japonica NMT*, with a similar *N*-methylation reaction involved in (*S*)-*N*-methylcoclaurine formation in the benzylisoquinoline pathway and has an identical full-length

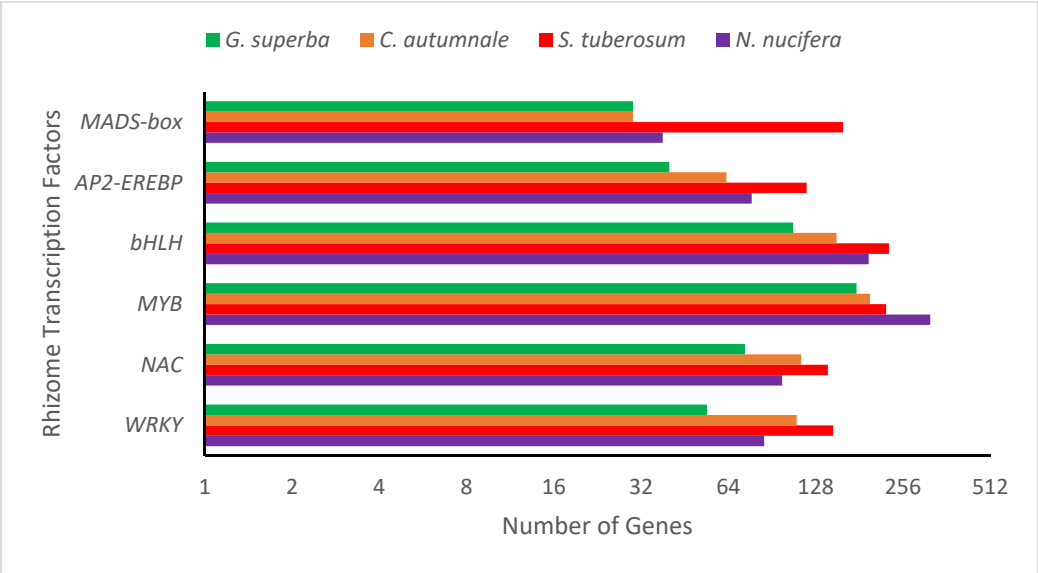

Figure 8. Total number of transcripts of transcription factors in *Gloriosa superba* and *Colchicum autumnale* transcriptomes and *G. superba* leaf cDNA library. *Solanum tuberosum* and *Nelumbo nucifera* were reference transcriptomes.

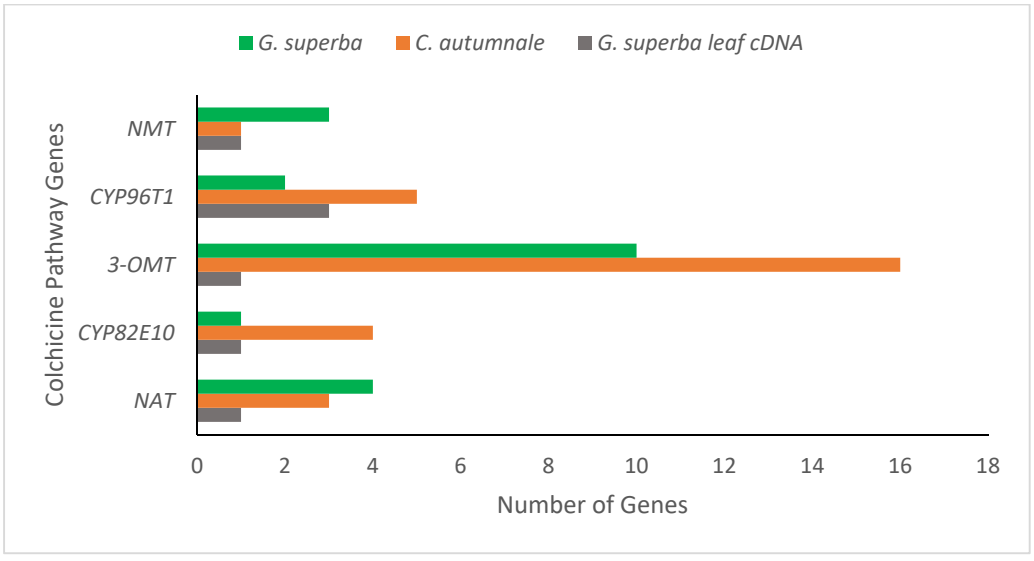

Figure 9. Possible candidate colchicine pathway genes in *Gloriosa superba* and *Colchicum autumnale* transcriptomes and *G. superba* leaf cDNA library.

Table 1. Predicted candidate colchicine pathway genes in *Gloriosa superba* and *Colchicum autumnale* transcriptomes.

Colchicine Pathway Genes	Candidate Gene ID	Reference Gene ID	% Identity	e-Value
Step 1: NMT	Gloriosa-20120814|64082_1	BAB71802.1 [Coptis japonica]	57.792	1.88E-63
	Gloriosa-20120814|31585_1	BAB71802.1 [Coptis japonica]	52.866	2.28E-123
	Colchicum_20101112|85_	BAB71802.1 [Coptis japonica]	51.862	4.03E-127
	Colchicum_20101112|85_	C3SBW0.1 [Thalictrum flavum subsp. glaucum]	48.703	3.10E-116
	Gloriosa-20120814|64082_1	C3SBW0.1 [Thalictrum flavum subsp. glaucum]	51.299	9.25E-53
	Gloriosa-20120814|12877_1	C3SBW0.1 [Thalictrum flavum subsp. glaucum]	51.29	7.38E-111
Step 2: P450	Colchicum_20101112|29610_	CYP719B1 [Papaver somniferum]	50	8.81E-17
	Colchicum_20101112|67076_	CYP96T1 [Narcissus pseudonarcissus]	51.282	9.95E-24
	Gloriosa-20120814|30999_1	CYP80G2 [Coptis japonica]	52.778	1.79E-06
	Gloriosa-20120814|85639_1	CYP80G2 [Coptis japonica]	52.206	8.23E-35
	Colchicum_20101112|76071_	CYP80G2 [Coptis japonica]	51.724	1.24E-27
	Colchicum_20101112|88526_	CYP80G2 [Coptis japonica]	50	2.01E-07
	Colchicum_20101112|99151_	CYP80G2 [Coptis japonica]	50	7.77E-31
Step 3: OMT	Gloriosa-20120814|31595_1	AFB74613.1 [Papaver somniferum]	55.814	6.62E-10
	Gloriosa-20120814|18437_1	AFB74613.1 [Papaver somniferum]	54.795	2.60E-21
	Gloriosa-20120814|13320_1	AFB74613.1 [Papaver somniferum]	52	5.10E-20
	Gloriosa-20120814|46280_1	AFB74613.1 [Papaver somniferum]	51.429	3.68E-07
	Gloriosa-20120814|6585_1	AFB74613.1 [Papaver somniferum]	51.282	4.34E-08
	Gloriosa-20120814|2006_1	AFB74613.1 [Papaver somniferum]	51.163	7.24E-08
	Gloriosa-20120814|16875_1	AFB74613.1 [Papaver somniferum]	50.909	1.37E-33
	Gloriosa-20120814|31794_1	AFB74613.1 [Papaver somniferum]	50.847	1.32E-15
	Gloriosa-20120814|36043_1	AFB74613.1 [Papaver somniferum]	50.588	4.80E-22
	Gloriosa-20120814|25318_1	Q9LEL5.1 [Coptis japonica]	62.162	5.4E-11
	Gloriosa-20120814|2864_1	Q9LEL5.1 [Coptis japonica]	50.725	4.88E-16
	Gloriosa-20120814|25750_1	Q9LEL5.1 [Coptis japonica]	50.685	1.15E-22
	Colchicum_20101112|51346_	Q9LEL5.1 [Coptis japonica]	53.333	1.16E-28
	Colchicum_20101112|72754_	Q9LEL5.1 [Coptis japonica]	52.809	1.86E-20
	Colchicum_20101112|59167_	Q9LEL5.1 [Coptis japonica]	50.485	9.05E-29
	Gloriosa-20120814|82787_1	A0A077EWA5.1 [Narcissus pseudonarcissus]	72.857	1.42E-26
	Gloriosa-20120814|42115_1	A0A077EWA5.1 [Narcissus pseudonarcissus]	68.636	2.31E-109
	Gloriosa-20120814|238_1	A0A077EWA5.1 [Narcissus pseudonarcissus]	64.017	3.14E-104
	Gloriosa-20120814|10195_1	A0A077EWA5.1 [Narcissus pseudonarcissus]	59.031	1.08E-92
	Colchicum_20101112|68404_	A0A077EWA5.1 [Narcissus pseudonarcissus]	71.25	2.12E-35
	Colchicum_20101112|14671_	A0A077EWA5.1 [Narcissus pseudonarcissus]	67.234	3.44E-114
	Colchicum_20101112|34468_	A0A077EWA5.1 [Narcissus pseudonarcissus]	65.236	3.19E-108
	Colchicum_20101112|9022_	A0A077EWA5.1 [Narcissus pseudonarcissus]	63.983	4.49E-108
	Colchicum_20101112|75393_	A0A077EWA5.1 [Narcissus pseudonarcissus]	63.194	2.07E-62
	Colchicum_20101112|29556_	A0A077EWA5.1 [Narcissus pseudonarcissus]	59.633	9.11E-89
	Colchicum_20101112|23611_	A0A077EWA5.1 [Narcissus pseudonarcissus]	58.15	1.12E-90
Step 4: P450	Gloriosa-20120814|5091_1	CYP82E5v2 [Nicotiana tabacum]	52.326	1.27E-24
	Colchicum_20101112|74349_	CYP82E5v2 [Nicotiana tabacum]	55.172	4.87E-29
	Colchicum_20101112|75506_	CYP82E5v2 [Nicotiana tabacum]	50.962	6.75E-34
	Gloriosa-20120814|5091_1	CYP82E10 [Nicotiana tabacum]	51.163	6.91E-25
	Colchicum_20101112|74349_	CYP82E10 [Nicotiana tabacum]	55.172	4.41E-29
	Colchicum_20101112|54622_	CYP82E10 [Nicotiana tabacum]	51.24	3.41E-34
	Colchicum_20101112|39301_	CYP82E10 [Nicotiana tabacum]	50	4.66E-22
	Colchicum_20101112|75506_	CYP82E10 [Nicotiana tabacum]	50	4.23E-34
	Gloriosa-20120814|5091_1	CYP82E4 [Nicotiana tabacum]	53.488	1.63E-24
	Colchicum_20101112|74349_	CYP82E4 [Nicotiana tabacum]	54.023	1.36E-27
	Colchicum_20101112|75506_	CYP82E4 [Nicotiana tabacum]	50	8.59E-32
Step 5: NAT	Gloriosa-20120814|32279_1	BAG90782.1 [Oryza sativa Japonica Group]	67.702	1.42E-74
	Colchicum_20101112|8688_	BAG90782.1 [Oryza sativa Japonica Group]	71.034	3.05E-67
	Gloriosa-20120814|70209_1	Q5KQI6.1 [Oryza sativa Japonica Group]	69.136	1.54E-108
	Gloriosa-20120814|51165_1	Q5KQI6.1 [Oryza sativa Japonica Group]	68.016	2.70E-106
	Gloriosa-20120814|30717_1	Q5KQI6.1 [Oryza sativa Japonica Group]	64.615	6.09E-104
	Colchicum_20101112|6813_	Q5KQI6.1 [Oryza sativa Japonica Group]	76.768	5.79E-109
	Colchicum_20101112|37095_	Q5KQI6.1 [Oryza sativa Japonica Group]	70.476	1.26E-102

transcript in the *G. superba* transcriptome (*Gloriosa*-20120814_33123). The putative *NMT* is a likely candidate for an *N*-methylation step that converts the trihydroxylated phenethylisoquinoline intermediate to (*S*) autumnaline, the first alkaloidal precursor formation step in the proposed colchicine pathway [52, 76]. Further, compared *G. superba* transcriptome to the reference sequence of *C. japonica* coclaurine *NMT* (BAB74802.1) gene, *Gloriosa*-20120814|64082_1 transcript showed an identity with 58%. Moreover, in *Thalictrum flavum*, pavine *NMT* (*PNMT*) converts (*S*)-tetrahydropapaverine to (*S*)-laudanosine in the benzylisoquinoline alkaloid pathway that is common in many plants [77]. *PNMT* is also considered a possible alternative reference sequence due to its capability of adding a methyl group to (S)-tetrahydropapaverine where the nitrogen is present, which is similar to the colchicine pathway mechanism. In addition, reticuline *NMT* involved in biosynthesis of the aporphine alkaloid magnoflorine in opium poppy roots [78]. When comparted *C. autumnale* transcriptome to the reference sequence of *T. flavum PNMT* (sp|C3SBW0.1) gene, *Colchicum*_20101112|85 transcript showed an identity of 48.7% (Table 1 and Figure 9).

Third, the cDNA clone (*Gloriosa* 8E03) identified as a putative P450 is a partial clone with 87% sequence homology to *Narcissus pseudonarcissus CYP96T1*, which catalyzes a C-C phenol coupling reaction in noroxomaritidine biosynthesis in the haemanthamine pathway and has an identical full-length transcript in the *C. autumnale* transcriptome (*Colchicum*-20101112_3005). This putative enzyme might be a candidate for the *para-para* phenol-tropolone oxidative coupling bridge-forming P450: *CYP96T1* that is NADPH and O_2-dependent and converts (*S*)-autumnaline to isoandrocymbine [79-80]. In *Papaver somniferum,* salutaridine synthase enzyme *CYP719B1* was responsible for the C-C phenol coupling converting (*R*)-reticuline to salutaridine by

connecting the 12 and 13 carbon (81-82). *P. somniferum* CYP719B1 enzyme could also considered an alternative reference gene because of its C-C phenol-coupling mechanisms in the morphine pathway. Furthermore, in *C. japonica*, cytochrome P450 enzyme *CYP80G2* has shown to convert (*S*)-reticuline to (*S*)-corytuberine through C-C phenol-coupling in an isoquinoline alkaloid pathway [83]. When compared *G. superba C. autumnale* and transcriptome with CYP80G2 (sp|A8CDR5.1) gene, which revealed *Gloriosa*-20120814|30999_1 and *Colchicum*_20101112|76071 transcripts showed an identity of 52.7% and 51.7% (Table 1 and Figure 9).

Fourth, the cDNA clone (*Gloriosa* 7D05) identified as a putative *OMT* is a partial clone with high sequence homology to *Narcissus* sp. aff. *pseudon6rcissus* norbelladine-4-OMT, which participates in a similar *O*-methylation reaction to form *O*-methylnorbelladine in the galanthamine pathway and has an identical full-length transcript in the *G. superba* transcriptome (*Gloriosa*-20120814_6585 and *Gloriosa*-20120814|82787_1). The putative *OMT* are a likely candidate for a *3-O*-methylation step that catalyzes the conversion of isoandrocymbine to *O*-methylandrocymbine [84]. Additionally, two reference *3-OMT* sequences such as AFB74613.1 from *P. somniferum* and sp|A0A077EWA5.1| from *N. pseudonarcissus* were considered when filtering *G. superba C. autumnale* transcriptomes for possible alternative *3-OMT* colchicine pathway genes. Sequence analysis of reference genes to *G. superba* transcriptome revealed, *Gloriosa*-20120814|16875_1 and *Gloriosa*-20120814|82787_1 showed an identity of 50.9% and 72.8%, respectively. The *Colchicum*_20101112|23611 showed an identity of 58.1% when compared to *N. pseudonarcissus* (Table 1 and Figure 9). *3-OMT* is a class I methyltransferase, which is involved in plant alkaloid specialized metabolism of several significant structural features such as, canonical 'Rossmann-

like' fold of their *S*-adenosylmethionion (SAM)-binding domain and a conserved S_n2 reaction, where an electrophilic methyl carbon on SAM is attacked by an election lone pair of a nucleophilic substrate. This including N-terminal dimerization domains, glycine-rich SAM-binding regions, and metal-binding sites [85]. In *P. somniferum 3-OMT* catalyzes the conversion of *4'-O*-desmethyl-3-*O*-acetylpapaveroxine to *3-O*-acetylpapaveroxine by forming a heterodimer with *2-OMT* in the noscapine pathway [86].

Fifth, the cDNA clone (*Gloriosa* 1F02) identified as another putative P450 is a partial clone with 79% sequence homology to *Fragaria x ananassa* and *Nicotiana tabacum CYP82E10*, which catalyze a *N*-demethylation reaction from nicotine to nornicotine and has a full-length transcript in the *C. autumnale* transcriptome (*Colchicum*-20101112_5364). This putative P450 enzyme might be a candidate for the *N*-demethylation that converts demecolcine to deacetylcolchicine [87]. Nevertheless, in *N. tabacum* three cytochrome P450 superfamily of monooxygenases that catalyze the *N*-demethylation of nicotine have been identified, *CYP82E4, CYP82E5v2*, and *CYP82E10* [88]. *CYP82E4* is the major nicotine demethylase enzyme responsible for converting nicotine to nornicotine [89]. The demethylase P450 enzymes *CYP82E5v2* and *CYP82E10* were expressed in nonsenescent green leaves and/or root tissue in *N. tabacum* [87, 89]. Therefore, the *G. superba C. autumnale* transcriptomes were compared with an alternative reference gene NP_001312976.1 (*CYP82E4*), which reveled *Gloriosa*-20120814|5091_1 and *Colchicum*_20101112|74349 showed an identity of 53.4%, and 54%, respectively (Table 1 and Figure 9). Thus, enzymes belonging to the predicated *NMT, 3-OMT*, *CYP96T1, CYP82E10*, and *NAT* families are likely to be early targets for cloning and characterization of colchicine pathway. However, the possibility exists that one or more of the candidate genes will no function as predicted. This case, other genes from the *G.*

superba and *C. autumnale* transcriptomes that homology to colchicine pathway enzymes could be also consider candidate genes.

Correlation of expression patterns is a widely used criterion to predict biological functions of genes, functional relatedness between genes, and gene regulatory networks [90-93]. Thus, an expanded set of genes, including the best candidates and several related (based on homology) genes for each step will be analyzed by qRT-PCR to validate the expression levels of the candidate genes in different parts of the developing biorhizome compared to other tissues in the plant. Once we have confirmed that specific genes are likely candidates for specific steps in the pathway, enzyme characterization, including pathway reconstruction in a heterologous expression system, will be performed to further validate their involvement in colchicine biosynthesis. The detailed analyses will involve biochemical characterization as previously performed to characterize *NMT, P450s, 3-OMT* and *NAT* enzymes in other alkaloid biosynthetic pathways [74-89].

What genes are candidate rhizome developmental genes from G. superba and *C. autumnale?* Underground rhizome (storage organ) development, enlargement molecular mechanism and dynamics are remains unclear in medicinal plants. Therefore, biorhizome can be a suitable system to study rhizome growth and biomedicine biomanufacturing. In addition, it will be helpful to identify candidate genes related to rhizome development, which is key resources to promote the improvement of biomass yield in biorhiozme. Expressions of the below known genes affect rhizome formation in different rhizomatous crops but have not been characterized in Colchicaceae. Understanding of rhizome developmental genes from the *G. superba* and *C. autumnale* could allow molecular intricacies that are involved in the biorhizome biomass production in biomanufacturing. To screen the rhizome developmental genes in colchicine producing species, transcriptome from

G. superba and *C. autumnale* were analyzed with the aspirations to track the rhizome metabolism. For annotation, all transcriptomes were mapped to the reference sequence. *GI* is a sensor gene that involved in seasonal growth, which includes rhizome development in potato [64, 94-96]. The reference *GI* sequence from *N. nucifera* transcriptome (gi|720040388|ref|XP_010268589.1|) and *S. tuberosum* transcriptome (PGSC0003DMT4000048370) were compared with *G. superba* and *C. autumnale* transcriptomes. Two potential full-length candidate genes *Gloriosa*-20120814|8521_1 and *Gloriosa*-20120814|11998_1 1 shared an identity of 70%, and 76%, respectively (Table 2 and Figure 10).

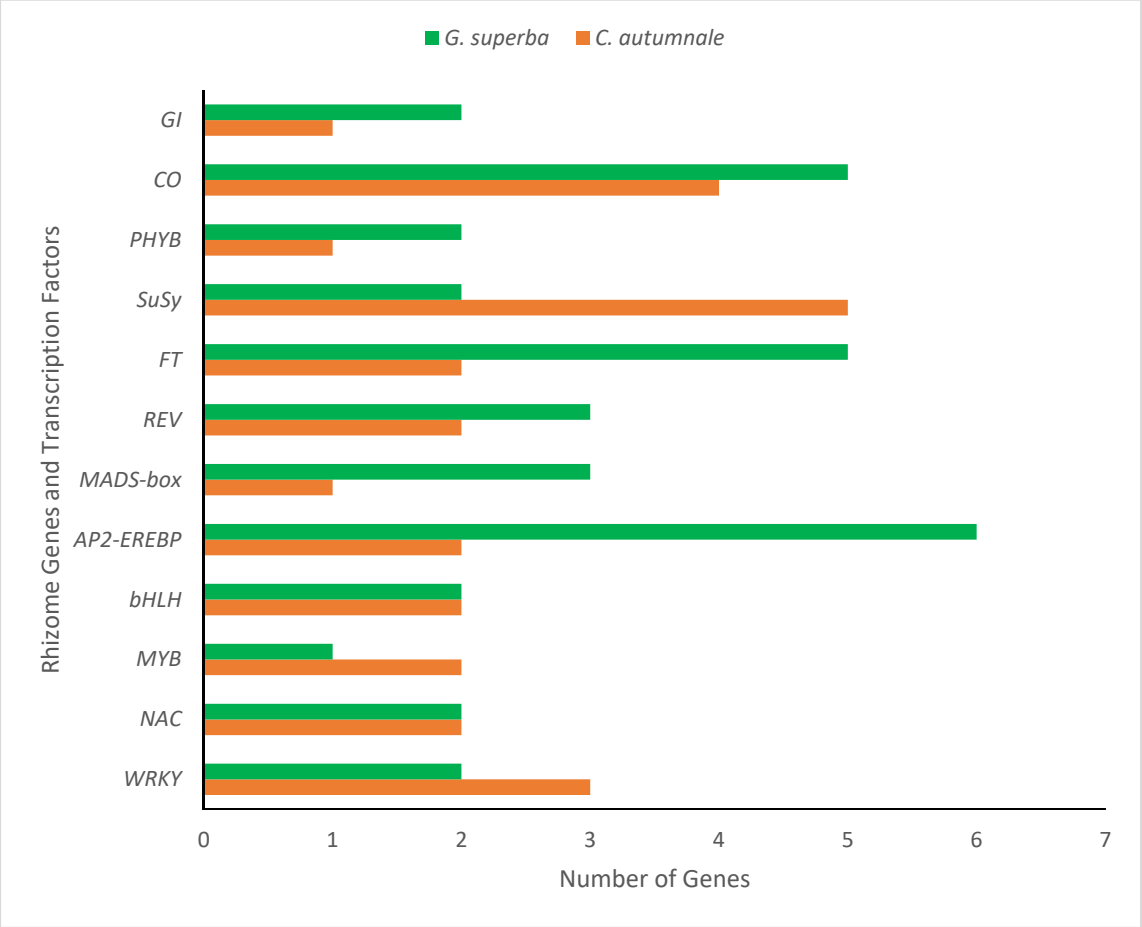

Figure 10. Possible candidate rhizome developmental genes and transcription factors in *Gloriosa superba* and *Colchicum autumnale* transcriptomes.

Table 2. Predicted candidate rhizome developmental genes in *Gloriosa superba* and *Colchicum autumnale* transcriptomes.

Rhizome Developmental Genes	Candidate Gene ID	Reference Gene ID	% Identity	e-Value
GIGANTEA	Colchicum_20101112|2617_	XP_010261025.1 *Nelumbo nucifera*	82.004	0
	Gloriosa-20120814l11998_1	XP_010268589.1 *Nelumbo nucifera*	76.132	0
	Gloriosa-20120814|8521_1	PGSC0003DMT4000048370 *Solanum tuberosum**	69.743	0
	Colchicum_20101112|2617_	PGSC0003DMT4000048370 *S. tuberosum**	82.609	1.68E-173
CONSTANS	Gloriosa-20120814|51159_1	XP_010261698.1 *Nelumbo nucifera*	84.091	7.14E-12
	Colchicum_20101112|3159_	XP_010257847.1 *Nelumbo nucifera*	81.818	1.79E-11
	Gloriosa-20120814|51159_1	PGSC0003DMT400067656 *S. tuberosum**	75.51	1.14E-13
	Colchicum_20101112|3159_	PGSC0003DMT400026065 *S. tuberosum**	81.818	4.61E-12
	Gloriosa-20120814|17037_1	NP_001274795.1 *Solanum tuberosum*	72.5	4.85E-30
	Gloriosa-20120814|51159_1	NP_001274795.1 *Solanum tuberosum*	68.627	3.75E-10
	Gloriosa-20120814|18001_1	NP_001274795.1 *Solanum tuberosum*	68.627	4.24E-10
	Gloriosa-20120814|20079_1	NP_001274795.1 *Solanum tuberosum*	67.442	2.97E-11
	Gloriosa-20120814|72100_1	NP_001274795.1 *Solanum tuberosum*	64	4.59E-29
	Colchicum_20101112|3159_	NP_001274795.1 *Solanum tuberosum*	81.818	2.96E-10
	Colchicum_20101112|5391_	NP_001274795.1 *Solanum tuberosum*	73.75	1.17E-23
	Colchicum_20101112|27602_	NP_001274795.1 *Solanum tuberosum*	79.365	7.90E-29
	Colchicum_20101112|26221_	NP_001274795.1 *Solanum tuberosum*	71.795	1.62E-12
Phytochrome B	Gloriosa-20120814|36050_1	XP_010267948.1 *Nelumbo nucifera*	81.445	0
	Colchicum_20101112|1221_	XP_010267948.1 *Nelumbo nucifera*	80.142	0
	Gloriosa-20120814|85746_1	PGSC0003DMT400061712 *S. tuberosum**	80.927	0
	Colchicum_20101112|1221_	PGSC0003DMT400061712 *S. tuberosum**	79.539	0
Sucrose Synthase	Gloriosa-20120814|69933_1	XP_010271909.1 *Nelumbo nucifera*	81.886	0
	Colchicum_20101112|13826_	XP_010271909.1 *Nelumbo nucifera*	81.39	0
	Colchicum_20101112|593_	XP_010271909.1 *Nelumbo nucifera*	80.15	0
	Colchicum_20101112|2693_	XP_010271909.1 *Nelumbo nucifera*	81.404	0
	Colchicum_20101112|33099_	XP_010271909.1 *Nelumbo nucifera*	71.535	0
	Colchicum_20101112|583_	XP_010271909.1 *Nelumbo nucifera*	71.411	0
	Gloriosa-20120814|9073_1	PGSC0003DMT400007506 *S. tuberosum**	76.923	0
	Colchicum_20101112|2693_	PGSC0003DMT400007506 *S. tuberosum**	76.289	0
Flowering Locus T	Colchicum_20101112|8829_	XP_010268289.1 *Nelumbo nucifera*	80.925	1.75E-109
	Colchicum_20101112|22914_	XP_010268289.1 *Nelumbo nucifera*	86.232	1.56E-92
	Gloriosa-20120814|55554_1	XP_010268289.1 *Nelumbo nucifera*	83.815	2.85E-113
	Colchicum_20101112|22914_	PGSC0003DMT400060057 *S. tuberosum**	78.788	9.86E-81
	Colchicum_20101112|8829_	PGSC0003DMT400060057 *S. tuberosum**	75	2.51E-97
	Gloriosa-20120814|55554_1	PGSC0003DMT400060057 *S. tuberosum**	75.595	4.17E-99
	Colchicum_20101112|8829_	NP_001274897.1 *Solanum tuberosum*	73.81	5.00E-93
	Colchicum_20101112|22914_	NP_001274897.1 *Solanum tuberosum*	77.273	1.98E-76
	Gloriosa-20120814|55554_1	NP_001274897.1 *Solanum tuberosum*	74.405	8.62E-95
	Gloriosa-20120814|55004_1	NP_001274897.1 *Solanum tuberosum*	67.066	3.03E-82
	Gloriosa-20120814|26010_1	NP_001274897.1 *Solanum tuberosum*	64.968	1.36E-71
	Gloriosa-20120814|45011_1	NP_001274897.1 *Solanum tuberosum*	72.034	4.46E-63
	Gloriosa-20120814|52745_1	NP_001274897.1 *Solanum tuberosum*	60	1.55E-66
REVOLUTA	Gloriosa-20120814|11056_1	XP_010247499.1 *Nelumbo nucifera*	86.244	0
	Colchicum_20101112|467_	XP_010247499.1 *Nelumbo nucifera*	85.221	0
	Colchicum_20101112|467_	PGSC0003DMT400030829 *S. tuberosum**	81.02	0
	Gloriosa-20120814|17486_1	PGSC0003DMT400030829 *S. tuberosum**	81.182	0
	Gloriosa-20120814|15982_1	AAY32332.1 *Phyllostachys praecox*	78.851	0
	Colchicum_20101112|6503_	AAY32332.1 *Phyllostachys praecox*	78.091	0

*Sequence from *Solanum tuberosum* transcriptome [https://plants.ensembl.org/Solanum_tuberosum/Info/Index]

CO gene play an important role in the circadian clock and rhizome development in radish and potato [97-98]. Five full-length sequences reveled *CO* functionality in *Gloriosa*-20120814|18001_1, *Gloriosa*-20120814|51159_1, *Gloriosa*-20120814|17037_1, *Gloriosa*-20120814|20079_1, and *Gloriosa*-20120814|72100_1. The reference sequence from *S. tuberosum* (NP_001274795.1) shared an identity with candidate sequences of 64%-73%. While *Colchicum*_20101112|3159, *Colchicum*_20101112|5391, *Colchicum*_20101112|26221, and *Colchicum*_20101112|27602 shared an identity of 72%-82%. The alternate reference sequence from *S. tuberosum* (PGSC0003DMT400067656) shared an identity with candidate sequences *Gloriosa*-20120814|51159_1 of 76% and *Colchicum*_20101112|3159 of 82%. The reference sequence from *N. nucifera* (gi|719967386|ref|XP_010261698.1|) shared an identity with candidate sequences *Gloriosa*-20120814|51159_1 of 84% and *Colchicum*_20101112|3159 of 82% (Table 2 and Figure 10).

FT and *FT*-like protein StSP6A are key components of tuberigen and a systemic floral inducer in potato [99]. By employing computational analysis using RNA-Seq data, we identified five full-length *FT* candidate genes such as *Gloriosa*-20120814|55554_1, *Gloriosa*-20120814|55004_1, *Gloriosa*-20120814|26010_1, *Gloriosa*-20120814|45011_1, and *Gloriosa*-20120814|52745_1), while two candidate genes from *Colchicum*_20101112|8829 and *Colchicum*_20101112|22914. The *FT* reference sequence from *S. tuberosum* (StSP6A NP_001274897.1) shared an identity range with candidate genes of 60%-74%. The alternative *FT* reference sequence from *S. tuberosum* (PGSC0003DMT400060057) shared an identity with candidate sequence *Gloriosa*-20120814|55554_1 of 76% and two sequences with *Colchicum*_20101112|8829 and *Colchicum*_20101112|22914 of 75% and 79%, respectively. The *FT* reference sequence from *N. nucifera* (gi|720039388|ref|XP_010268289.1|) shared an identity with *Gloriosa*-

20120814|55554_1 of 84% and two candidate genes with *Colchicum*_20101112|8829 and *Colchicum*_20101112|22914 of 81% and 86%, respectively (Table 2 and Figure 10).

The photoreceptor *PHYB* is involved tuber induction and microRNA, miR172 highly expressed in potato tuber [100]. It has been shown in *N. nucifera* that *PHYB* and/or other phytochromes might measure the length of the light period to affect rhizome girth enlargement [64]. The *PHYB* candidate gene *Gloriosa*-20120814|36050_1 and *Colchicum*_20101112|1221 shared an identity of 81% with the reference sequence from *N. nucifera* (gi|720038316|ref|XP_010267948.1|). The *PHYB* reference sequence from *S. tuberosum* (PGSC0003DMT400061712) shared an identity of 81% with candidate sequence *Gloriosa*-20120814|85746_1) and 80% with *Colchicum*_20101112|1221 (Table 2 and Figure 10).

SuSy play a key role in tubers biomass allocation, starch biosynthesis and storage, which also cleaves sucrose into fructose and UDP-glucose [101-102]. Two full-length *SuSy* sequences were identified in *G. superba* such as *Gloriosa*-20120814|9073_1 and *Gloriosa*-20120814|69933_1); whereas five candidate genes in *C. autumnale* *Colchicum*_20101112|583, *Colchicum*_20101112|593, *Colchicum*_20101112|2693, *Colchicum*_20101112|13826, and *Colchicum*_20101112|33099, which shared an identity of 81-82% with *N. nucifera* (gi|720050864|ref|XP_010271909.1|). The reference sequence from *S. tuberosum* (PGSC0003DMT400007506) shared an identity of 78% with candidate sequence *Colchicum*_20101112|2693 (Table 2 and Figure 10).

REV plays an important role during morphogenesis and controlling the apical meristem formation during rhizome development in bamboo [60, 103]. Since, biorhizomes are compressed scale leaves that replicate rhizomes possessing vegetative buds [47], at the molecular level an intricate

regulatory network determines initial biorhizome leaves development and *REV* could antagonistically regulates lamina structures that triggers specialized organ. The *REV* reference sequence from *N. nucifera* (gi|720097995|ref|XP_010247499.1|) shared an identity with candidate *Gloriosa*-20120814|11056_1 sequence of 86% and *Colchicum*_20101112|467 of 85%. The reference sequence from *S. tuberosum* (PGSC0003DMT400030829) shared an identity with *Gloriosa*-20120814|17486_1 of 81% and *Colchicum*_20101112|467 of 81%. Additionally, the *REV*-like homolog (*PpHB1)* was found in bamboo, which highly expressed in the tips of lateral buds at several developmental stages [104]. BLAST analysis of *Gloriosa*-20120814|15982_1 and *Colchicum*_20101112|6503 showed a 79% and 78% identity when compared to *PpHB1*, respectively (Table 2 and Figure 10).

What are the transcription factors in G. superba and *C. autumnale?* Currently, no information is known about the transcription factors in *G. superba* and *C. autumnale*. Besides rhizome developmental genes, some important subset of rhizome genes transcription factors were also identified. It was suggested that *MADS-box* (an acronym for mini chromosome maintenance 1), *AP2-EREBP* (APETALA2/ethylene-responsive element binding protein), *bHLH* (basic helix-loop-helix), *MYB* (myeloblastosis related), *NAC* (no apical meristem), and *WRKY* (contain the highly conserved amino acid sequence WRKYGQK and the zinc-finger-like motifs Cys(2)-His(2) or Cys(2)-HisCys, and bind to the TTGAC(C/T) W-box *cis*-element in the promoter of their target genes) proteins can act as rhizome developmental transcription factors [105-108]. Several genome-wide analyses have been conducted on these transcription factors in *S. tuberosum, N. nucifera* and *Sinopodophyllum hexandrum* [66, 105, 109]. The below screened candidate transcription factors could contribute to our understanding of the molecular mechanism of biorhizome development,

by identifying candidate colchicine biosynthetic gene family members together with the rhizome developmental genes and transcription factors in *G. superba* and *C. autumnale*.

MADS-box genes were present on all 12 potato chromosomes, the StMADS1 and StMADS13 were most likely to be downstream target of StSP6A and involved in tuber development [105]. Therefore, the known *S. tuberosum* and *N. nucifera MADS-box* sequences were used as query to perform BLAST against the *G. superba* and *C. autumnale* protein databases. One full-length candidate sequence *Gloriosa*-20120814|19828_1 shared an identity of 77% with *N. nucifera* (gi|720053055|ref|XP_010272608.1|). The reference sequence from *S. tuberosum* (PGSC0003DMT400000026) shared an identity of 81 and 82% with candidate sequences *Colchicum*_20101112|22985 and *Gloriosa*-20120814|35037_1, respectively (Table 3 and Figure 10).

AP2 (APETALA2) and *EREBPs* are one of the largest and the prototypic family of transcription factors unique to plant, which is upregulated in rhizomes [110-111]. Potential four *AP2-EREBP* candidate genes were *Gloriosa*-20120814|77954_1, *Gloriosa*-20120814|7267_1, *Gloriosa*-20120814|54352_1, and *Gloriosa*-20120814|45526_1 screened from the reference sequences *N. nucifera* (gi|720011136|ref|XP_010259468.1|) with an identity 66%, 69%, 82%, and 86%, respectively . In addition, two full-length alternative sequences were identified using reference sequence from *S. tuberosum* (PGSC0003DMT400016585) from the *C. autumnale* such as *Colchicum*_20101112|4936 and *Colchicum*_20101112|10921 with an identity of 76% and 82%. Of the six *AP2-EREBP* candidate genes, two were found (*Gloriosa*-20120814|7267_1 and Gloriosa-20120814|77954_1) within *G. superba* leaf tissue cDNA (*Gloriosa* 8E03), which both had an identity of 100% (Table 3 and Figure 10).

Table 3. Predicted candidate transcription factors in *Gloriosa superba* and *Colchicum autumnale* transcriptomes.

Transcription Factors	Candidate Gene ID	Reference Gene ID	% Identity	e-Value
MADS-box	Gloriosa-20120814|19828_1	XP_010272608.1 *Nelumbo nucifera*	77.366	2.02E-138
	Gloriosa-20120814|35037_1	PGSC0003DMT400000026 *Solanum tuberosum**	82.432	3.32E-42
	Gloriosa-20120814|39825_1	PGSC0003DMT400000026 *Solanum tuberosum**	60.317	9.37E-23
	Colchicum_20101112|22985_	PGSC0003DMT400000026 *Solanum tuberosum**	81.159	3.61E-37
AP2-EREBP	Colchicum_20101112|4936_	XP_010263474.1 *Nelumbo nucifera*	73.684	3.88E-88
	Gloriosa-20120814|45526_1	XP_010271250.1 *Nelumbo nucifera*	86.047	9.56E-22
	Gloriosa-20120814|54352_1	XP_010271250.1 *Nelumbo nucifera*	82.353	1.69E-16
	Gloriosa-20120814|7267_1	XP_010271250.1 *Nelumbo nucifera*	68.831	3.04E-32
	Gloriosa-20120814|58739_1	XP_010271250.1 *Nelumbo nucifera*	66.667	1.15E-18
	Gloriosa-20120814|77954_1	XP_010271250.1 *Nelumbo nucifera*	65.854	4.44E-33
	Gloriosa-20120814|23228_1	PGSC0003DMT400016585 *Solanum tuberosum**	80.905	1.16E-97
	Colchicum_20101112|10921_	PGSC0003DMT400016585 *Solanum tuberosum**	75.701	6.49E-99
	Colchicum_20101112|4936_	PGSC0003DMT400016584 *Solanum tuberosum**	83.929	5.12E-64
bHLH	Gloriosa-20120814|6279_1	XP_010257880.1 *Nelumbo nucifera*	89.744	3.68E-18
	Colchicum_20101112|7042_	XP_010273628.1 *Nelumbo nucifera*	81.818	1.93E-48
	Gloriosa-20120814|53746_1	PGSC0003DMT400022702 *Solanum tuberosum**	88.889	1.27E-11
	Colchicum_20101112|21881_	PGSC0003DMT400022701 *Solanum tuberosum**	86.364	9.32E-08
MYB	Gloriosa-20120814|15461_1	XP_010277005.1 *Nelumbo nucifera*	90.698	8.02E-86
	Colchicum_20101112|11314_	XP_010253806.1 *Nelumbo nucifera*	87.097	1.43E-14
	Gloriosa-20120814|15461_1	PGSC0003DMT400012203 *Solanum tuberosum**	90.698	1.34E-87
	Colchicum_20101112|20668_	PGSC0003DMT400017709 *Solanum tuberosum**	86.301	2.98E-41
NAC	Gloriosa-20120814|28710_1	XP_010243512.1 *Nelumbo nucifera*	84.397	1.10E-89
	Colchicum_20101112|44481_	XP_010270427.1 *Nelumbo nucifera*	71.429	1.09E-07
	Gloriosa-20120814|18396_1	PGSC0003DMT400045294 *Solanum tuberosum**	81.212	2.64E-101
	Colchicum_20101112|794_	PGSC0003DMT400079789 *Solanum tuberosum**	71.429	7.85E-91
WRKY	Colchicum_20101112|22998_	XP_010258216.1 *Nelumbo nucifera*	80.952	5.53E-07
	Gloriosa-20120814|20164_1	XP_010258120.1 *Nelumbo nucifera*	76.596	7.13E-50
	Gloriosa-20120814|66904_1	PGSC0003DMT400028529 *Solanum tuberosum**	77.419	3.60E-28
	Colchicum_20101112|22998_	PGSC0003DMT400072835 *Solanum tuberosum**	80.952	5.16E-07
	Colchicum_20101112|49701_	ALD83482.1 *Sinopodophyllum hexandrum*	80	2.46E-14
	Colchicum_20101112|3957_	ALD83482.1 *Sinopodophyllum hexandrum*	76	2.38E-19

*Sequence from *Solanum tuberosum* transcriptome [https://plants.ensembl.org/Solanum_tuberosum/Info/Index]

The *bHLH* superfamily of proteins is the second largest transcription factor family in plants and *StbHLH76* and *StbHLH86* had a relatively high expression level in the potato tuber compared with in other tissues [112]. The full-length sequence of *Gloriosa*-20120814|6279_1 demonstrated possible *bHLH* functionality with reference sequence from *N. nucifera* (gi|720006121|ref|XP_010257880.1|) with an identity of 90%. The reference sequence from *S. tuberosum* (PGSC0003DMT400022702) shared an identity 89% with *bHLH* candidate sequence *Gloriosa*-20120814|53746_1). Additionally, the reference sequence from *N. nucifera* (gi|719971259|ref|XP_010273628.1|) shared an identity of 82% with *bHLH* candidate sequence

with *C. autumnale Colchicum*_20101112|7042. Likewise, *S. tuberosum* (PGSC0003DMT400022701) shared an identity of 86% with candidate sequence *Colchicum*_20101112|21881 (Table 3 and Figure 10)

The *MYB* transcription factors are consider potentially important regulators of secondary metabolism and 1R-, R2R3-, 3R-, and 4R-, where the R2R3-*MYB* proteins are specific to plants [113-115]. The reference sequence from *N. nucifera* (gi|719972251|ref|XP_010277005.1|) shared an identity of 91% with *MYB* candidate sequence (Gloriosa-20120814|15461_1). The reference sequence from *S. tuberosum* (PGSC0003DMT400012203) shared and identity of 91% with *MYB* candidate sequence *Gloriosa*-20120814|15461_1. In *C. autumnale,* candidate sequence *Colchicum*_20101112|11314 shared an identity of 87% with *N. nucifera* (gi|719993192|ref|XP_010253806.1|). Likewise, in *S. tuberosum* (PGSC0003DMT400017709) shared and identity of 86% with candidate sequence *Colchicum*_20101112|20668 (Table 3 and Figure 10).

There is abundant evidence indicating that *NAC* proteins play crucial roles in hormone signaling, lateral root development and upregulated during rhizome formation [64, 116-117]. *NAC* candidate sequence *Gloriosa*-20120814|28710_1 shows 84% identity with reference gene from *N. nucifera* (gi|720085419|ref|XP_010243512.1|). Alternative gene *Gloriosa*-20120814|18396_1 shared identity of 81% with *S. tuberosum* (PGSC0003DMT400045294). The reference sequence from *N. nucifera* (gi|720046186|ref|XP_010270427.1|) shared an identity of 71% with candidate sequence *Colchicum*_20101112|44481 and *S. tuberosum* (PGSC0003DMT400079789) shared an identity 71% with *Colchicum*_20101112|794 (Table 3 and Figure 10)

WRKY can regulate diverse responses in rhizome network of genes including sprouting mechanism in potato and secondary metabolisms [118-119]. Two full-length candidate sequences *Colchicum*_20101112|3957 and *Colchicum*_20101112|49701 shared *WRKY* functionality with the reference sequence from *Sinopodophyllum hexandrum* (ALD83482.1) of 76% and 80%, respectively. Among them *Gloriosa*-20120814|20138_1 shared an identity of 72%. The reference sequence from *N. nucifera* (gi|720006863|ref|XP_010258120.1|) shared an identity of 77% with *Gloriosa*-20120814|20164_1. The reference sequence from *S. tuberosum* (PGSC0003DMT400028529) shared an identity of 77% with candidate sequence *Gloriosa*-20120814|66904_1. The alternative reference sequence from *N. nucifera* (gi|720007162|ref|XP_010258216.1|) shared an identity of 81% with candidate sequence Colchicum_20101112|22998 and *S. tuberosum* (PGSC0003DMT400072835) shared and identity percentage with Colchicum_20101112|22998_ of 81% (Table 3 and Figure 10).

Phylogenetic analysis of colchicine pathway, rhizome developmental, and transcription factor protein in G. superba and C. autumnale: A rooted tree of total possible 169 full-length amino-acid candidate sequences such as colchicine pathway, rhizome developmental, and transcription factors of *G. superba* and *C. autumnale* was constructed using the iTOL Neighbor joining software (Figure 11). A total of 17 and 22 colchicine pathway enzymes, 29 and 20 rhizome developmental, and 47 and 34 transcription factors were phylogenetically mapped in *G. superba* and *C. autumnale*, respectively. The transcription factor clade represented the largest protein cluster followed by the rhizome developmental and colchicine pathway genes. The phylogenetic tree indicates that the colchicine pathway, rhizome developmental, and transcription factor gene families of *G. superba* and *C. autumnale* were found uniquely to Colchicaceae. This suggests that the predicted candidate

genes not only are involved in colchicine biosynthesis in rhizomes but are also involved in the rhizome metabolism. In comparison with tuber producing *S. tuberosum* and *N. nucifera*, rhizome developmental and transcription factor genes from *G. superba* and *C. autumnale* showed nearly the same transcriptional regulation map that are possible downstream targets of rhizome initiation or biomass production. The annotated *G. superba* and *C. autumnale* can contributes to discovering the candidate genes in biosynthesis of different groups of secondary metabolites in biorhizome and rhizome developmental genes, which could serve as a comprehensive resource for molecular mechanism research of colchicine biosynthesis in *G. superba*. This provides a molecular platform and resource for future genetic and functional rhizomatous medicinal crops genomic research.

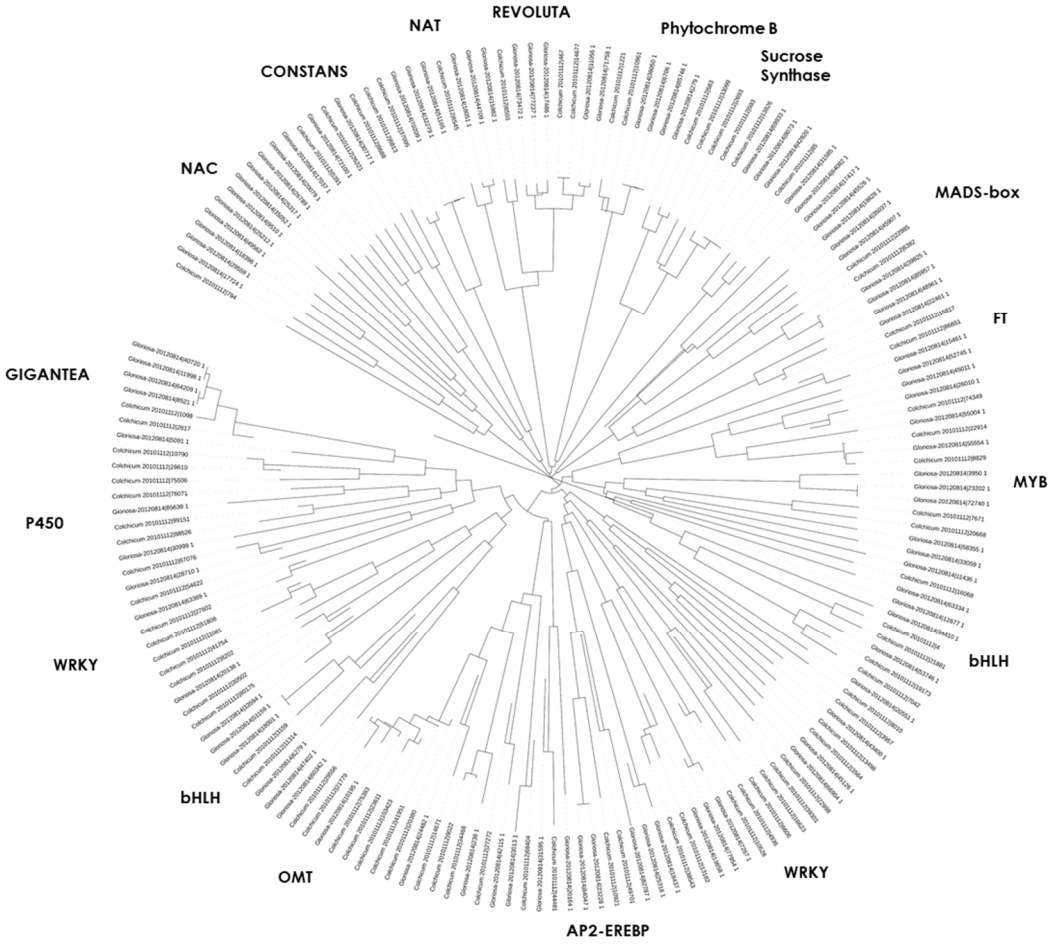

Figure 11. Phylogenetic tree of *Gloriosa superba* and *Colchicum autumnale* possible candidate colchicine pathway, rhizome developmental genes, and transcription factors.

MATERIALS AND METHODS

cDNA library

A full length cDNA library was constructed from a month old *G. superba* biorhizome derived leaves. The total RNA isolation, cDNA library construction were performed per previous protocol [120]. The assembled cDNA sequences were BLASTed against the National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/), and *G. superba* and *C. autumnale* transcriptomes.

Prediction of colchicine pathway and rhizome developmental proteins

The *G. superba* and *C. autumnale* transcriptomes were obtained from the Medicinal Plant Transcriptome database (http://www.medplantrnaseq.org). BUSCO analysis of transcriptomes were processed according to Waterhouse et al. [121]. Colchicine pathway and rhizome developmental along with its transcription factors encoding proteins were predicted from the NCBI database.

Functional annotation of transcriptomes, colchicine pathway and rhizome developmental proteins

The *G. superba* and *C. autumnale* transcriptomes GO classification of the identified know plant proteins was performed using the web-accessible Blast2GO v5 annotation system (https://www.blast2go.com/) [122]. Blast2GO is an all-in-one bioinformatics software for protein functional prediction and the genome-wide analysis of annotation data. KEGG pathway enrichment analysis was used to analyze the functional significance of biochemical pathways using BlastKOALA and GhostKOALA. The first step in Blast2GO is to cDNA nucleotide sequences against the NCBI non-redundant database by Basic Local Alignment

Search Tool protein (BLASTp/BLASTn) with an expectation value of $1e^{-5}$. Next, transcriptomes FASTA protein sequences were uploaded to Blast2GO for BLAST analysis to identify homologous sequences, mapping and annotation were performed. The significantly enriched biological processes, molecular function, cellular component and KEGG pathway were identified by *p* value less than threshold value 0.05. The colchicine pathway candidate genes with their specific enzymatic function and high homology were identified through alkaloids metabolic reference genes such as *NMT; 3-OMT; two cytochrome P450s: CYP96T1 and CYP82E10*; and *NAT* [74-89]. For rhizome developmental genes namely *GI, CO, PHYB, SuSy, FT and REV* as well as transcription factors such as *MADS-box, AP2-EREBP, bHLH, MYB, NAC*, and *WRKY* were designated as reference genes from the *S. tuberosum* and *N. nucifera* due to their known nature as rhizome producers [https://plants.ensembl.org/Solanum_tuberosum/Info/Index and ftp://ftp.ncbi.nih.gov/genomes/].

Phylogenetic analysis of possible total colchicine pathway, rhizome developmental, and transcription factor proteins

Gloriosa superba and *Colchicum autumnale* possible candidate colchicine pathway, rhizome developmental, and transcription factor protein sequences were aligned using ClustalX version 1.83. The phylogenetic tree was constructed from iTOL using Neighbor joining method [https://itol.embl.de/] [105].

CONCLUSIONS

In this study, we used two different transcriptomes such as *G. superba* and *C. autumnale* (evolutionarily diverse, different from *Gloriosa* species) to predict genes that encode proteins in colchicine biosynthesis and rhizome metabolism, a rhizomatous medicinal plants that lacks a sequenced genome. Collectively, transcriptomes and cDNA approaches were applied to uncovered candidate genes for each predicted colchicine pathway step, which could help to elucidate the colchicine biosynthetic pathway. Additionally, our work would be useful to establish the basic information of rhizome developmental genes and transcription factors in *G. superba* and *C. autumnale*, which are aimed to improve the biorhizome biomass. The predicted genes in this work need to be cloned, expressed, or their functions validated, which are an important resource for metabolic engineering or synthetic biotechnology that could improve the colchicine biomanufacturing.

ABBREVIATIONS

AP2-EREBP	APETALA2/ethylene-responsive element binding protein
bHLH	Basic-Helix-Loop-Helix
BLAST	Basic Local Alignment Search Tool
BUSCO	Benchmarking universal single-copy orthologue
cDNA	Complementary DNA
CO	CONSTANS
FT	Flowering Locus T
GI	GIGANTEA
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
MADS	Mini chromosome maintenance 1
MYB	Myeloblastosis related
NAC	No apical meristem
NAT	*N*-acetyltransferase
NCBI	National Center for Biotechnology Information
NMT	*N*-methyltransferase
OMT	*O*-methyltransferase
PHYB	Phytochrome B
qRT-PCR	Quantitative Real Time Polymerase Chain Reaction
REV	REVOLUTA
SuSy	Sucrose Synthase

FUTURE DIRECTION

In the future, the proposed colchicine pathway and rhizome developmental genes should be validated through qRT-PCR and transgenic knockouts. For the colchicine pathway genes, cloning into different heterologous systems, like bacteria or yeast, could help identify their structure and prove their function. For rhizome developmental and transcription factors, transgenic plants will be used to identify whether the proposed genes are involved in the suspected roles in rhizome development and initiation. Another possibility would be to created transcriptomes of each plant that were tissue specific, which would help identify which genes are specifically involved in each tissue. This could also help evaluate which genes are being upregulated versus down-regulated based on transcriptomic annotation. Our goal is to identify which genes are involved in the biosynthesis of colchicine and which genes are involved in rhizome development. After elucidating the colchicine pathway, we can mass produce colchicine as a pharmaceutical for the eventual treatment of other disease than gout, like cancer and cardiovascular ailments.

REFERENCES

1. Ozmen N, Kaya-Sezginer E, Bakar-Ates F. The cellular uptake and apoptotic efficiency of colchicine is correlated with down-regulation of MMP-9 mRNA expression in SW480 colon cancer cells. *Anticancer Agents Med Chem.* 2018, doi: 10.2174/1871520618666180821102047.
2. Blasco V, Cuñat AC, Sanz-Cervera JF, Marco JA, Falomir E, Murga J, Carda M. Arylureas derived from colchicine: Enhancement of colchicine oncogene downregulation activity. *Eur J Med Chem.* 2018, 150: 817-828.
3. Dybowska M, Szturmowicz M, Kuca P, Kazanecka B, Burakowski J, Czajka C, Grzegorczyk F, Langfort R, Burakowska B, Rudziński P, Tomkowski W. Intrapericardial cisplatin combined with oral colchicine resulted in long term control of malignant pericardial effusion in the course of metastatic renal cancer. *Adv Respir Med.* 2018, doi: 10.5603/ARM.a2018.0030.
4. Dasgeb B, Kornreich D, McGuinn K, Okon L, Brownell I, Sackett DL. Colchicine: An ancient drug with novel applications. *Br J Dermatol.* 2018, 178: 350-356.
5. Roitman A, Ben-Zvi I, Mendel L, Livneh A. Inflammation and cardiovascular disease in familial Mediterranean fever. An analysis of hospital admissions for acute cardiovascular event. *Clin Exp Rheumatol.* 2018, PMID: 30299242.
6. Cimmino G, Tarallo R, Conte S, Morello A, Pellegrino G, Loffredo FS, Calì G, De Luca N, Golino P, Trimarco B, Cirillo P. Colchicine reduces platelet aggregation by modulating cytoskeleton rearrangement via inhibition of cofilin and LIM domain kinase 1. *Vascul Pharmacol.* 2018, pii: S1537-1891(18)30053-3.
7. Vaidya K, Martínez G, Patel S. The role of colchicine in acute coronary syndromes. *Clin Ther.* 2018, pii: S0149-2918(18)30330-8.

8. Schenone AL, Menon V. Colchicine in pericardial disease: From the underlying biology and clinical benefits to the drug-drug interactions in cardiovascular medicine. *Curr Cardiol Rep.* 2018, 20: 62.
9. Chaldakov GN. Colchicine, a microtubule-disassembling drug, in the therapy of cardiovascular diseases. *Cell Biol Int.* 2018, 42: 1079-1084.
10. Arbel Y, Abuzeid W, Rosenson RS, Weisman A, Farkouh ME. Old drugs for new indications in cardiovascular medicine. *Cardiovasc Drugs Ther.* 2018, 32: 223-232.
11. Akodad M, Fauconnier J, Lacampagne A, Roubille F. Reply concerning "Colchicine in coronary artery disease: Role of anti-inflammatory medications redefined": Prime time for anti-inflammatory agents for the management of cardiovascular diseases. *Int J Cardiol.* 2018, 254: 52.
12. Ashish K, Bandyopadhyay D, Mondal S, Ghosh RK. Colchicine in coronary artery disease: Role of anti-inflammatory medications redefined. *Int J Cardiol.* 2018, 254: 51.
13. Liantinioti G, Argyris AA, Protogerou AD, Vlachoyiannopoulos P. The role of colchicine in the treatment of autoinflammatory diseases. *Curr Pharm Des.* 2018, 24: 690-694.
14. Vaidya K, Arnott C, Martínez GJ, Ng B, McCormack S, Sullivan DR, Celermajer DS, Patel S. Colchicine therapy and plaque stabilization in patients with acute coronary syndrome: A CT coronary angiography study. *ACC Cardiovasc Imaging,* 2018, 11: 305-316.
15. Slobodnick A, Krasnokutsky S, Lehmann RA, Keenan RT, Quach J, Francois F, Pillinger MH. Colorectal cancer among gout patients undergoing colonoscopy. *Clin Rheumatol.* 2018, doi: 10.1097/RHU.0000000000000893.
16. Pascart T, Richette P. Colchicine in gout: An update. *Curr Pharm Des.* 2018, 24: 684-689.
17. Slobodnick A, Shah B, Krasnokutsky S, Pillinger MH. Update on colchicine, 2017. *Rheumatology (Oxford)*, 2018, 57: 4-11.

18. Capistrano R, Vangestel C, Wouters A, Dockx Y, Pauwels P, Stroobants S, Apers S, Lardon F, Pieters L, Staelens S. Efficacy screening of *Gloriosa superba* extracts in a murine pancreatic cancer model using (18)F-FDG PET/CT for monitoring treatment response. *Cancer Biother Radiopharm.* 2016, 31: 99-109.
19. Capistrano I R, Vangestel C, Vanpachtenbeke H, Fransen E, Staelens S, Apers S, Pieters L. Coadministration of a *Gloriosa superba* extract improves the *in vivo* antitumoural activity of gemcitabine in a murine pancreatic tumour model. *Phytomedicine*, 2016, 23: 1434-1440.
20. Budchart P, Khamwut A, Sinthuvanich C, Ratanapo S, Poovorawan Y, Thienprasert NP. Partially purified *Gloriosa superba* peptides inhibit colon cancer cell viability by inducing apoptosis through p53 upregulation. *Am J Med Sci.* 2017, 354: 423-429.
21. Johnson L, Goping IS, Rieger A, Mane JY, Huzil T, Banerjee A, Luduena R, Hassani B, Winter P, Tuszynski JA. Novel colchicine derivatives and their anti-cancer activity. *Curr Top Med Chem*. 2017, 17: 2538-2558.
22. Sivakumar G. Colchicine semisynthetics: Chemotherapeutics for cancer? *Curr Med Chem.* 2013, 20: 892-898.
23. Arnst KE, Wang Y, Hwang DJ, Xue Y, Costello T, Hamilton D, Chen Q, Yang J, Park F, Dalton JT, Miller DD, Li W. A potent, metabolically stable tubulin inhibitor targets the colchicine binding site and overcomes Taxane resistance. *Cancer Res.* 2018, 78: 265-277.
24. Worachartcheewan A, Songtawee N, Siriwong S, Prachayasittikul S, Nantasenamat C, Prachayasittikul V. Rational design of colchicine derivatives as anti-HIV agents via QSAR and molecular docking. *Med Chem.* 2018, doi: 10.2174/1573406414666180924163756.
25. Ghawanmeh AA, Chong KF, Sarkar SM, Bakar MA, Othaman R, Khalid RM. Colchicine prodrugs and codrugs: Chemistry and bioactivities. *Eur J Med Chem.* 2018, 144: 229-242.

26. Kesselheim AS, Franklin JM, Kim SC. Reductions in use of colchicine after FDA enforcement of market exclusivity in a commercially insured population. *J Gen Intern Med.* 2015, 30: 1633-1638.
27. Yeh LC, Banerjee A, Prasad V, Tuszynski JA, Weis AL, Bakos T, Yeh IT, Ludueña RF, Lee JC. Effect of CH-35, a novel anti-tumor colchicine analogue, on breast cancer cells overexpressing the βIII isotype of tubulin. *Invest New Drugs,* 2016, 34: 129-137.
28. Johnson L, Goping IS, Rieger A, Mane JY, Huzil T, Banerjee A, Luduena R, Hassani B, Winter P, Tuszynski JA. Novel colchicine derivatives and their anti-cancer activity. *Curr Top Med Chem*. 2017, 17: 2538-2558.
29. Sun Y, Lin X, Chang H. Proliferation inhibition and apoptosis of breast cancer MCF-7 cells under the influence of colchicine. *J Buon.* 2016, 21: 570-575.
30. Kumar A, Singh B, Mahajan G, Sharma PR, Bharate SB, Mintoo MJ, Mondhe DM. A novel colchicine-based microtubule inhibitor exhibits potent antitumor activity by inducing mitochondrial mediated apoptosis in MIA PaCa-2 pancreatic cancer cells. *Tumour Biol.* 2016, 37: 13121-13136.
31. Deftereos SG, Vrachatis DA, Angelidis C, Vrettou AR, Sarri EK, Giotaki SG, Varytimiadi E, Kossyvakis C, Kotsia E, Deftereos GS, Doudoumis K, Giannopoulos G. The role of colchicine in treating postoperative and post-catheter ablation atrial fibrillation. *Clin Ther.* 2018, pii: S0149-2918(18)30371-0.
32. Vrachatis DA, Kossyvakis C, Angelidis C, Panagopoulou V, Sarri EK, Giotaki S, Siasos G, Vavuranakis M, Deftereos SG. Colchicine in post-operative Atrial Fibrillation: A review. *Curr Pharm Des.* 2018, 24: 695-701.
33. Marinaki S, Skalioti C, Boletis JN. Colchicine in renal diseases: Present and future. *Curr Pharm Des.* 2018, 24: 675-683.

34. Solak Y, Siriopol D, Yildiz A, Yilmaz MI, Ortiz A, Covic A, Kanbay M. Colchicine in renal medicine: New virtues of an Ancient friend. *Blood Purif.* 2017, 43: 125-135.

35. Demidowich AP, Davis AI, Dedhia N, Yanovski JA. Colchicine to decrease NLRP3-activated inflammation and improve obesity-related metabolic dysregulation. *Med Hypotheses*. 2016, 92: 67-73.

36. Mubayed L, Muller BA, Jacobson JL, Hast HA, Nguyen HH. Acute pediatric colchicine toxicity is associated with marked bradydysrhythmias. *J Emerg Med.* 2018, S0736-4679(18)30246-4.

37. Alkadi H, Khubeiz MJ, Jbeily R. Colchicine: A review on chemical structure and clinical usage. *Infect Disord Drug Targets*. 2018, 18: 105-121.

38. Chen B, Liu X, Hu YJ, Zhang DM, Deng L, Lu I, Min L, Ye WC, Li CC. Enantioselective total synthesis of (-)-colchicine, (+)-demecolcinone and metacolchicine: determination of the absolute configurations of the latter two alkaloids. *Chem Sci.* 2017, 8: 4961-4966.

39. Chien CW, Teng YG, Honda T, Ojima I. Synthesis of colchicinoids and allocolchicinoids through Rh(I)-catalyzed [2+2+2+1] and [2+2+2] cycloadditions of o-phenylenetriynes with and without CO. *J Org Chem.* 2018, 83: 11623-11644.

40. Gumustas M, Polat DÇ, Kiliç CS, Akalin K, Ozkan SA, Coşkun M. Comparison of seeds of *Colchicum speciosum* and *Gloriosa superba* in respect to colchicine and colchicoside contents by RP-LC. *Nat Prod Commun.* 2016, 11: 397-400.

41. Misra A, Shukla PK, Kumar B, Chand J, Kushwaha P, Khalid M, Singh Rawat AK, Srivastava S. High-performance thin-layer chromatographic-densitometric quantification and recovery of bioactive compounds for identification of elite chemotypes of *Gloriosa superba* L. collected from Sikkim Himalayas (India). *Pharmacogn Mag.* 2017, 13: S700-S705.

42. Arun Kumar P, Elangaimannan R. Evaluation of *Gloriosa superba* for yield attributing characters and quantification of colchicine originated from different agro climatic zones of Tamil Nadu and Andhra Pradesh. *International Journal of Pharmaceutical and Clinical Research*, 2017; 9: 183-192.

43. Pandey DK, Malik T, Dey A, Singh J Banik RM. Improved growth and colchicine concentration in *Gloriosa Superba* on mycorrhizal inoculation supplemented with phosphorus-fertilizer. *Afr J Tradit Complement Altern Med.* 2014, 11: 439-446.

44. Jana S, Shekhawat GS. Critical review on medicinally potent plant species: *Gloriosa superba*. *Fitoterapia*, 2011, 82: 293-301.

45. Zarev Y, Foubert K, Ionkova I, Apers S, Pieters L. Isolation and structure elucidation of glucosylated colchicinoids from the seeds of *Gloriosa superba* by LC-DAD-SPE-NMR. *J Nat Prod.* 2017, 80: 1187-1191.

46. Nguyen PA, Kim JS, Kim JH. The complete chloroplast genome of colchicine plants (*Colchicum autumnale* L. and *Gloriosa superba* L.) and its application for identifying the genus. *Planta*. 2015, 242: 223-237.

47. Sivakumar G. Upstream biomanufacturing of pharmaceutical colchicine. *Crit Rev Biotechnol.* 2018, 38: 83-92.

48. Sivakumar G, Kamran A, Phillips GC. Biorhizome: A biosynthetic platform for colchicine manufacturing. *Front Plant Sci.* 2017, 8: 1137.

49. Nasreen A, Gundlach H, Zenk MH. Incorporation of phenethylisoquinolines into colchicine in isolated seed of *Colchicum autumnale*. *Phytochemistry*, 1997, 46: 107-115.

50. Nasreen A Rueffer M, Zenk MH. Cytochrome P-450 dependent formation of isoandrocymbine from autumnaline in colchicine biosynthesis. *Tetrahedron Lett.* 1996, 37: 8161-8164.

51. Herbert RB, Kattah AE, Knagg E. The biosynthesis of the phenethylisoquinoline alkaloid colchicine. Early and intermediate stage. *Tetrahedron Lett.* 1990, 46: 7119-7138.

52. Rueffer M, Zenk MH. Microsome-mediated transformation of O-methylandrocymbine to demecolcine and colchicine. *FEBS Lett.* 1998, 438: 111-113.

53. Samanani N, Park SU, Facchini PJ. Cell type-specific localization of transcripts encoding nine consecutive enzymes involved in protoberberine alkaloid biosynthesis. *Plant Cell.* 2005, 7: 915-26.

54. Liao D, Wang P, Jia C, Sun P, Qi J, Zhou L, Li X. Identification and developmental expression profiling of putative alkaloid biosynthetic genes in *Corydalis yanhusuo*bulbs. *Sci Rep.* 2016, 6: 19460.

55. Woodhouse RN, McDonald E, Ramage R, Battersby AR. Biosynthesis. Part 29. Colchicine: studies on the ring expansion step focusing on the fate of the hydrogens at C-3 of autumnaline. *J. Chem. Soc. Perkin 1,* 1998, 18: 2995-3001

56. Karamanou M, Tsoucalas G, Pantos K, Androutsos G. Isolating colchicine in 19th century: An old drug revisited. *Curr Pharm Des.* 2018, 24: 654-658.

57. Cheng L, Li S, Xu X, Hussain J, Yin J, Zhang Y, Li L, Chen X. Identification of differentially expressed genes relevant to corm formation in *Sagittaria trifolia*. *PLoS One*, 2013, 8: e54573.

58. Novak SD, Whitehouse GA. Auxin regulates first leaf development and promotes the formation of protocorm trichomes and rhizome-like structures in developing seedlings of *Spathoglottis plicata* (Orchidaceae). *AoB Plants*, 2013, 5: pls053.

59. Masuda J, Ozaki Y, Okubo H. Rhizome transition to storage organ is under phytochrome control in lotus (*Nelumbo nucifera*). *Planta*, 2007, 226: 909-915.

60. Wang K, Peng H, Lin E, Jin Q, Hua X, Yao S, Bian H, Han N, Pan J, Wang J, Deng M, Zhu M. Identification of genes related to the development of bamboo rhizome bud. *J Exp Bot.* 2010, 61: 551-561.

61. Reddy AS, Day IS, Narasimhulu SB, Safadi F, Reddy VS, Golovkin M, Harnly MJ. Isolation and characterization of a novel calmodulin-binding protein from potato. *J Biol Chem.* 2002, 277: 4206-4214.

62. Yang M, Zhu L, Pan C, Xu L, Liu Y, Ke W, Yang P. Transcriptomic analysis of the regulation of rhizome formation in temperate and Tropical Lotus (*Nelumbo nucifera*). *Sci Rep.* 2015, 5: 13059.

63. Navarro C, Abelenda JA, Cruz-Oró E, Cuéllar CA, Tamaki S, Silva J, Shimamoto K, Prat S. Control of flowering and storage organ formation in potato by Flowering Locus T. *Nature*, 2011, 478: 119-122.

64. Cheng L, Li S, Yin J, Li L, Chen X. Genome-wide analysis of differentially expressed genes relevant to rhizome formation in Lotus root (*Nelumbo nucifera* Gaertn). *PLoS One*, 2013, 8: e67116.

65. Kim MJ, Nelson W, Soderlund C, Gang DR. Next-generation sequencing-based transcriptional profiling of Sacred Lotus "China Antique". *Trop Plant Biol.* 2013, 6: 161-179.

66. Yang M, Zhu L, Pan C, Xu L, Liu Y, Ke W, Yang P. Transcriptomic analysis of the regulation of rhizome formation in temperate and tropical lotus (*Nelumbo nucifera*). *Sci Rep.* 2015, 5: 13059.

67. Yang CQ, Fang X, Wu XM, Mao YB, Wang LJ, Chen XY. Transcriptional regulation of plant secondary metabolism. *J Integr Plant Biol.* 2012, 54: 703-712.

68. Koo HJ, Gang DR. Suites of terpene synthases explain differential terpenoid production in ginger and turmeric tissues. *PLoS One*, 2012, 7: e51481.

69. Koo HJ, McDowell ET, Ma X, Greer KA, Kapteyn J, Xie Z, Descour A, Kim H, Yu Y, Kudrna D, Wing RA, Soderlund CA, Gang DR. Ginger and turmeric expressed sequence tags identify signatures for rhizome identity and development and biosynthesis of curcuminoids, gingerols and terpenoid. *BMC Plant Biol.* 2013, 13: 27.

70. Chen H, Deng C, Nie H, Fan G, He Y. Transcriptome analyses provide insights into the difference of alkaloids biosynthesis in the Chinese goldthread (*Coptis chinensis* Franch.) from different biotopes. *PeerJ.* 2017, 5: e3303.

71. Samanani N, Park SU, Facchini PJ. Cell type-specific localization of transcripts encoding nine consecutive enzymes involved in protoberberine alkaloid biosynthesis. *Plant Cell.* 2005, 7: 915-26.

72. Liao D, Wang P, Jia C, Sun P, Qi J, Zhou L, Li X. Identification and developmental expression profiling of putative alkaloid biosynthetic genes in *Corydalis yanhusuo*bulbs. *Sci Rep.* 2016, 6: 19460.

73. Zonga Y, Huanga L, Zhanga T, Qina Q, Wanga W, Zhaoa X, Hub F, Fua B, Lia Z. Differential microRNA expression between shoots and rhizomes in *Oryza longistaminata* using high-throughput RNA sequencing. *Crop J.* 2014, 2: 102-109.

74. Kang K, Lee K, Park S, Byeon Y, Back K. Molecular cloning of rice serotonin *N*-acetyltransferase, the penultimate gene in plant melatonin biosynthesis. *J Pineal Res.* 2013, 55: 7-13.

75. Byeon Y, Lee HY, Back K. Cloning and characterization of the serotonin *N-acetyltransferase*-2 gene (SNAT2) in rice (*Oryza sativa*). *J Pineal Res.* 2016, 61: 198-207.

76. Choi KB, Morihige T, Shitan N, Yazaki K, Sato F. Molecular cloning and characterization of coclaurine *N*-methyltransferase from cultured cells of *Coptis japonica*. *J Biol Chem.* 2002, 277: 830-835.

77. Torres MA, Hoffarth E, Eugenio L, Savtchouk J, Chen X, Morris JS, Facchini PJ, Ng KK. Structural and functional studies of pavine N-Methyltransferase from *Thalictrum flavum* reveal insights into substrate recognition and catalytic mechanism. *J Biol Chem.* 2016, 291: 23403-23415

78. Morris JS, Facchini PJ. Isolation and characterization of reticuline N-Methyltransferase involved in biosynthesis of the aporphine alkaloid magnoflorine in opium poppy. *J Biol Chem.* 2016, 291: 23416-23427.

79. Maier UH, Zenk MH. Colchicine is formed by *para-para* phenol coupling from autumnaline. *Tetrahedron Lett.* 1997, 38: 7357-7360.

80. Kilgore MB, Augustin MM, May GD, Crow JA, Kutchan TM. CYP96T1 of Narcissus sp. aff. pseudonarcissus catalyzes formation of the para-para C-C phenol couple in the Amaryllidaceae alkaloids. *Front Plant Sci.* 2016, 7: 225.

81. Gesell A, Rolf M, Ziegler J, Dı´az Cha´vez ML, Huang FC, Kutchan TM. CYP719B1 is salutaridine synthase, the C-C phenol-coupling enzyme of morphine biosynthesis in opium poppy. *J. Biol. Chem.* 2009, 284: 24432-24442

82. Gesell A, Rolf M, Ziegler J, Dı´az Cha´vez ML, Huang FC, Kutchan TM. CYP719B1 is salutaridine synthase, the C-C phenol-coupling enzyme of morphine biosynthesis in opium poppy. *J Biol Chem.* 2009, 284: 24432-24442.

83. Ikezawa N, Iwasa K, Sato F. Molecular cloning and characterization of CYP80G2, a cytochrome P450 that catalyzes an intramolecular C-C phenol coupling of (S)-reticuline in magnoflorine biosynthesis, from cultured *Coptis japonica* cells. *J Biol Chem.* 2008, 283: 8810-8821.

84. Kilgore MB, Augustin MM, Starks CM, O'Neil-Johnson M, May GD, Crow JA, Kutchan TM. Cloning and characterization of a norbelladine 4'-*O*-methyltransferase involved in the biosynthesis

of the Alzheimer's drug galanthamine in Narcissus sp. aff. pseudonarcissus. *PLoS One*, 2014, 9: e103223.

85. Park MR, Chen X, Lang DE, Ng KKS, Facchini PJ. Heterodimeric O-methyltransferases involved in the biosynthesis of noscapine in opium poppy. *Plant J.* 2018, 95 :252-267.
86. Li Y, Li S, Thodey K, Trenchard I, Cravens A, Smolke CD. Complete biosynthesis of noscapine and halogenated alkaloids in yeast. *Proc Natl Acad Sci USA*. 2018, 115: E3922-E3931.
87. Cai B, Siminszky B, Chappell J, Dewey RE, Bush LB. Enantioselective demethylation of nicotine as a mechanism for variable nornicotine composition in tobacco leaf. *J Biol Chem.* 2012, 287: 42804-42811.
88. Lewis RS, Bowen SW, Keogh MR, Dewey RE. Three nicotine demethylase genes mediate nornicotine biosynthesis in *Nicotiana tabacum* L.: functional characterization of the CYP82E10 gene. *Phytochemistry,* 2010, 71: 1988-1998.
89. Liedschulte V, Schwaar JD, Laparra H, Vuarnoz A, Philippon B, Bakaher N, Sierro N, Bovet L, Lang G, Goepfert S. Identification of CYP82E21 as a functional nicotine N-demethylasein tobacco flowers. *Phytochemistry,* 2016, 131: 9-16.
90. Kellner F, Kim J, Clavijo BJ, Hamilton JP, Childs KL, Vaillancourt B, Cepela J, Habermann M, Steuernage B, Clissold L, McLay K, Buell CR, O'Connor SE. Genome-guided investigation of plant natural product biosynthesis. *Plant J*. 2015, 82: 680-692.
91. Goossens A, Häkkinen ST, Laakso I, Seppänen-Laakso T, Biondi S, De Sutter V, Lammertyn F, Nuutila AM, Söderlund H, Zabeau M, Inzé D, Oksman-Caldentey KM. A functional genomics approach toward the understanding of secondary metabolism in plant cells. *PANS*, 2003, 100: 8595-600.

92. Kautsar SA, Suarez Duran HG, Medema MH. Genomic identification and analysis of specialized metabolite biosynthetic gene clusters in plants using PlantiSMASH. *Methods Mol Biol.* 2018, 1795:173-188.

93. Kautsar SA, Suarez Duran HG, Blin K, Osbourn A, Medema MH. PlantiSMASH: automated identification, annotation and expression analysis of plant biosynthetic gene clusters. *Nucleic Acids Res.* 2017, 45: W55-W63.

94. Ding J, Böhlenius H, Rühl MG, Chen P, Sane S, Zambrano JA, Zheng B, Eriksson ME, Nilsson O. GIGANTEA-like genes control seasonal growth cessation in Populus. *New Phytol.* 2018, 218: 1491-1503.

95. de Montaigu A, Coupland G. The timing of GIGANTEA expression during day/night cycles varies with the geographical origin of Arabidopsis accessions. *Plant Signal Behav.* 2017, 12: e1342026.

96. Tang W, Yan H, Su ZX, Park SC, Liu YJ, Zhang YG, Wang X, Kou M, Ma DF, Kwak SS, Li Q. Cloning and characterization of a novel GIGANTEA gene in sweet potato. *Plant Physiol Biochem.* 2017, 116: 27-35.

97. Hu T, Wei Q, Wang W, Hu H, Mao W, Zhu Q, Bao C. Genome-wide identification and characterization of CONSTANS-like gene family in radish (*Raphanus sativus*). *PLoS One,* 2018, 13: e0204137.

98. Abelenda JA, Cruz-Oró E, Franco-Zorrilla JM, Prat S. Potato StCONSTANS-like1 suppresses storage organ formation by directly activating the FT-like StSP5G repressor. *Curr Biol.* 2016, 26: 872-881.

99. Teo CJ, Takahashi K, Shimizu K, Shimamoto K, Taoka KI. Potato tuber induction is regulated by interactions between components of a tuberigen complex. *Plant Cell Physiol.* 2017, 58: 365-374.

100. Martin A, Adam H, Díaz-Mendoza M, Zurczak M, González-Schain ND, Suárez-López P. Graft-transmissible induction of potato tuberization by the microRNA miR172. *Development,* 2009, 136: 2873-2881.

101. Hastilestari BR, Lorenz J, Reid S, Hofmann J, Pscheidt D, Sonnewald U, Sonnewald S. Deciphering source and sink responses of potato plants (*Solanum tuberosum* L.) to elevated temperatures. *Plant Cell Environ.* 2018, 41): 2600-2616.

102. Akakpo R, Scarcelli N, Chaïr H, Dansi A, Djedatin G, Thuillet AC, Rhoné B, François O, Alix K, Vigouroux Y. Molecular basis of African yam domestication: analyses of selection point to root development, starch biosynthesis, and photosynthesis related genes. *BMC Genomics*, 2017, 18: 782.

103. Merelo P, Paredes EB, Heisler MG, Wenkel S. The shady side of leaf development: the role of the REVOLUTA/KANADI1 module in leaf patterning and auxin-mediated growth promotion. *Curr Opin Plant Biol.* 2017, 35: 111-116.

104. Peng HZ, Lin EP, Sang QL, Yao S, Jin QY, Hua XQ, Zhu MY. Molecular cloning, expression analyses and primary evolution studies of REV- and TB1-like genes in bamboo. *Tree Physiol.* 2007, 27: 1273-1281.

105. Gao H, Wang Z, Li S, Hou M, Zhou Y, Zhao Y, Li G, Zhao H, Ma H. Genome-wide survey of potato MADS-box genes reveals that StMADS1 and StMADS13 are putative downstream targets of tuberigen StSP6A. *BMC Genomics*, 2018, 19: 726.

106. Huang Q, Huang X, Deng J, Liu H, Liu Y, Yu K, Huang B. Differential gene expression between leaf and rhizome in *Atractylodes lancea*: A comparative transcriptome analysis. *Front Plant Sci.* 2016, 7: 348.

107. Su X, Lu G, Guo H, Zhang K, Li X, Cheng H. The dynamic transcriptome and metabolomics profiling in *Verticillium dahliae* inoculated *Arabidopsis thaliana*. *Sci Rep.* 2018, 8: 15404.

108. Bakshi M, Oelmüller R. WRKY transcription factors. *Plant Signal Behav.* 2014, 9; e27700.

109. Kumari A, Singh HR, Jha A, Swarnkar MK, Shankar R, Kumar S. Transcriptome sequencing of rhizome tissue of S*inopodophyllum hexandrum* at two temperatures. *BMC Genomics*, 15: 871.

110. Zhang T, Huang L, Wang Y, Wang W, Zhao X, Zhang S, Zhang J, Hu F, Fu B, Li Z. Differential transcriptome profiling of chilling stress response between shoots and rhizomes of *Oryza longistaminata u*sing RNA sequencing. *PLoS One.* 2017, 12: e0188625.

111. Charfeddine M, Saïdi MN, Charfeddine S, Hammami A, Gargouri Bouzid R. Genome-wide analysis and expression profiling of the ERFtranscription factor family in potato (*Solanum tuberosum* L.). *Mol Biotechnol.* 2015, 57: 348-358.

112. Wang R, Zhao P, Kong N, Lu R, Pei Y, Huang C, Ma H, Chen Q. Genome-wide identification and characterization of the potato bHLH transcription factor family. *Genes (Basel)*, 2018, 9: pii: E54.

113. Kodama M, Brinch-Pedersen H, Sharma S, Holme IB, Joernsgaard B, Dzhanfezova T, Amby DB, Vieira FG, Liu S, Gilbert MTP. Identification of transcription factor genes involved in anthocyanin biosynthesis in carrot (*Daucus carota* L.) using RNA-Seq. *BMC Genomics*, 2018, 19: 811.

114. Liu F, Yang Y, Gao J, Ma C, Bi Y. A comparative transcriptome analysis of a wild purple potato and its red mutant provides insight into the mechanism of anthocyanin transformation. *PLoS One*. 2018, 13: e0191406

115. Kumar P, Jaiswal V, Pal T, Singh J, Chauhan RS. Comparative whole-transcriptome analysis in Podophyllum speciesidentifies key transcription factors contributing to biosynthesis of podophyllotoxin in *P. hexandrum*. *Protoplasma*, 2017, 254: 217-228.

116. Xie Q, Guo HS, Dallman G, Fang S, Weissman AM, Chua NH. SINAT5 promotes ubiquitin-related degradation of NAC1 to attenuate auxin signals. *Nature*, 2002, 419: 167-170.

117. Zhang Y, Li D, Wang Y, Zhou R, Wang L, Zhang Y, Yu J, Gong H, You J, Zhang X. Genome-wide identification and comprehensive analysis of the NAC transcription factor family in *Sesamum indicum*. *PLoS One,* 2018, 13: e0199262.

118. Phukan UJ, Jeena GS, Shukla RK. WRKY transcription factors: Molecular regulation and stress responses in plants. *Front Plant Sci.* 2016, 7: 760.

119. Li LQ, Zou X, Deng MS, Peng J, Huang XL, Lu X, Fang CC, Wang XY. Comparative morphology, transcription, and proteomics study revealing the key molecular mechanism of camphor on the potato tuber sprouting effect. *Int J Mol Sci.* 2017, 18: 2280.

120. Teoh KH, Sivakumar G. An expressed sequence tag database analysis of fatty acid genes in *Stichococcus bacillaris* strain siva2011. In: *Biotechnology Advances*. (Ed. Kim SK), Academic Press, Elsevier, 2015, pp 429-438.

121. Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol*. 2017, 35: 543-548.

122. Götz S, García-Gómez J, Terol J, Williams T, Nagaraj S, Nueda M, Robles M, Talón M, Dopazo J, Conesa A. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 2008, 36: 3420–3435.