OPEN ACCESS Freely available online

PLOS | ONE

Transcriptome Analysis of Buds and Leaves Using 454 Pyrosequencing to Discover Genes Associated with the Biosynthesis of Active Ingredients in *Lonicera japonica* Thunb.

Liu He[1]ᵠ, Xiaolan Xu[1]ᵠ, Ying Li[1], Chunfang Li[1], Yingjie Zhu[1], Haixia Yan[1], Zhiying Sun[1], Chao Sun[1], Jingyuan Song[1], Yu'an Bi[2], Juan Shen[2], Ruiyang Cheng[3], Zhenzhong Wang[2], Wei Xiao[2]*, Shilin Chen[1,4]*

1 Institute of Medicinal Plant Development (IMPLAD), Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, **2** State Key Laboratory of New-tech for Chinese Medicine Pharmaceutical Process; Jiangsu Kanion Pharmaceutical Co. LTD, Lianyungang, China, **3** Beijing University of Chinese Medicine, Beijing, China, **4** Institute of Chinese Materia Medica, Academy of Chinese Medical Sciences, Beijing, China

Abstract

Background: *Lonicera japonica* Thunb. is a plant used in traditional Chinese medicine known for its anti-inflammatory, anti-oxidative, anti-carcinogenic, and antiviral pharmacological properties. The major active secondary metabolites of this plant are chlorogenic acid (CGA) and luteoloside. While the biosynthetic pathways of these metabolites are relatively well known, the genetic information available for this species, especially the biosynthetic pathways of its active ingredients, is limited.

Methodology/Principal Findings: We obtained one million reads (average length of 400 bp) in a whole sequence run using a Roche/454 GS FLX titanium platform. Altogether, 85.69% of the unigenes covering the entire life cycle of the plant were annotated and 325 unigenes were assigned to secondary metabolic pathways. Moreover, 2039 unigenes were predicted as transcription factors. Nearly all of the possible enzymes involved in the biosynthesis of CGA and luteoloside were discovered in *L. japonica*. Three hydroxycinnamoyl transferase genes, including two hydroxycinnamoyl-CoA quinate hydroxycinnamoyl transferase genes and one hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase (HCT) gene featuring high similarity to known genes from other species, were cloned. The *HCT* gene was discovered for the first time in *L. japonica*. In addition, 188 candidate cytochrome P450 unigenes and 245 glycosyltransferase unigenes were found in the expressed sequence tag (EST) dataset.

Conclusion: This study provides a high quality EST database for *L. japonica* by 454 pyrosequencing. Based on the EST annotation, a set of putative genes involved in CGA and luteoloside biosynthetic pathways were discovered. The database serves as an important source of public information on genetic markers, gene expression, genomics, and functional genomics in *L. japonica*.

Citation: He L, Xu X, Li Y, Li C, Zhu Y, et al. (2013) Transcriptome Analysis of Buds and Leaves Using 454 Pyrosequencing to Discover Genes Associated with the Biosynthesis of Active Ingredients in *Lonicera japonica* Thunb. PLoS ONE 8(4): e62922. doi:10.1371/journal.pone.0062922

Editor: Turgay Unver, Cankiri Karatekin University, Turkey

Received October 5, 2012; **Accepted** March 29, 2013; **Published** April 25, 2013

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

Funding: This work was supported by the Key National Natural Science Foundation of China (Grant no. 81130069), the National Basic Research Program of China (Grant no.2010CB735604), and the Program for Changjiang Scholars and Innovative Research Team in University of Ministry of Education of China (Grant no. IRT1150). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Four authors (YB, J. Shen, ZW and WX) are employed in Jiangsu Kanion Pharmaceutical Co. LTD. All the authors are cooperation partners in this study about transcriptome analysis of L. japonica. There are no competing interests in employment, consultancy, patents, products in development or marketed products etc. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: wzhzh-nj@tom.com (WX); slchen@implad.ac.cn (SC)

ᵠ These authors contributed equally to this work.

Introduction

Lonicera japonica Thunb. is a perennial, evergreen, twining vine. It has double-tongued flowers that open white and fade to yellow. This plant is called Jinyinhua (literally "gold silver flower") in Chinese and is cultivated worldwide as an ornamental plant because of its numerous sweet smelling flowers. However, *L. japonica* is believed to be invasive to the ecology of some American countries because of its strong viability [1]. It is widely cultivated in China as well as in other Asian countries, such as Japan and Korea, as a commercially valuable plant. Flos Lonicerae Japonicae (FLJ), the dried bud of the *L. japonica*, has been used for thousands of years in Chinese medicine for its antipyretic, antidotal, and anti-inflammatory properties. It has been recorded in the "Ben Cao Gang Mu" (Compendium of Materia Medica) as early as the 17th century. Thus far, FLJ is a popular drug for the treatment of influenza virus. Since the outbreak of SARS and avian influenza viruses in China, the use of *L. japonica* has significantly increased.

Several studies have investigated *L. japonica* to improve its applications and a large number of active ingredients have been extracted from the plant, including phenolic acids [2–4], flavones [5–8], triterpenoid saponins [9–13] and volatile oils [14,15]. These

ingredients mediate multiple properties, including antioxidant [16–18], anti-inflammatory [19], anti-carcinogenic [20–26], and antiviral [27,28] effects. Among these ingredients, chlorogenic acid (CGA) and luteoloside are the primary active components that have attracted the most attention from researchers. The contents of these two components in *L. japonica* are the main valuation criteria of this plant.

CGA was first found in sunflower seeds [29] and is an important phenolic acid that comes from secondary metabolism pathways in many plants [30,31]. CGA is often used in medicines and foods because of its high anti-oxidative activity [32–34].The biosynthetic pathway of CGA is controversial, although three metabolic pathways have been postulated based on previous research. The first mechanism indicates that CGA is produced from quinic acid and caffeoyl CoA and catalyzed by the hydroxycinnamoyl-CoA quinate hydroxycinnamoyl transferase (HQT) [35–39]. The second mechanism indicates that CGA comes from quinic acid and caffeoyl-D-glucose and is catalyzed by hydroxycinnamoyl D-glucose: quinate hydroxycinnamoyl transferase (HCGQT) [40–42]. The third mechanism proposes that CGA comes from p-coumaroyl quinic acid and is catalyzed by hydroxycinnamoyl CoA shikimate/quinate hydroxycinnamoyl transferase (HCT) [43,44]. Recent research has demonstrated that HQT is an indispensable enzyme in the synthesis of CGA in *L. japonica* [38].

Luteoloside belongs to a group of natural flavonoids and exists in many plants. It comes from luteolin by 7-O-beta-glucosyltransferase, which transfers the glucosyl moiety to the 7-O-position of the substrate [45–50]. The biosynthetic pathway of luteolin is clear in other species [51] but all of the genes involved in the biosynthesis of luteolin in *L. japonica* have not yet been found. *L. japonica* is a very important plant material in the study of CGA and luteoloside because of their high content in this species.

Limited genomic and transcriptomic information is available for *L. japonica* in GenBank (September 2012). The 100 million sequence reads obtained from *L. japonica* and *L. japonica* Thunb. var. chinensis (Watts) using Illumina Genome Analyzer II and the approximately 6000 expressed gene tags for each of the three flower development stages from FLJ are not available in NCBI [52]. In addition, most enzymes involved in the biosynthetic pathways of active compounds have not been reported.

Besides Illumina GA II, the Roche/454 GS FLX platform, another high-throughput sequencing platform, is also frequently used to sequence the transcriptomes of medicinal plants. This platform aims to discover genes and analyze EST-SSRs [53–55]. In this study, we sequenced the transcriptome of *L. japonica* buds and leaves using the Roche 454 GS FLX Titanium platform. Our purpose is to discover all candidate genes encoding enzymes and putative transcription factors (TFs) in the CGA and luteoloside biosynthetic pathways to allow for the future synthesis of CGA and luteoloside by heterologous expression in other cell lines.

Results

454 cDNA Sequencing and EST Assembly

In order to find more genes involved in the biosynthetic pathway of active components, we used buds and leaves for transcriptome sequencing and analysis. Buds are the primary medicinal parts of *L. japonica* [56] and leaves are developing medical resources [57]. Two cDNA libraries were constructed from the total RNA of fresh *L. japonica* buds and leaves using SMART technology. The libraries were subjected to a sequencing run on the 454 GS FLX Titanium platform, each library for half a run. The length distribution of all reads (≥50 bp) of the two libraries can evaluate the quality of the sequence (Figure 1A). After initial quality filtering with the default parameters, the entire run for *L. japonica* buds and leaves produced over one million high-quality (HQ) reads with a total length of 450.9 Mb. The 454-derived ESTs in this study were independently assembled using GS *De Novo* v.2.6 assembler software. After the entire run was assembled, the quality of the assembly (e.g., the ratio of aligned reads, average Contig length, N50 Contig size, largest Contig size) was found to be better than the individual results (Table 1). Overall, 92% of the HQ reads exceeded our minimal quality standards (e.g., SMART primer filtering; length threshold of 50 bp) and were thus used in the assembly. The length distribution of Contigs is shown in Figure 1B. The assembly of the two libraries was used in the following analysis to completely explore their genetic information.

Functional Annotation and Categorization

To produce the most informative and complete annotation, all unique sequences were annotated by BLAST [58] against a series of nucleotide and protein databases, including NCBI nucleotide (Nt, released on 03/2012), NCBI non-redundant protein (Nr, released on 03/2012), UniProtKB/SwissProt (released on 03/2012), Kyoto Encyclopedia of Genes and Genomes (KEGG 58) [59,60], and the *Arabidopsis thaliana* proteome databases (TAIR10) [61]. A total of 64,184 unigenes were compared against these databases with a significance threshold of E-value ≤1e-5. The number of unigenes only can be found in each organ was 32,907 for the buds and 14,566 for the leaves (Figure S1A). Of the 64,184 unigenes, 51,500 (80.24%) unigenes had at least one hit within these databases (Table 2). The remaining unigenes (19.76%) that were not annotated are likely to comprise *L. japonica*-specific genes, as well as genes with homologs in other species, whose corresponding biological functions have not yet been investigated.

The *L. japonica* unigenes were further analyzed and categorized using Gene Ontology (GO) based on the InterPro scan results. GO is an international classification system for standardized gene

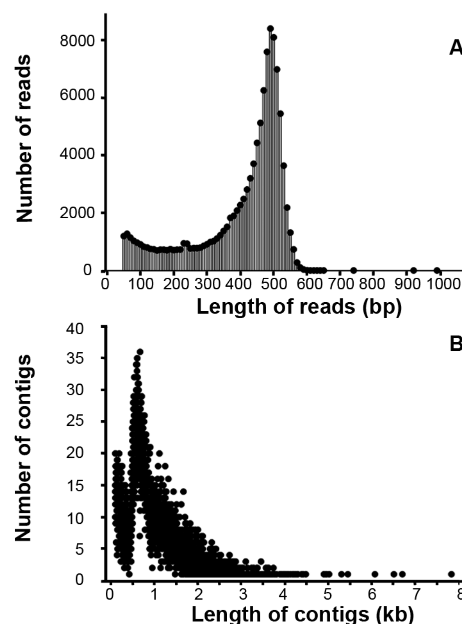

Figure 1. Distribution of the 454 HQ read and contig lengths from *L. japonica*. A. Size distribution of 454 sequencing HQ reads. B. Length distribution of contigs in the EST datasets.
doi:10.1371/journal.pone.0062922.g001

Table 1. Summary of *L. japonica* leaves and buds EST sequencing derived from 454 GS FLX Titanium.

	Buds	Leaves	All
Total bases	235250941	215654256	450905197
Total reads	587840	535470	1123310
Average reads length	404	407	406
Aligned bases	215517769	204915360	420433129
Aligned reads	531719	499708	1031427
Average Contig length	1007	879	1127
N50 Contig Size	1130	933	1282
Largest Contig size	5302	5539	7830
Number of Contigs	21863	14382	20940

doi:10.1371/journal.pone.0062922.t001

functions [62]. A total of 19,785 unigenes were classified into three main independent GO categories: molecular function, biological process and cellular component, including 34 subcategories (Figure 2A). Given that a transcript may have multiple hits, some unigenes were classified into two or three categories. The category of biological process, which included 29,623 unigenes, was the largest, followed by the function category (25,306 unigenes) and the cellular location category (16,342 unigenes). Of these categories, protein binding (44.8%) and metabolic process (35.3%) were the two largest subcategories. The percentage of each subcategory in the two organs was shown in Figure S1B. The unigenes only found in the buds or leaves may be related to the development in the two organs. GO analysis showed that the unigenes identified by our sequence run function in various biological processes.

Cluster of orthologous groups (COG) [63] classification is based on comparing protein sequences encoded in complete genomes and presenting major phylogenetic lineages. Only 17.8% (11,414) of the unigenes have been annotated. A total of 9816 of these unigenes clustered into 24 functional categories based on the COG phylogenetic classification (Figure 2B). Among these categories, the "general function prediction only" group was the largest. The "secondary metabolites biosynthesis" group comprised 232 (2.4%) unigenes (Figure 2B). This category includes important factors in the biosynthesis of active secondary metabolites in *L. japonica*.

To conduct a more detailed analysis of the unigenes involved in biosynthetic pathways, the unigenes were assigned to the KEGG metabolic pathway category [64]. We mapped 43,129 unigenes to 249 pathways using the KEGG database. Of these unigenes, 9,594 were involved in metabolic pathways and 325 were related to secondary metabolism and mapped to 11 pathways (Figure 3). The phenylpropanoid biosynthesis pathway (ID: ko00940, 39.4%) was the largest group, containing 128 unigenes. Approximately 14.4% of the unigenes were found in the "flavonoid biosynthesis" (ID: ko00941, 13.2%) and "flavone and flavonol biosynthesis" pathways (ID: ko0094). These three categories feature the biosynthesis of the main active compounds, including CGA and luteoloside, in *L. japonica*.

Discovery of Unigenes Encoding Putative TFs

L. japonica has strong adaptability and is widely cultivated in the rock-desertification environment of southwestern China. Some progress has been made in understanding the physiological characteristics of their tolerance to low temperature and drought stress [65–67]. However, the molecular mechanism of this resistance remains unclear. Secondary metabolites can protect plants against microbial, herbivore, and ultraviolet attack. The regulation of TFs involved in secondary metabolism in plants is a crucial area of research for exploring plant defense responses in stresses [68,69]. Various TFs participate in different defense signaling pathways. Thus, the identification of putative TF genes is helpful for realizing the molecular mechanism of *L. japonica* response to environment changes.

In this study, annotation on TAIR was performed to search against the AGRIS (Arabidopsis Gene Regulatory Information Server) [70]. A total of 2,039 unigenes were annotated to TFs in *L. japonica*. These unigenes were annotated in 788 independent coding sequences of *Arabidopsis* TFs that belong to 50 known TF families (Table 3). The zinc finger C2H2 TF family is the most abundant within the *L. japonica* ESTs, including 276 unigenes, followed by the bHLH, C3H, Homeobox, MYB, and AP2/EREBP TF families, which include, 211, 202, 145, 100, and 90 unigenes, respectively, in our data set.

The C2H2, MYB, and AP2/EREBP TF families are important for regulating the response to abiotic stress in plants [71–73]. In addition, members of MYB and AP2/EREBP have been found in two plant metabolic pathways, leading to the biosynthesis of phenylpropanoids and terpenoid indole alkaloids, respectively [74–76]. The high expression level of TFs in *L. japonica* may be linked to their defense responses and biosynthesis of secondary metabolites. The discovery of these putative TFs in this study may provide useful information for future research.

Putative Unigenes Encoding Enzymes Involved in the Biosynthesis of CGA and Luteoloside

CGA and luteoloside are both derived from phenylalanine. We screened the EST dataset and, based on the annotations results,

Table 2. Summary of *L. japonica* leaves and buds EST annotation.

	Total Unigene	Annotated	Percent (%)	Unannotated	Percent (%)
Nt	64184	42595	66.4	21589	33.6
Nr	64184	47121	73.4	17063	26.6
Swissprot	64184	27671	43.1	36513	56.9
KEGG	64184	43129	67.2	21055	32.8
TAIR	64184	45344	70.6	18840	29.4

doi:10.1371/journal.pone.0062922.t002

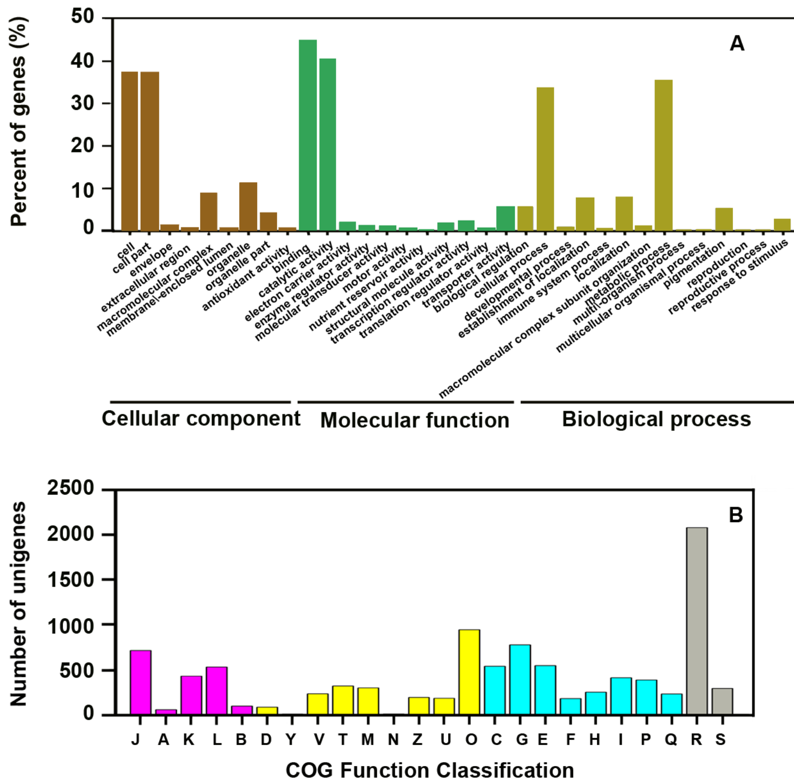

Figure 2. GO and COG classifications of unigenes from *L. japonica*. A. GO function classification of transcriptome. B. COG function classification of transcriptome. Purple boxes represent information storage and processing. J: Translation, ribosomal structure, and biogenesis. A: RNA processing and modification. K: Transcription. L: Replication, recombination and repair. B: Chromatin structure and dynamics. Yellow boxes represent cellular processes and signaling. D: Cell cycle control, cell division, and chromosome partitioning. Y: Nuclear structure. V: Defense mechanisms. T: Signal transduction mechanisms. M: Cell wall/membrane/envelope biogenesis. N: Cell motility. Z: Cytoskeleton. W: Extracellular structures. U: Intracellular trafficking, secretion, and vesicular transport. O: Posttranslational modification, protein turnover, and chaperones. Blue boxes represent metabolism. C: Energy production and conversion. G: Carbohydrate transport and metabolism. E: Amino acid transport and metabolism. F: Nucleotide transport and metabolism. H: Coenzyme transport and metabolism. I: Lipid transport and metabolism. P: Inorganic ion transport and metabolism. Q: Secondary metabolite biosynthesis, transport, and catabolism. Gray boxes represent poorly characterized. R: General function prediction only. S: Function unknown.
doi:10.1371/journal.pone.0062922.g002

discovered almost all of the enzymes involved in these biosynthetic pathways. Figure 4 shows the number of corresponding unigenes with each enzyme.

The three possible routes of CGA biosynthesis are shown in the order in which they were observed. A total of 61 unigenes were found for CGA biosynthesis (Figure 4, Table S3). Two enzymes that are specific to the second route are not in this EST database. Unigenes encoding *C4H*, *C3′H*, and *HCT* were found in *L. japonica* for the first time. Three unigenes matched *HQT*, one of which is partially *LjHQT* (GQ847546) [38]. HCT and HQT are both hydroxycinnamoyl transferases and their amino acid sequences have high similarity. Twenty unigenes encoding HCT and HQT were all subjected to the following analysis to facilitate the cloning and characterization of *HCT* and *HQT* genes.

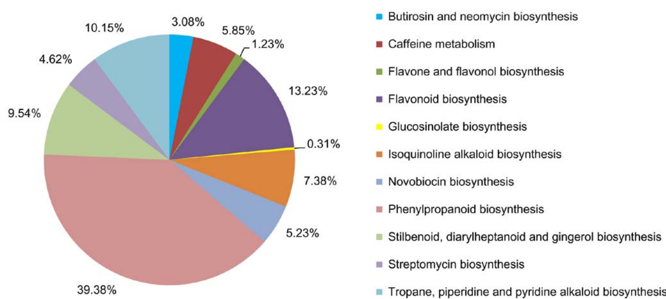

Figure 3. Percentages of *L. japonica* unigenes in 11 subcategories of the metabolic pathway category.
doi:10.1371/journal.pone.0062922.g003

Table 3. Top 10 putative transcription factor families in *L. japonica*.

TF family	No. of unigenes	Percent (%)
C2H2	276	13.5
bHLH	211	10.3
C3H	202	9.9
Homeobox	145	7.1
MYB	100	4.9
AP2-EREBP	90	4.4
GRAS	81	4.0
MYB-related	81	4.0
Trihelix	77	3.8
NAC	72	3.5

doi:10.1371/journal.pone.0062922.t003

Luteolin is normally combined with glucose to form luteoloside in plants. Luteoloside is a type of flavonoid that has an important function in the interaction between plants and their environment. All enzymes involved in luteoloside biosynthesis were found in this study (Table S4). One unigene encoding FSII and two unigenes encoding F3′H were identified in *L. japonica* for the first time. FSII and F3′H belong to the CYP93B and CYP75B families, respectively. Cytochrome P450 (CYP) is a large superfamily that has important functions in the oxidation and hydroxylation reactions in plants. In this study, a total of 188 unique unigenes were annotated as putative CYPs and further classified into 22 families and 30 subfamilies (Table S1). The last step in the biosynthesis of luteoloside is catalyzed by 7-O-beta-glucosyltransferase. Glycosylation can stabilize luteolin in the cells. Glycosyltransferase belongs to another large multigene family in plants. This study found 245 unigenes encoding glycosyltransferase in *L. japonica*. These unigenes are all regarded as important enzymes in many secondary metabolism pathways.

HQT

We further analyzed the 17 putative *HCT* unigenes and 3 *HQT* unigenes according to protein domains. In this manner, 10 putative *HQT/HCT* genes were identified using bioinformatics analysis. 5′ RACE-PCR and 3′ RACE-PCR were performed according to the unigene sequences in the EST dataset. All of the full-length cDNA sequences of the 10 genes were acquired (Table S2). Among these, singleton HDUSP and BQY55 are two partial sequences of the same gene; contig07826 and singleton H2O5B are two partial sequences of the same gene. We finally obtained 8 ORF sequences and named them: *Lj_07643*, *Lj_07826*, *Lj_08086*, *Lj_08422*, *Lj_09545*, *Lj_12999*, *GIA04,* and *HDUSP* based on their contig and singleton names. These eight sequences were aligned with *Solanum lycopersicum* (NP001234850) and *Nicotiana tabacum* (CAE46932) to identify the conserved regions. The amino acid sequences of Lj_08086, Lj_08422, and HDUSP have the two conserved domains, HXXXDG and DFGWG [39,77,78] (Figure 5). Therefore, *Lj_08086*, *Lj_08422*, and *HDUSP* are believed to be putative *HQT* or *HCT* genes in *L. japonica* (Figure 5A).

To further understand the relationship between HQT and HCT in *L. japonica* and the same enzymes in other species, 19 sequences from eight species were retrieved from NCBI. All sequences were divided into two groups, representing HQTs and HCTs in the phylogenetic tree (Figure 5B). HDUSP is believed to be a novel *HCT* gene in *L. japonica* and is named LjHCT1 in this study. In the HCT group, LjHCT1 and *Cynara cardunculus* var. scolymus, which share an identity of 83%, were the most similar (Figure S2).

The amino acid sequence of Lj_08086 is identical to that of LjHQT (ACZ52698) and is named LjHQT1 in this study.

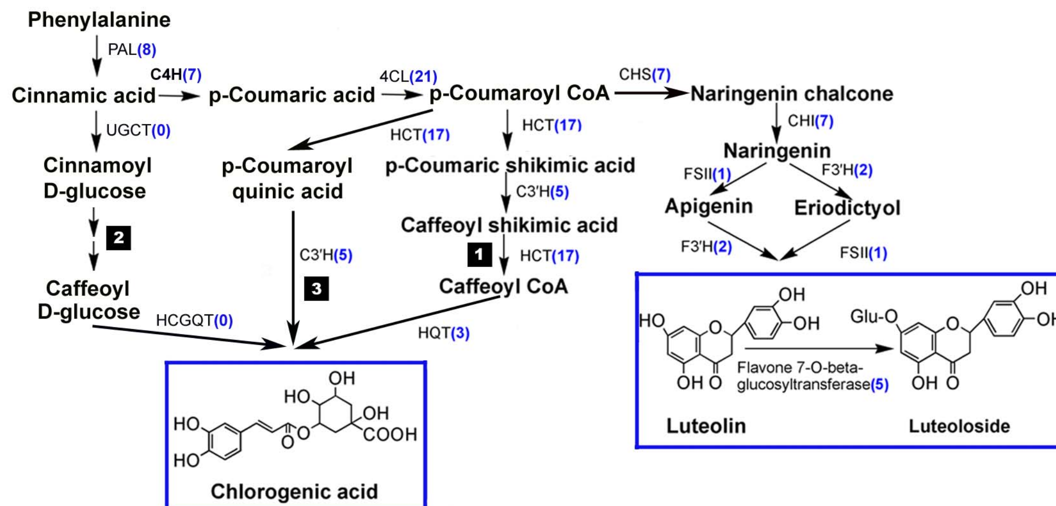

Figure 4. Proposed pathways for the biosynthesis of CGA and luteoloside in *L. japonica*. The three different routes of CGA biosynthesis are labeled 1, 2, and 3. Enzyme names are shown in the pictures. Each enzyme is annotated with the number of corresponding unigenes shown in parentheses. PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-hydroxycinnamoyl CoA ligase/4-coumarate-CoA ligase; HCT, hydroxycinnamoyl CoA shikimate/quinate hydroxycinnamoyltransferase; C3H, *p*-coumarate 3′-hydroxlase; HQT, hydroxycinnamoyl CoA quinate hydroxycinnamoyl transferase; UGCT, UDP glucose: cinnamate glucosyl transferase; HCGQT, hydroxycinnamoyl D-glucose: quinate hydroxycinnamoyl transferase. CHS, chalcone synthesis; CHI, chalcone isomerase; FSII, flavonol synthase; F3′H, Flavonoid 3′-monooxygenase.
doi:10.1371/journal.pone.0062922.g004

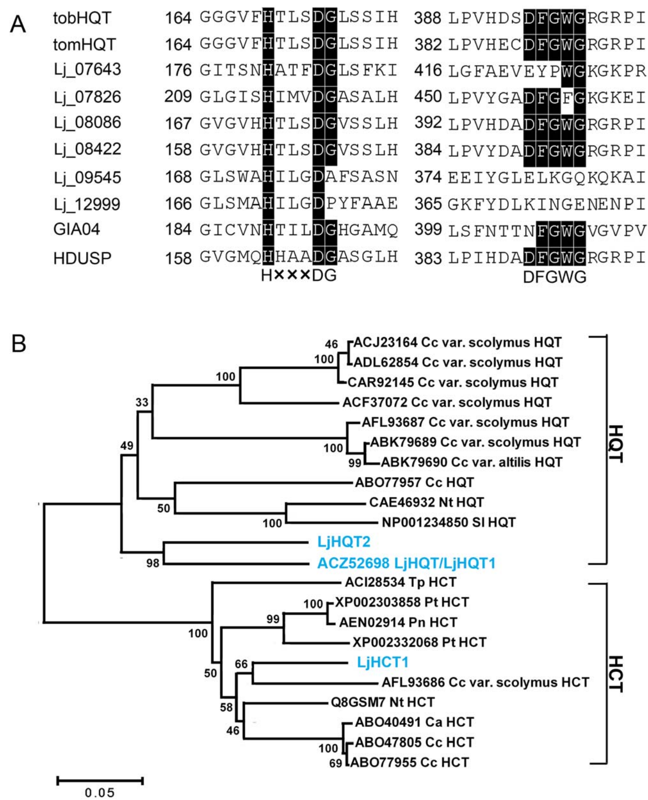

Figure 5. Protein sequence alignment of putative *L. japonica* HQTs and HCTs with representative members of the HQT and HCT families. A. Sequence alignment of the conserved structure motifs in eight *L. japonica* putative HQTs and HCTs with *Solanum lycopersicum* (NP001234850) and *Nicotiana tabacum* (CAE46932). Black boxes indicate the conserved region. B. Neighbor-joining tree of HQTs and HCTs from *L. japonica* and other plants. The HQTs and HCTs used in phylogenetic analysis were retrieved from NCBI, including *Cynara cardunculus* var. scolymus (ACJ23164, ADL62854, CAR92145, ACF37072, AFL93687, ABK79689, ABK79690, AFL93686), *Coffea canephora* (ABO77957, ABO47805, ABO77955), *Nicotiana tabacum* (CAE46932, Q8GSM7), *Solanum lycopersicum* (NP001234850), *Trifolium pratense* (ACI28534), *Populus trichocarpa* (XP002303858, XP002332068), *P. nigra* (AEN02914) and *C. arabica* (ABO40491).
doi:10.1371/journal.pone.0062922.g005

Lj_08422, a novel *HQT* gene reported in this study, is named *LjHQT2*. LjHQT/LjHQT1 and LjHQT2 were clustered into one sub-group with the highest degree of identity (79%). The shared identity between two HQT proteins in the closed sub-group including *Coffea canephora*, *N. tabacum* and *S. lycopersicum* ranged from 69% to 72% (Figure S3).

Discussion

L. japonica has pharmacological attributes, including anti-inflammatory and anti-nociceptive properties, and has been believed to be an effective Chinese medicine for thousands of years. Numerous studies have reported the isolation and pharmacological action of the bioactive components in *L. japonica*, including phenolic acids and flavonoids. The potential molecular mechanisms that produce the bioactive biosynthesis of CGA and luteoloside are still not comprehensively understood in *L. japonica*. To gain more genetic information on the secondary metabolites of *L. japonica*, we sequenced two cDNA libraries from *L. japonica* buds and leaves using the Roche/454 GS FLX platform. We identified 51,500 unigenes that we annotated based on Nr, Nt, Swissprot, KEGG, and TAIR, significantly more than that obtained in a previous work. Yuan *et al.* [52] acquired over 32 million reads and over 6000 ESTs from a library made from *L. japonica* buds. They also correlated gene expression profiles to known metabolic activities by comparing them with a transcriptome of a variety (*L. japonica* Thunb. var. chinensis (Watts)) [52]. However, reports describing the key enzymes involved in the biosynthesis pathways of active compounds, such as C4H, C3'H, HCT, and HQT, are unavailable.

We found all of the putative genes encoding the possible enzymes involved in the first and third possible routes of CGA biosynthesis (Figure 4). Regarding the second route, i.e., only one protein, HCGQT, was first purified in the 1970s from sweet potatoes and demonstrated to catalyze caffeoyl-D-glucose and quinic acid to produce CGA *in vitro* [40–42]. However, CGA: glucaric acid caffeoyltransferase (CQT) has been found in tomatoes and shown to participate in the transfer of caffeic acid from CGA to glucaric and galactaric acids [79]. Therefore, a route for the recycling of CGA may exist. Whether or not HCGQT is absolutely necessary for the biosynthesis of CGA has yet to be confirmed. Over the past 10 years, the first route involving HQT has attracted the most research attention. *HQT* genes have been discovered in several species, including tobacco, tomato, artichoke, and *L. japonica*. The catalysis of quinic acid and caffeoyl CoA to produce CGA has been identified *in vivo* [35–39]. Future experiments to produce CGA through heterologous expression of the two sets of genes in yeast may shed light on the biosynthetic pathway of CGA in *L. japonica*.

Many other related metabolic processes, including carbohydrate metabolic processes as well as amino acid and other secondary metabolic processes, have been mapped based on the KEGG

analysis. *L. japonica* contains a large number of volatile constituents that are synthesized through terpenoid metabolism, which may include many CYPs and glycotransferases. In addition, a number of putative TFs have been found in *L. japonica,* some of which are involved in the regulation of metabolic pathway genes. These metabolic pathways and the EST dataset will allow the possibility of intensive studies of *L. japonica*. In addition, the ESTs and unique sequences obtained from this study could provide an important resource for the scientific community interested in the molecular genetics and functional genomics of *L. japonica*.

Conclusion

We produced one million reads in a whole sequence run using with the Roche/454 GS FLX platform. Based on BLAST, we established an HQ EST database containing 64,184 unigenes for *L. japonica* buds and leaves using 454 GS FLX Titanium sequencing technology. Among these unigenes, 51,500 were annotated and all of the unigenes encoding enzymes involved in the CGA and luteoloside biosynthetic pathways were found. The *HCT* gene in *L. japonica* was cloned for the first time in this study and two *HQT* genes were found in *L. japonica*. In addition, 188 putative CYP unigenes and 245 putative glycosyltransferase unigenes were discovered. In *L. japonica*, *F3'H* and *FSII* encode CYP75B and CYP93B, respectively, which are involved in the biosynthesis of luteoloside. *C3'H* encodes CYP98A, which is involved in the biosynthesis of CGA. CYP may participate in the biosynthesis of triterpenoids in *L. japonica*. Using KEGG analysis, five putative 7-O-beta-glycosyltransferases unigenes that participate in the process of conversion from luteolin to luteoloside were discovered. We initiated a large-scale investigation of the transcriptome of *L. japonica* in terms of functional genomics. A large number of novel putative genes involved in CGA and luteoloside biosynthesis were identified in our EST dataset.

Materials and Methods

Plant Materials

L. japonica buds and leaves were collected from cultivated fields in Zhengcheng town, Shandong Province, China, in May 2011. The samples used in this study were obtained from local authorities (Linyi Jintai Yaoye Co., Ltd) isolated, rinsed three times with water, snap frozen in liquid nitrogen, and stored at -80°C until RNA isolation.

RNA Preparation and cDNA Synthesis

Total RNA was isolated using the Universal Plant RNA Isolation Mini Kit (BioTeke, China) according to the manufacturer's protocol. RNA quality was tested using EtBr-stained 0.8% agarose gels, and the concentration was assessed using a Nanodrop 2000 (Thermo, USA). RNA was treated with TURBO DNase I (Ambion, USA) prior to cDNA synthesis. First-strand cDNA was produced from 1 μg total RNA according to the protocol of the SMARTer^{TM} PCR cDNA Synthesis Kit (Clontech, USA). We used a modified synthetic poly (T) primer (5′-AAG CAG TGG TAT CAA CGC AGA GTG CAG T [20] VN-3′) containing a *Bsg*I digestion site upstream of the poly (T) segment. For double-stranded cDNA synthesis, the cDNA was amplified using the PCR Advantage II polymerase (Clontech) with the following thermal cycle profile: 1 min at 95°C; 17 cycles of 95°C for 15 s, 62°C for 30 s; and finally 68°C for 6 min. The PCR products were purified using the PureLink^{TM} PCR Purification Kit (Invitrogen, USA). The cDNA was then treated overnight with *Bsg*I (NEB, USA) and recovered using the QIAquick PCR Purification Kit (Qiagen, German). Approximately 500 ng of ds cDNA was used for pyrosequencing with the Roche 454 GS FLX Titanium Kit.

454 Library Preparation and Sequencing

Approximately 500 ng of amplified cDNA was sheared to produce random fragments of approximately 600 bp to 900 bp. The two libraries of buds and leaves were constructed for Roche 454 sequencing according to the manufacturer's recommendation. Each library was sent for a 1/2 run using the 454 GS FLX shotgun sequencing platform (454 Life Sciences, Roche).

EST Assembly

The raw read sequences were processed to screen and filter for weak signals and low-quality reads to yield HQ sequences. The resulting HQ reads were then submitted to the Short Read Archive at NCBI and assigned the accession numbers SRX189761 and SRX189762. The primer and adapter sequences were trimmed from the HQ reads prior to assembly. Sequences shorter than 50 bp were also removed. The remaining HQ ESTs were employed for assembly using the GS FLX *De Novo* Assembly Software, V2.6 (default parameters). After assembly, all sequences, including contigs (obtained from one cluster) and singletons (appeared only once), were named as "unigenes" for subsequent annotation. This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GAAY00000000. The version described in this paper is the first version, GAAY01000000.

Annotation and Classification

The unigenes were annotated by a BLAST search against a series of protein and nucleotide databases, including Nr (ftp: //ftp.ncbi.nih.gov/blast/db/FASTA/nr.gz), Nt (ftp: //ftp.ncbi.nih.gov/blast/db/FASTA/nt.gz), Swissprot (http: //www.uniprot.org/downloads), KEGG (ftp: //ftp.genome.jp/pub/kegg/release/archive/kegg/58/), and TAIR (http: //www.arabidopsis.org). The unigenes were compared against these databases with a significance threshold of E-value ≤1e-5. The search was limited to the first five significant hits for each query to maximize computational speed. The definition line of the top BLAST hit was used as a description of the putative function of the queried unigene. Customized Perl scripts were used to parse the BLAST outputs.

Full-length cDNA Cloning of Putative HQT and HCT Genes

RNA samples of young leaves and buds for gene cloning were converted to first-strand cDNA of the 5' and 3' ends according to the SMART^{TM} RACE cDNA protocol. All genes were amplified at 95°C for 3 min, followed by 25 cycles of 95°C for 30 s, 57°C for 30 s, and 72°C for 1 min 30 s, and a final step at 72°C for 10 min. The recycled products were integrated into a pMD18-T vector (Takara, Dalian, China) and transferred into *Escherichia coli* DH5α competent cells (Transgene, Beijing, China). The isolated clones were sequenced on a 3730XL sequencing platform (ABI, USA).

Phylogenetic Analysis

All of the hydrocinnamate transferase amino acid sequences were aligned with Clustal X [80] using the following parameters: gap opening penalty, 10; gap extension penalty, 0.1; and delay divergent cutoff, 25%. For the phylogenetic analysis, neighbor-joining tree [81] was constructed under the Jones–Taylor–Thornton substitution model [82] in MEGA4.0 [83] with a bootstrap of 1000 replicates.

Supporting Information

Figure S1 The analysis of genes expression in buds and leaves respectively. A.Venn diagram of the unigenes in the buds and leaves of *L. japonica*. B.Functional classification of unigenes in the two *L. japonica* organs based on GO categories. Unique sequences were classified into three major categories: cellular components, molecular functions and biological processes on the basis of the TAIR GO slim.
(TIF)

Figure S2 Alignment of HCT amino acid sequence from *Lonicera japonica* and other species. *Trifolium pratense* (ACI28534), *Populus trichocarpa* (XP002303858, XP002332068), *Cynara cardunculus* var. scolymus (AFL93686), *Nicotiana tabacum* (Q8GSM7), *Coffea canephora* (ABO77957, ABO47805, and ABO77955). *Lonicera japonica* (LjHCT1).
(TIF)

Figure S3 Alignment of HQT amino acid sequence from *Lonicera japonica* and other species. *Cynara cardunculus* var. scolymus (ACJ23164, ADL62854, CAR92145, ACF37072, AFL93687, ABK79689, and AFL93686), *Coffea canephora* (ABO77957), *Nicotiana tabacum* (CAE46932), *Solanum lycopersicum* (NP001234850). *Lonicera japonica* (ACZ52689/LjHQT1, LjHQT2).
(TIF)

Table S1 Classification of the candidate CYP genes.
(DOC)

Table S2 Summary of the 10 candidate HQT or HCT genes in L. japonica.
(DOC)

Table S3 List of putative unigenes related to chlorogenic acid biosynthesis.
(DOC)

Table S4 List of putative unigenes related to luteoloside biosynthesis.
(DOC)

Author Contributions

Conceived and designed the experiments: SC WX J. Song CS ZW. Performed the experiments: LH XX CL. Analyzed the data: YL YZ RC. Contributed reagents/materials/analysis tools: LH HY ZS YB J. Shen. Wrote the paper: LH XX.

References

1. Schierenbeck KA (2004) Japanese Honeysuckle (*Lonicera japonica*) as an invasive species; history, ecology, and context. CRIT REV PLANT SCI 23: 391–400.
2. Chang WC, Hsu FL (1992) Inhibition of platelet activation and endothelial cell injury by polyphenolic compounds isolated from *Lonicera japonica* Thunb. PLEFA 45: 307–312.
3. Clifford MN (2000) Chlorogenic acids and other cinnamates-nature, occurrence, dietary burden, absorption and metabolism. J Sci Food Agr 80: 1033–1043.
4. Lu HT, Jiang Y, Chen F (2004) Application of preparative high-speed counter-current chromatography for separation of chlorogenic acid from Flos Lonicerae. J Chromatogr A 1026: 185–190.
5. Inagaki I, Sakushima A, Hisada S, Nishibe S, Morita N (1974) Comparison of lonicerin and veronicastroside. Yakugaku Zasshi 94: 524–525.
6. Kumar N, Singh B, Bhandari P, Gupta AP, Uniyal SK, et al. (2005) Biflavonoids from *Lonicera japonica*. Phytochem 66: 2740–2744.
7. Chen J, Li SL, Li P, Song Y, Chai XY, et al. (2005) Qualitative and quantitative analysis of active flavonoids in Flos Lonicerae by capillary zone electrophoresis coupled with solid-phase extraction. JSS 28: 365–372.
8. Qian ZM, Li HJ, Li P, Chen J, Tang D (2007) Simultaneous quantification of seven bioactive components in Caulis *Lonicerae Japonicae* by high performance liquid chromatography. Biomed Chromatogr 21: 649–654.
9. Xing J, Li P (1999) Research on chemical constituents of Lonicera: A review and prospects. Zhong Yao Cai 22: 366–370.
10. Kwak WJ, Han CK, Chang HW, Kim HP, Kang SS, et al. (2003) Loniceroside C, an antiinflammatory saponin from *Lonicera Japonica*. Chem Pharm Bull 51: 333–335.
11. Chai XY, Li SL, Li P (2005) Quality evaluation of Flos Lonicerae through a simultaneous determination of seven saponins by HPLC with ELSD. J Chromatogr A 1070: 43–48.
12. Choi CW, Jung H, Kang S, Choi J (2007) Antioxidant constituents and a new triterpenoid glycoside from Flos Lonicerae. Arch Pharm Res 30: 1–7.
13. Lin LM, Zhang XG, Zhu JJ, Gao HM, Wang ZM, et al. (2008) Two new triterpenoid saponins from the flowers and buds of *Lonicera japonica*. J Asian Nat Products Res 10: 925–929.
14. Schlotzhauer WS, Pair SD, Horvat RJ (1996) Volatile constituents from the flowers of Japanese Honeysuckle (*Lonicera japonica*). J Agric Food Chem 44: 206–209.
15. Li H, Zhang Z, Li P (2002) Comparative study on volatile oils in flower and stem of *Lonicera Japonica*. Zhong Yao Cai 25: 476–477.
16. Tsuchiya T, Suzuki O, Igarash K (1996) Protective effects of chlorogenic acid on paraquat-induced oxidative stress in rats. BBB 60: 765–768.
17. Wu L, Zhang ZJ, Zhang ZS (2007) Characterization of antioxidant activity of extracts from Flos Lonicerae. Drug Dev Ind Pharm 33: 841–847.
18. Tang D, Li HJ, Chen J, Guo CW, Li P (2008) Rapid and simple method for screening of natural antioxidants from Chinese herb Flos Lonicerae Japonicae by DPPH-HPLC-DAD-TOF/MS. J Sep Sci 31: 3519–3526.
19. Lee S, Shin E, Son K, Chang H, Kang S, et al. (1995) Anti-inflammatory activity of the major constituents of *Lonicera Japonica*. Arch Pharm Res 18: 133–135.
20. Mori H, Tanaka T, Shima H (1986) Inhibitory effect of chlorogenic acid on methylazoxymethanol acetate-induced carcinogenesis in large intestine and liver of hamsters. Cancer Lett 30: 49–54.
21. Tanaka T, Nishikawa A, Shima H, Sugie S, Shinoda T, et al. (1990) Inhibitory effects of chlorogenic acid, reserpine, polyprenoic acid (E-5166), or coffee on hepatocarcinogenesis in rats and hamsters. Basic Life Sci 52: 429–440.
22. Tanaka T, Kojima T, Kawamori T, Wang A, Suzui M, et al. (1993) Inhibition of 4-nitroquinoline-1-oxide-induced rat tongue carcinogenesis by the naturally occurring plant phenolics caffeic, ellagic, chlorogenic and ferulic acids. Carcinogenesis 14: 1321–1325.
23. Kurata R, Adachi M, Yamakawa O, Yoshimoto M (2006) Growth suppression of human cancer cells by polyphenolics from sweetpotato (*Ipomoea Batatas* L.) leaves. J Agric Food Chem 55: 185–190.
24. Bandyopadhyay G, Biswas T, Roy KC, Mandal S, Mandal C, et al. (2004) Chlorogenic acid inhibits Bcr-Abl tyrosine kinase and triggers p38 mitogen-activated protein kinase-dependent apoptosis in chronic myelogenous leukemic cells. Blood 104: 2514–2522.
25. Feng R, Lu Y, Bowman LL, Qian Y, Castranova V, et al. (2005) Inhibition of activator protein-1, NF-κB, and MAPKs and induction of phase 2 detoxifying enzyme activity by chlorogenic acid. J Biol Chem 280: 27888–27895.
26. Qiu F, Li ZX, He L, Wang D (2012) HPLC-ESI-MS/MS analysis and pharmacokinetics of luteoloside, a potential anticarcinogenic component isolated from *Lonicera japonica*, in beagle dogs. Biomed Chromatogr doi: 10.1002/bmc.2793.
27. Wang GF, Shi LP, Ren YD, Liu QF, Liu HF, et al. (2009) Anti-hepatitis B virus activity of chlorogenic acid, quinic acid and caffeic acid in vivo and in vitro. Antiviral Res 83: 186–190.
28. Noriaki K, Yukari K, Kazuya I, Kyo M, Tokio F (2005) In vitro antibacterial, antimutagenic and anti-influenza virus activity of caffeic acid phenethyl esters. Biocontrol Sci 10: 155–161.
29. Osborne TB, Campbell GF (1897) The proteids of sunflower seed. J Am Chem Soc 19: 487–492.
30. Bradfield AE, Flood AE, Hulme AC, Williams AH (1952) Chlorogenic acids in fruit trees. Nature 170: 168–169.
31. Dickinson D, Gawler JH (1954) The chemical constituents of victoria plums: Preliminary qualitative analysis. J Sci Food Agric 5: 525–529.
32. Zucker M, Levy CC (1959) Some factors which affect the synthesis of chlorogenic acid in disks of potato tuber. Plant Physiol 34: 108–112.
33. Stöckigt J, Zenk MH (1974) Enzymatic synthesis of chlorogenic acid from caffeoyl coenzyme A and quinic acid. FEBS Lett 42: 131–134.
34. Ulbrich B, Zenk MH (1979) Partial purification and properties of hydroxycinnamoyl-CoA: quinate hydroxycinnamoyl transferase from higher plants. Phytochem 18: 929–933.
35. Niggeweg R, Michael AJ, Martin C (2004) Engineering plants with increased levels of the antioxidant chlorogenic acid. Nat Biotechnol 22: 746–754.
36. Lepelley M, Cheminade G, Tremillon N, Simkin A, Cailet V, et al. (2007) Chlorogenic acid synthesis in coffee: An analysis of CGA content and real-time RT-PCR expression of HCT, HQT, C3H1, and CCoAOMT1 genes during grain development in *C.canephora*. Plant Sci 172: 978–996.
37. Comino C, Hehn A, Moglia A, Menin B, Bourqaud F, et al. (2009) The isolation and mapping of a novel hydroxycinnamoyl transferase in the globe artichoke chlorogenic acid pathway. BMC Plant Bio 9: 30.
38. Peng XX, Li WD, Wang WQ, Bai G (2010) Cloning and characterization of a cDNA coding a hydroxycinnamoyl-coA quinate hydroxycinnamoyl transferase

involved in chlorogenic acid biosynthesis in *Lonicera japonica*. Planta Med 76: 1921–1926.
39. Sonnante G, Amore RD, Blanco E, Pierri CL, De Palma M, et al. (2010) Novel hydroxycinnamoyl-coenzyme A quinate transferase genes from artichoke are involved in the synthesis of chlorogenic acid. Plant Physiol 153: 1224–1238.
40. Uritani I, Kojima M (1972) Studies on chlorogenic acid biosynthesis in sweet potato root tissue using trans-cinnamic acid-2-14C and quinic acid-G-3H. Plant Cell Physiol 13: 311–319.
41. Villegas RJA, Kojima M (1986) Purification and characterization of hydroxycinnamoyl D-glucose quinate hydroxycinnamoyl transferase in the root of sweet potato, *Ipomoea batatas* Lam. J Biol Chem 261: 8729–8733.
42. Strack D, Gross W (1990) Properties and activity changes of chlorogenic acid: glucaric acid caffeoyltransferase from tomato (*Lycopersicon esculentum*). Plant Physiol 92: 41–47.
43. Schoch G, Goepfert S, Morant M, Hehn A, Meyer D, et al. (2001) CYP98A3 from Arabidopsis thaliana is a 3′ hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. J Biol Chem 276: 36566–36574.
44. Franke R, Humphreys JM, Hemm MR, Denault JW, Ruegger MO, et al. (2002) The Arabidopsis REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. Plant J 30: 33–45.
45. Schulz M, Weissenböck G (1988) Three specific UDP-glucuronate: Flavone-glucuronosyl-transferases from primary leaves of *Secale cereale*. Phytochem 27: 1261–1267.
46. McIntosh CA, Mansell RL (1990) Biosynthesis of naringin in *Citrus paradise* UDP-glucosyltransferase activity in grapefruit seedlings. Phytochem 29: 1533–1538.
47. McIntosh CA, Latchinian L, Mansell RL (1990) Flavanone-specific 7-O-glucosyltransferase activity in Citrus paradisi seedlings: purification and characterization. Arch Biochem Biophys 282: 50–57.
48. Vellekoop P, Lugones L, van Brederode J (1993) Purification of an UDP-glucose: flavone, 7-O-glucosyltransferase, from *Silene latifolia* using a specific interaction between the enzyme and phenyl-Sepharose. FEBS Lett 330: 36–40.
49. Stich K, Halbwirth H, Wurst F, Forkmann G (1997) UDP-glucose: flavonol 7-O-glucosyltransferase activity in flower extracts of *Chrysanthemum segetum*. Z Naturforsch 52c: 153–158.
50. Hirotani M, Kuroda R, Suzuki H, Yoshikawa T (2000) Cloning and expression of UDP-glucose: flavonoid 7-O-glucosyltransferase from hairy root cultures of *Scutellaria baicalensis*. Planta 210: 1006–1013.
51. Kitada C, Gong Z, Tanaka Y, Yamazaki M, Saito K (2001) Differential expression of two cytochrome P450s involved in the biosynthesis of flavones and anthocyanins in chemo-varietal forms of *Perilla frutescens*. Plant Cell Physiol 42(12): 1338–1344.
52. Yuan Y, Song LP, Li MH, Liu GM, Chu YN, et al. (2012) Genetic variation and metabolic pathway intricacy govern the active compound content and quality of the Chinese medicinal plant *Lonicera japonica* thunb. BMC Genomics 13: 195.
53. Wu B, Li Y, Yan HX, Ma YM, Luo HM, et al. (2012) Comprehensive transcriptome analysis reveals novel genes involved in cardiac glycoside biosynthesis and mlncRNAs associated with secondary metabolism and stress response in *Digitalis purpurea*. BMC Genomics 13: 15.
54. Nicolai M, Pisani C, Bouchet JP, Vuylsteke M, Palloix A (2012) Discovery of a large set of SNP and SSR genetic markers by high-throughput sequencing of pepper (*Capsicum annuum*). Genet Mol Res 11: 2295–2300.
55. Caniard A, Zerbe P, Legrand S, Cohade A, Valot N, et al. (2012) Discovery and functional characterization of two diterpene synthases for sclareol biosynthesis in *Salvia sclarea* (L.) and their relevance for perfume manufacture. BMC Plant Biol 12(1): 119.
56. Wang LM, Li MT, Yan YY, Ao MZ, Wu G, et al. (2009) Influence of flowering stage of *Lonicera japonica* Thunb. on variation in volatiles and chlorogenic acid. J Sci Food Agric 89: 953–957.
57. Liu BC, Sun RY, Li RQ (2011) Study on the extraction and content determination of chlorogenic acid in different organs of *Lonicera japonica* Thunb. Medicinal Plant 2: 35–37.
58. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402.
59. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M (2012) KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res 40: D109–D114.
60. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, et al. (1999) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 27: 29–34.
61. Berardini TZ, Mundodi S, Reiser R, Huala E, Garcia-Hernandez M, et al. (2004) Functional annotation of the Arabidopsis genome using controlled vocabularies. Plant Physiol 135(2): 745–755.
62. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. Nat Genet 25: 25–29.
63. Tatusov RL, Koonin EV, Lipman DJ (1997) A genomic perspective on protein families. Science, 278: 631–637.
64. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28: 27–30.
65. Li Q, Cao JH, Y LJ, Li MT, Liao JJ, et al. (2012) Effects on physiological characteristics of Honeysuckle (*Lonicera japonica* Thunb) and the role of exogenous calcium under drought stress. Plant Omics Journal 5: 1–5.
66. Cao Y, Liu MG, Li MJ (2009) Contrast Analysis on physiological indicators of cold resistance in *Lonicera japonica*. Shandong Forestry Science and Technology 3: 29–31.
67. Liu Z, He X, Chen W, Yuan F, Yan K, et al. (2009) Accumulation and tolerance characteristics of cadmium in a potential hyperaccumulator–*Lonicera japonica* Thunb. J Hazard Mater 169: 170–5.
68. Nakashima K, Ito Y, Yamaguchi-Shinozaki K (2009) Transcriptional regulatory networks in response to abiotic stresses in Arabidopsis and grasses. Plant Physiol 149: 88–95.
69. Yamaguchi-Shinozaki K, Shinozaki K (2006) Transcriptional regulatory networks in cellularresponses and tolerance to dehydration and cold stresses. Annu Rev Plant Biol 57: 781–803.
70. Yilmaz A, Mejia-Guerra MK, Kurz K, Liang X, Welch L, et al. (2011) AGRIS: the Arabidopsis gene regulatory information server, an update. Nucleic Acids Res 39: D1118–1122.
71. Sakamoto H, Maruyama K, Sakuma Y, Meshi T, Iwabuchi M, et al. (2004) Arabidopsis Cys2/His2-type zinc-finger proteins function as transcription repressors under drought, cold, and high-salinity stress conditions. Plant Physiol 136: 2734–2746.
72. Cominelli E, Galbiati M, Vavasseur A, Conti L, Sala T, et al. (2005) A guard-cell-specific MYB transcription factor regulates stomatal movements and plant drought tolerance. Curr Biol 15: 1196–1200.
73. Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, et al. (2002) DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. Biochem and Biophys Res Commun 290: 998–1009.
74. Borevitz JO, Xia Y, Blount J, Dixon RA, Lamb C (2000) Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. Plant Cell 12: 2383–2393.
75. Menke FLH, Champion A, Kijne JW, Memelink J (1999) A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene Str interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. EMBO J 18: 4455–4463.
76. Van der Fits L, Memelink J (2000) ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. Science 289: 295–297.
77. St-Pierre B, De Luca V (2000) Evolution of acyltransferase genes: origin and diversification of the BAHD superfamily of acyltransferases involved in secondary metabolism. Recent Adv Phytochem 34: 285–315.
78. D'Auria JC (2006) Acyltransferases in plants: a good time to be BAHD. Curr Opin Plant Biol 9: 331–340.
79. Strack D, Gross W (1990) Properties and activity changes of chlorogenic acid: glucaric acid caffeoyltransferase from tomato (*Lycopersicon esculentum*). Plant Physiol 92: 41–47.
80. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25: 4876–4882.
81. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425.
82. Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8: 275–282.
83. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24 : 1596–1599.