

# Transcriptome analysis of a medicinal plant, *Pistacia chinensis*

Ki-Young Choi · Duck Hwan Park · Eun-Soo Seong · Sang Woo Lee · Jin Hang · Li Wan Yi · Jong-Hwa Kim · Jong-Kuk Na

Received: 12 November 2019 / Revised: 2 December 2019 / Accepted: 2 December 2019  
© Korean Society for Plant Biotechnology

**Abstract** *Pistacia chinensis* Bunge has not only been used as a medicinal plant to treat various illnesses but its young shoots and leaves have also been used as vegetables. In addition, *P. chinensis* is used as a rootstock for *Pistacia vera* (pistachio). Here, the transcriptome of *P. chinensis* was sequenced to enrich genetic resources and identify secondary metabolite biosynthetic pathways using Illumina RNA-seq methods. *De novo* assembly resulted in 18,524 unigenes with an average length of 873 bp from 19 million RNA-seq reads. A Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation tool assigned KO (KEGG orthology) numbers to 6,553 (36.2%) unigenes, among which 4,061 unigenes were mapped into 391 different metabolic pathways. For terpenoid backbone and carotenoid biosynthesis pathways, 44 and 22 unigenes encode enzymes corresponding to 30 and 16 entries, respectively. Twenty-two unigenes encode proteins for 16 entries of the carotenoid biosynthesis pathway. As for the phenylpropanoid and

flavonoid biosynthesis pathways, 63 and 24 unigenes were homologous to 17 and 14 entry proteins, respectively. Mining of simple sequence repeat identified 2,599 simple sequence repeats from *P. chinensis* unigenes. The results of the present study provide a valuable resource for in-depth studies on comparative and functional genomics to unravel the underlying mechanisms of the medicinal properties of *Pistacia* L.

**Keywords** Transcriptome, Medicinal plant, *Pistacia*

## Introduction

The genus *Pistacia* L., belonging to the family Anacardiaceae, is consisting of more than 13 species (Rauf et al. 2017). *P. vera* is a well-known species from the genus *Pistacia* and has been studied in various aspects due to its agronomical importance. *P. chinensis* is distributed in China, Taiwan, and Pakistan as well as North America. This small tree has been used for landscape and shade tree and its seed oil possesses not only biodiesel properties, but also pest repelling properties (Rashed et al. 2016). Tender burgeon of *P. chinensis* has been used as vegetables, and seeds can be used to make confectionery or vegetable oil. Also, *P. chinensis* has been used as a rootstock for *P. vera* because it has a strong adaptability and resistance to adverse environment (Tang et al. 2012).

Several *Pistacia* species, including *P. chinensis*, have been used as folk medicines to treat various illnesses (Akhtar et al. 2013; Bozorgi et al. 2013; Rauf et al. 2017). *P. vera* has been used to treat abdominal ailments and rheumatism (Bozorgi et al. 2013; Rauf et al. 2017). *P. khinjuk* and *P. integerrima* have been used to treat hepatitis and liver disorder (Rauf et al. 2017). *P. lentiscus* has been used to relieve various illnesses such as coughs, sore throats, eczema, and kidney stones (Rauf et al. 2017). Similarly,

<sup>†</sup>Equal contribution

K.-Y. Choi<sup>†</sup> · J.-K. Na (✉)  
Department of Controlled Agriculture, Kangwon National University, Chuncheon, Kangwon 24341, Republic of Korea  
e-mail: jongkook@kangwon.ac.kr

D. H. Park<sup>†</sup>  
Division of Bioresource Sciences, Kangwon National University, Chuncheon, Kangwon 24341, Republic of Korea

E.-S. Seong  
Department of Medicinal Plants, Suwon Women's University, Suwon 18333, Republic of Korea

S. W. Lee  
International Biological Material Research Center, KRIBB, Daejeon 34141, Republic of Korea

J. Hang · L. W. Yi  
Yunnan Academy of Agricultural Sciences, Yunnan 650223, China

J.-H. Kim  
Department of Horticulture, Kangwon National University, Chuncheon, Kangwon 24341, Republic of Korea

all parts of *P. chinensis* can be used for medicinal purpose to relieve dysentery, inflammatory swelling, psoriasis and rheumatism (Tang et al. 2012) and also to treat jaundice and liver diseases (Akhtar et al. 2013). Other *Pistacia* species such as *P. terebinthus*, *P. palaestina*, *P. eurycarpa*, *P. weinmannifolia*, and *P. atlantica* have also been used to relieve illnesses (Rauf et al. 2017).

Various studies have shown that most *Pistacia* species contain diverse and valuable secondary metabolites (Bozorgi et al. 2013; Rauf et al. 2017). Like other *Pistacia* species, various metabolites were identified from *P. chinensis*. Identified metabolites include new *N*-phenyl-pyrrolidone derivative (Liu et al. 2008), gallic acid and 6-O-galloyl arbutin-quercitrin (Shi and Zuo 1992),  $\beta$ -sitosterol, luepol, myricetin 3-O- $\alpha$ -rhamnoside (Rashed et al. 2016), 4-aryldihydrocoumarins (Nishimura et al. 2000). These chemicals have been used or have potentials for pharmaceutical drugs. Even though several *Pistacia* species are known to have valuable pharmaceutical and industrial properties, genomic and genetic studies of *Pistacia* species have not been vigorously studied except for *P. vera*. Therefore, it is essential to enrich genomic and genetic resources to delve into genetic insight of the genus *Pistacia* and also to identify not only key genes but also biosynthetic pathways involved in useful metabolite biosynthesis.

Advanced RNA sequencing transcriptome analysis tool is in preference to get transcriptome information of the target organism because it is cost- and time-effective (Lee et al. 2015). Because RNA sequencing method generates transcriptome information in a short time, this tool is frequently utilized to analyze transcriptome information of non-model organisms including medicinal plants (Bae et al. 2018; Eum et al. 2019). An increasing number of previously unexplored medicinal plants have been sequenced through this advanced sequencing technology, providing genomic resources for unravelling genes and biosynthetic pathways involved in metabolite biosynthesis in various medicinal plants (Bae et al. 2018; Eum et al. 2019; Kotwal et al. 2016; Loke et al. 2016; Rai et al. 2016). However, a vast majority of medicinal plants are yet to be studied.

In this study, RNA-seq transcriptome analysis was performed to characterize genomic features of *P. chinensis*. Generated RNA-seq data were used *de novo* assembly, and resulting unigenes were used for gene ontology (GO) analysis, KEGG metabolic pathway search, and SSR mining. In addition, the reference transcriptome of *P. chinensis* would be very useful resources for enriching and facilitating genetic/genomic studies, molecular marker discovery, and various genetic/biological studies of not only *P. chinensis*

but also the genus *Pistacia* and close relatives.

## Materials and Methods

### Plant materials

Fresh leaf tissues of a fully grown *P. chinensis* Bunge were harvested June 2015 at Wunnan, China and submerged into liquid N<sub>2</sub>. For long-distance transportation, the leaf samples were transferred into RNAlater solution (Ambion Ins, USA), which was stored in -20°C freezer prior to mRNA extraction for RNA-seq analysis.

### Preparation of RNA-seq library

Leaf samples of *P. chinensis* were used for total RNA extraction using TRIzol reagent by following the manufacturer's instructions. The remaining procedures were performed according to the methods described by Bae et al. (2018) and Eum et al. (2019).

### *De novo* assembly and unigene annotation, and

For assembly, Trimmomatic tool (Bolger et al. 2014) was used to remove low quality reads (< Q20) and the read with a length < 50 bp. The following methods were performed according to the methods described by Bae et al. (2018). Briefly, three different assemblers were used for *de novo* assembly of the trimmed raw reads. Identified unigenes were annotated by comparing similarity to known proteins deposited to the NCBI non-redundant (NR) protein database.

### GO analysis and KEGG pathway search

The Blast2GO analysis tool was used for GO annotation of unigenes, of which annotation information was used for functional classification of unigenes using WEGO software. For KEGG pathway search (<http://www.genome.jp/kegg>), all unigene sequences were run on the annotation by BlastKOALA (KEGG Orthology And Links Annotation), from which assigned K numbers were used to construct KEGG pathway search.

### Identification of SSR and repetitive sequence in *P. chinensis* unigenes

To examine SSR accumulation in the unigenes from *P.*

*chinensis*, SSRs were identified using MISA, MicroSatellite identification tool ([pgrc.ipk-gatersleben.de/misa](http://pgrc.ipk-gatersleben.de/misa)). Criteria of MISA were set to a minimum of four motif repeats and  $\geq 12$  bp. To examine a content of repetitive sequence in the unigenes, RepeatMasker (v. 4.0.7) was run on the unigenes with a default mode using the reference library, RepBaseRepeatMaskerEdition-20170127 ([www.girinst.org](http://www.girinst.org)). To compare SSR distributions of *P. chinensis* to other plant species, transcriptome data from several plant species (*Mangifera indica*, *Sterculia lanceolata*, *Clausena excavata*, *Arabidopsis thaliana*, and *Oryza sativa*) were used. Transcripts of *Mangifera indica* (Mango) were retrieved from NCBI nucleotide database. Transcript sequences of both *Sterculia lanceolata* and *Clausena excavata* were obtained from NCBI transcriptome shotgun assembly data. Transcripts of *Oryza sativa* (version 7.0) and *Arabidopsis thaliana* (ATH\_cDNA\_sequences\_20101108) were retrieved from the Rice Genome database (<http://rice.plantbiology.msu.edu>) and from the TAIR database (<https://www.arabidopsis.org>), respectively.

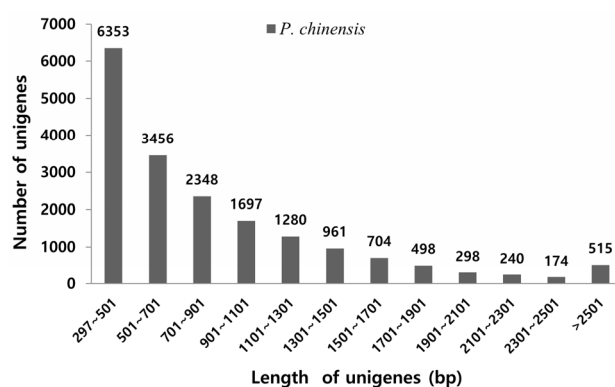
## Results

*De novo* assembly of RNA-seq data and transcriptome annotation

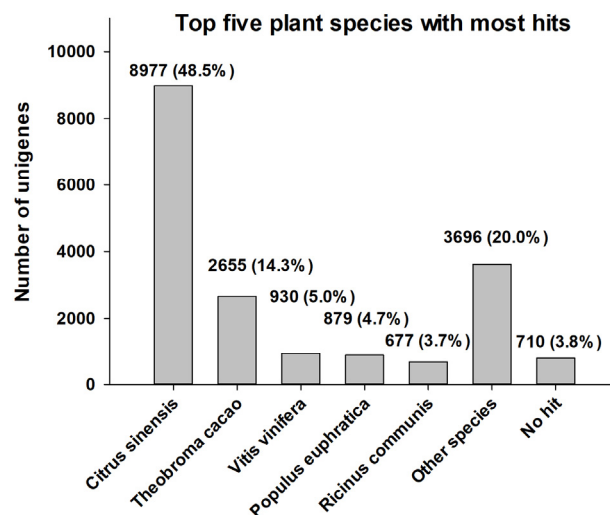
RNA-seq whole transcriptome sequencing generated 20.8 million of raw reads (~2.6 Gb) from *P. chinensis*. *De novo* assembly using trimmed raw reads by Trimmomatic trimmer (Bolger et al. 2014) resulted in 19 million clear reads with a total of read length of ~2.36 Gb. From assembly using clear reads, a total of 18,524 unigenes with a length of 16,174,683 bp were generated (Table 1), of which GC content was 40.7%. N50 was 1,104 bp, and average length

of unigenes was 873 bp (Table 1). The length of the unigenes ranged from 300 to 9,942 bp. A total of 6,353 (34.3%) unigenes was found between 297~500 bp, followed by 3,456 (18.7%) between 401~500 bp, and 2,348 (12.7%) between 701~900 bp (Fig. 1). A total of 515 unigenes (2.8%) were longer than 2,501 bp (Fig. 1).

To annotate the unigenes, protein sequences of the unigenes were searched for similarity against NCBI non-redundant (NR) protein database. Among 18,524 unigenes, 17,814 unigenes (96.2%) were aligned to protein sequences from other organisms, whereas 710 unigenes (3.8%) did not show similarity to other known proteins (Fig. 2). Top-five plant species with most hits with annotated unigenes were *Citrus sinensis* with 8,977 unigenes (48.5%), *Theobroma cacao* with 2,655 (14.3%), *Vitis vinifera* 930 (5.0%), *Populus euphratica* 879 (4.7%), and *Ricinus communis* 677 (3.7%), respectively. As shown, a half of annotated unigenes showed higher similarity to *Citrus sinensis* (Fig. 2).



**Fig. 1** Length distribution of unigenes from transcriptome of *Pistacia chinensis*



**Fig. 2** Top five plant species with higher homologous genes

**Table 1** Summary of sequencing and assembly data

Data description	Data summary
Total number of raw reads	10,621,059
Total length of raw reads (bp)	2,676,506,868
Number of filtered reads used for assembly	9,625,676
Total length of filtered reads (bp)	2,358,007,393
Number of assembled contigs (Unigenes)	18,524
Total length of assembled contigs (bp)	16,174,683
Average length (bp)	873
Length of largest contig (bp)	9,942
N50 (bp)	1,104
GC content (%)	40.7

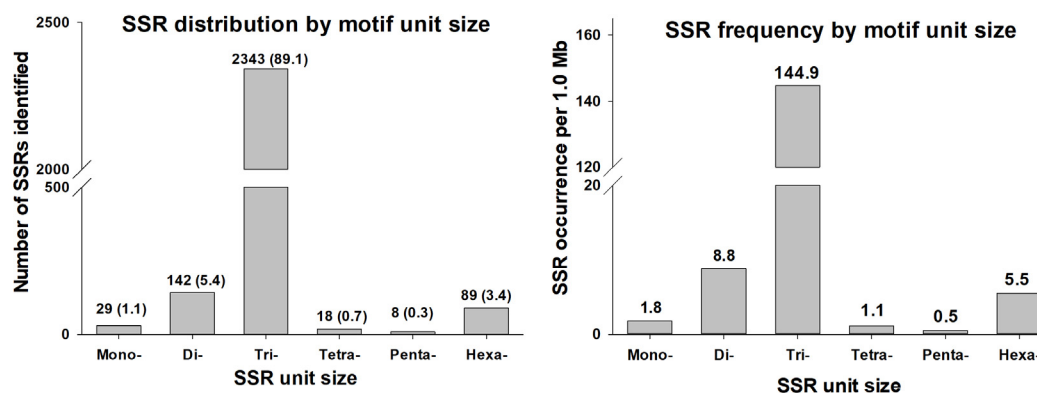


Fig. 3 SSR distribution and frequencies by repeat unit size

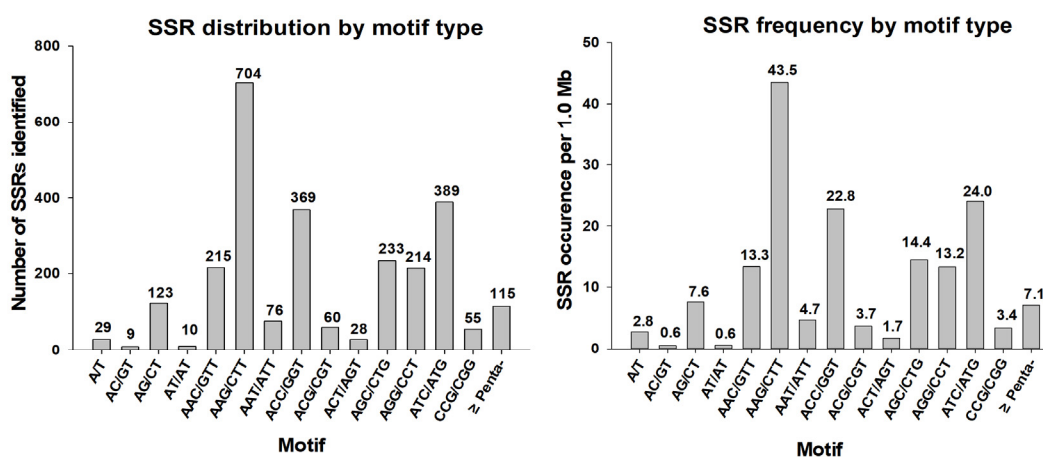


Fig. 4 SSR distribution by motif types

Accumulation of repetitive sequences and simple sequence repeats

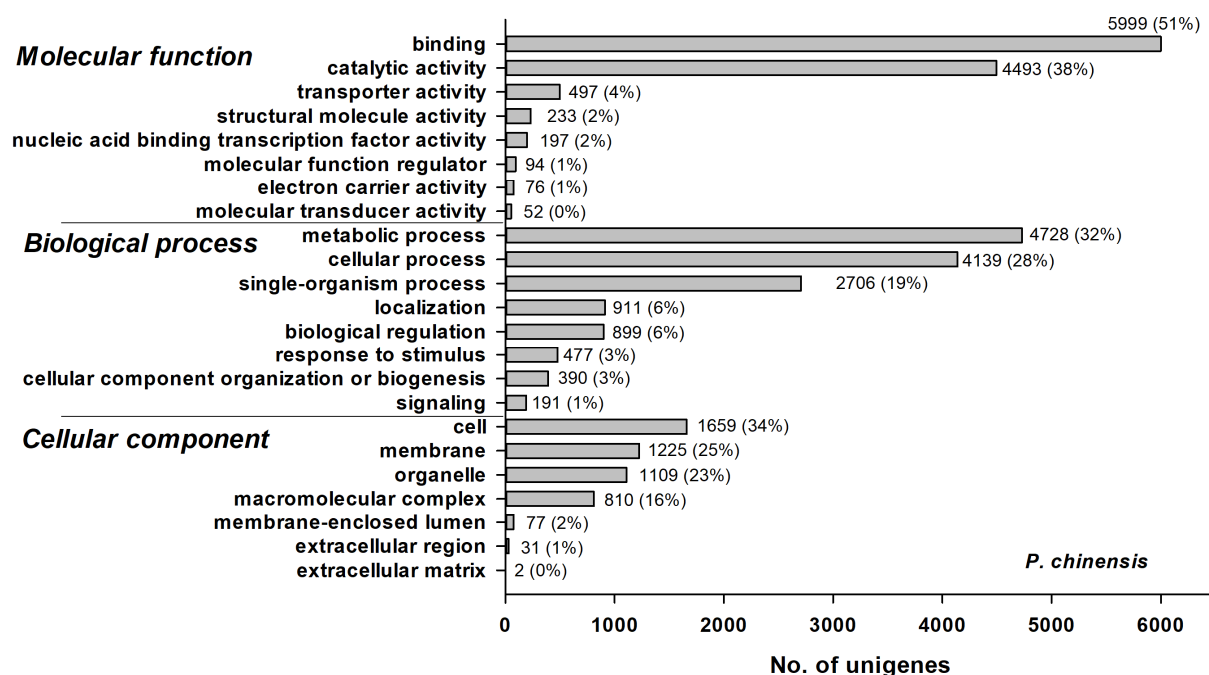
To examine the content of repetitive sequences, Repeat-Masker (<http://www.repeatmasker.org>) was run on unigenes, resulting in that 149,205 bp (0.92%) were occupied by repetitive and low complexity sequences (Supplementary Table S1). Simple sequence repeats were the most abundant repetitive sequences among identified repetitive sequences: 2,670 elements occupying 116,432 bp (0.72%) of the *P. chinensis* transcriptome (Supplementary Table S1). A total of 18 interspersed repeat elements were identified from *P. chinensis* unigenes.

SSR search using MISA SSR search tool identified a total of 2,629 perfect SSRs from 2,041 unigenes (Supplementary Table S2). Among SSR-containing unigenes, 393 had more than one SSR. The frequency of all identified SSRs was 162.5 per one million base pairs (Mbp). Tri-nucleotide SSRs were the most abundant SSRs with 2,343 (89.1%) occurrences, followed by di-nucleotide SSRs with 142 (5.4%) occurrences. Frequency of tri-nucleotide SSRs

were 144.9 per Mbp (Fig. 3). The highest SSR frequency by motif type was AAG/CTT motif with 43.5 occurrences per Mbp, followed by ACC/GGT motif with 22.8 per Mbp (Fig. 4; Supplementary Table S2). Among di-nucleotide repeats, AG/CT motif showed the highest 7.6 occurrences per Mbp, while AAAG/CTTT motif showed the highest frequency with 0.6 per Mbp among tetra-nucleotide motifs (Fig. 4; Supplementary Table S2).

#### Functional classification of unigenes by GO analysis

To examine the functional classification of annotated genes, unigenes were assigned by the GO functional term using Blast2GO software (Conesa et al. 2005) and classified using WEGO tool. By functional categorization of annotated unigenes, 9,020 were classified into molecular functions at level one, 6,133 into biological processes, and 2,622 unigenes classified into cellular component, respectively. Most unigenes belonging to molecular function at level one were classified into two major categories, binding with 5,999 and catalytic activity with 4,493 unigenes (Fig. 5). As for unigenes in



**Fig. 5** Gene Ontology functional categorization of annotated unigenes

**Table 2** Categories of KEGG metabolic pathways and their associated entries and unigenes

Category	No. of sub categories	No. of pathways	No. of entry	No. of associated genes
Metabolism	12	138	3324	6715
Genetic Information Processing	4	22	859	1520
Environmental Information Processing	3	35	424	1242
Cellular Processes	5	31	549	1280
Organismal Systems	10	84	607	1608
Human Diseases	12	81	935	2131

biological process, most genes were classified into three sub-categories, metabolic process with 4,728 genes, cellular process with 4,139, and single-organism process with 2,706 (Fig. 5). In cellular component, a majority of genes were classified into four sub-categories, cell with 1,659 genes, membrane with 1,225, organelle with 1,109, and macromolecular complex with 810 (Fig. 5; Supplementary Table S3).

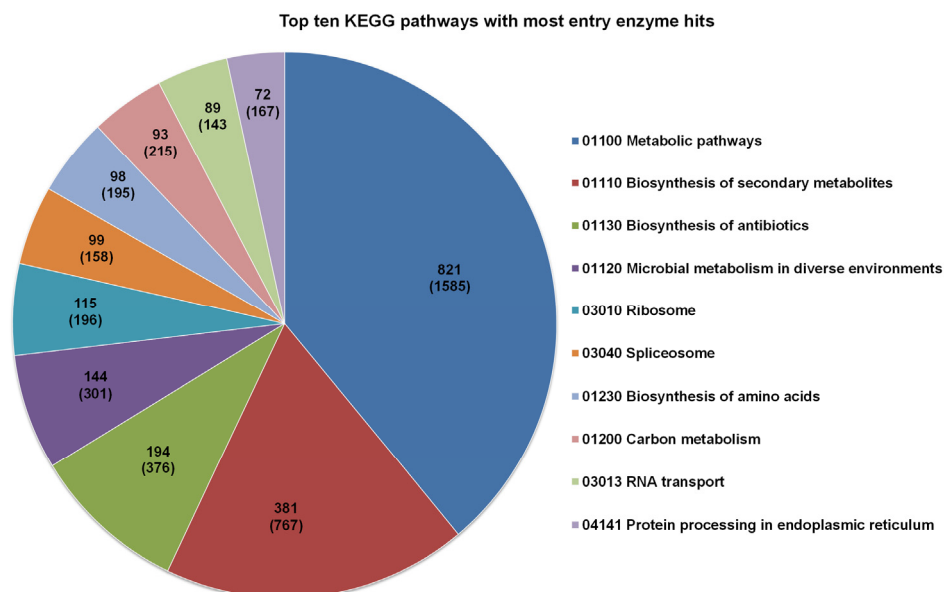
#### Analysis of KEGG metabolic pathway of unigenes

To identify unigenes involved in the KEGG metabolic pathway, the KEGG BlastKOALA online tool was used to assign KEGG Orthology (KO) number, by which a total of 6,553 unigenes of *P. chinensis* were assigned. Among them, 4,061 unigenes were mapped into 391 different KEGG metabolic pathways (Table 2), in which 2,561 unigenes were shown to be involved in more than one pathway. Major portions of unigenes (6,715) and pathways (138)

were found to be included into the pathways of metabolism category (Table 2).

Among all pathways, top four pathways with most entry enzyme hits were metabolic pathways with 821 entry enzymes, biosynthesis of secondary metabolites with 381, biosynthesis of antibiotics with 194, microbial metabolism in diverse environments with 144 entries, respectively (Fig. 6).

Biosynthesis of many secondary metabolites is tightly correlated to some of metabolic pathways including the pathways involved in the metabolism of terpenoid, the biosynthesis of phenylpropanoid, or flavonoid. Twenty-two unigenes were found for 16 entries of carotenoid biosynthesis pathway (Table 3). For phenylpropanoid biosynthesis pathway, 63 unigenes were found to encode enzymes for 17 entries, while 24 unigenes were found for 14 entries of the flavonoid biosynthetic pathway (Table 3; Supplementary Table S4).



**Fig. 6** Top 10 KEGG pathways with the most entry enzymes identified from *P. chinensis* transcriptome. Number for each pathway denotes total entry enzymes identified and the number in parenthesis indicates total unigenes encoding entry enzymes

**Table 3** KEGG metabolic pathways related to the biosynthesis of various medicinal metabolites

Metabolic pathways	KEGG map ID	Number of entry	No of unigenes*
<b><u>Metabolism of terpenoids</u></b>			
Carotenoid biosynthesis	00906	16	22
Sesquiterpenoid and triterpenoid biosynthesis	00909	5	8
Diterpenoid biosynthesis	00904	5	6
Monoterpenoid biosynthesis	00902	2	5
<b><u>Biosynthesis of phenylpropanoid and flavonoid</u></b>			
Phenylpropanoid biosynthesis	00940	17	63
Flavonoid biosynthesis	00941	14	24
Isoflavonoid biosynthesis	00943	1	2
Flavone and flavonol biosynthesis	00944	1	1

\* Number of unigenes indicates that they can encode enzymes for corresponding entry.

## Discussion

Medicinal plants are normally rich in traditional knowledge about medicinal usage, but there is very limited genetic information available for most traditional medicinal plants except for well-known medicinal plants. Medicinal plants are getting more interest to identify new metabolic compounds that possess important medicinal properties. In the absence of genomic information, however, it is very difficult to identify new lead molecules for pharmaceutical drug development from medicinal plants and to delve into how those molecules are synthesized in those plants. Therefore,

enrichment of genomic resources as well as genetic information is crucial for studying medicinal properties and for identifying potential lead molecules from unexplored medicinal plant species. As medicinal plants are getting more interest, increasing number of those species are getting sequenced by advanced sequencing technologies.

Advanced RNA-seq technology has spurred transcriptome analysis of increasing number of medicinal plants that have not been of interest previously. In this study, transcriptome *P. chinensis* was analyzed by RNA-seq technology. *De novo* assembly generated 18,524 unigenes (Table 1), of which 17,814 unigenes were annotated. The number of identified

unigenes in the present study is significantly lower than that from previous study in which 127,545 unigenes were identified from *P. chinensis* (Dong et al. 2016). However, compared to reports from other plant species, the number of unigenes reported by Dong et al. (2016) is very high, so further verification may be needed. On the other hand, the number of unigenes in our study is about 60% of the number of predicted genes from *P. vera* genome, in which 31,784 genes were reported by Zeng et al. (2019).

Many unigenes from the present study showed higher similarity to those of *Citrus sinensis*. Roughly, half of the annotated unigenes, 8,977 unigenes (48.5%), showed similarity to proteins from *Citrus sinensis* (Fig. 2), suggesting that *P. chinensis* is closely related to *C. sinensis*. Genome evolution analysis using gene families from *P. vera* showed a similar results among nine species from different genus, in which *P. chinensis* and *C. sinensis* showed that they diverged more recently among tested nine species (Zeng et al. 2019).

Many SSR markers have been developed mostly from *P. vera* (Zaloglu et al. 2015; Ziya Motalebipour et al. 2016), many of which were successful for cross-amplification among various *Pistacia* species (Zaloglu et al. 2015; Ziya Motalebipour et al. 2016). In this study, a total of 2,629 perfect SSRs were identified from *P. chinensis*, of which frequency was 162.5 per one Mbp (Supplementary Table S2). Frequency of tri-nucleotide SSRs was 144.9 per one Mbp (Fig. 3) and was the most abundant SSRs as reported in various other studies (Zhang et al. 2019; Bae et al. 2018; Eum et al. 2019; Kotwal et al. 2016). Repetitive sequence analysis showed that repetitive and low complexity sequences occupied 149,205 bp (0.92%) of the *P. chinensis* transcriptome, and the most abundant repetitive sequences were SSRs (Supplementary Table S1).

Various secondary metabolites possess important medicinal effect, and many metabolites are known to be derived from pathways involved in terpenoid metabolism or secondary metabolite biosynthesis pathways. A total of 4,061 unigenes were assigned into 391 different metabolic pathways through KEGG pathway analysis. Among the assigned unigenes, 131 genes were found to be involved in some of metabolite biosynthesis pathways including terpenoid, phenylpropanoid, and flavonoid (Table 3; Supplementary Table S4). Various genes involved in these pathways were identified from the present study as shown in Supplementary (Supplementary Table S4). Reconstruction of flavonoid and phenylpropanoid biosynthesis using the unigene information in Supplementary Table S4 revealed that most key genes involved in those pathways are well conserved in *P. chinensis* (Supplementary Fig. S1 and S2). Such genes include those that encode en-

zymes required for biosynthesis of kaempferol and quercetin known to have anti-cancer activity (Batra and Sharma 2013). Also, the biosynthetic pathway of luteoforol (Deoxyeucocyanidin) is well conserved, which is known to have an antimicrobial activity (Spinelli et al. 2005). Nevertheless, since *P. chinensis* contains various medicinal compounds having potentially medicinal effects (Liu et al. 2008; Nishimura et al. 2000; Rashed et al. 2016; Shi and Zuo 1992), our data will be invaluable to study how those metabolic compounds are synthesized.

## Conflict of Interests

The authors declare that there is no conflict of interests.

## Acknowledgment

This study was supported by KRIBB initiative program of Republic of Korea and 2017 research grant to JKN from Kangwon National University.

## Data Deposit

This transcriptome shotgun sequencing data was deposited at BioProject: PRJNA566127 in NCBI GenBank. SRA accession number is SRR10136265.

## References

- Akhtar N, Rashid A, Murad W, Bergmeier E (2013) Diversity and use of ethno-medicinal plants in the region of Swat, North Pakistan. *Journal of Ethnobiology and Ethnomedicine* 9 (1):25. doi:10.1186/1746-4269-9-25
- Bae DY, Eum SM, Lee SW, Paik JH, Kim SY, Park M, Lee C, Tran TB, Do VH, Heo JY, Seong ES, Kim IS, Choi KY, Hong JS, Ramekar RV, Choi S, Na JK (2018) Enrichment of genomic resources and identification of simple sequence repeats from medicinally important *Clausena excavata*. *3 Biotech* 8 (133):1–10. doi:doi.org/10.1007/s13205-018-1162-x
- Batra P, Sharma AK (2013) Anti-cancer potential of flavonoids: recent trends and future perspectives. *3 Biotech* 3 (6):439–459. doi:10.1007/s13205-013-0117-5
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30 (15): 2114–2120. doi:10.1093/bioinformatics/btu170
- Bozorgi M, Memariani Z, Mobli M, Salehi Surmaghi MH, Shams-Ardekani MR, Rahimi R (2013) Five *Pistacia* species (*P. vera*, *P. atlantica*, *P. terebinthus*, *P. khinjuk*, and *P. lentiscus*): a



- review of their traditional uses, phytochemistry, and pharmacology. *TheScientificWorldJournal* 2013:219815. doi:10.1155/2013/219815
- Dong S, Liu Y, Xiong B, Jiang X, Zhang Z (2016) Transcriptomic Analysis of a Potential Bioenergy Tree, *Pistacia chinensis* Bunge, and Identification of Candidate Genes Involved in the Biosynthesis of Oil. *BioEnergy Research* 9 (3):740–749. doi:10.1007/s12155-016-9716-4
- Eum SM, Kim S-Y, Hong JS, Roy NS, Choi S, Paik J, Lee SW, Tran TB, Do VH, Kim KS, Seong E-S, Park K-C, Yu CY, Eom SH, Choi K-Y, Kim J-H, Na J-K (2019) Transcriptome analysis and development of SSR markers of ethnobotanical plant *Sterculia lanceolata*. *Tree Genetics & Genomes* 15 (3):37. doi:10.1007/s11295-019-1348-3
- Kotwal S, Kaul S, Sharma P, Gupta M, Shankar R, Jain M, Dhar MK (2016) De Novo Transcriptome Analysis of Medicinally Important *Plantago ovata* Using RNA-Seq. *PloS one* 11 (3):e0150273. doi:10.1371/journal.pone.0150273
- Lee BY, Kim HS, Choi BS, Hwang DS, Choi AY, Han J, Won EJ, Choi IY, Lee SH, Om AS, Park HG, Lee JS (2015) RNA-seq based whole transcriptome analysis of the cyclopoid copepod *Paracyclops nana* focusing on xenobiotics metabolism. *Comparative biochemistry and physiology Part D, Genomics & proteomics* 15:12–19. doi:10.1016/j.cbd.2015.04.002
- Liu JJ, Geng CA, Liu XK (2008) A new pyrrolidone derivative from *Pistacia chinensis*. *Chinese Chemical Letters* 19 (1):65–67. doi:https://doi.org/10.1016/j.ccl.2007.10.037
- Loke KK, Rahnamaie-Tajadod R, Yeoh CC, Goh HH, Mohamed-Hussein ZA, Mohd Noor N, Zainal Z, Ismail I (2016) RNA-seq analysis for secondary metabolite pathway gene discovery in *Polygonum minus*. *Genomics data* 7:12–13. doi:10.1016/j.gdata.2015.11.003
- Nishimura S, Taki M, Takaishi S, Iijima Y, Akiyama T (2000) Structures of 4-aryl-coumarin (neoflavone) dimers isolated from *Pistacia chinensis* BUNGE and their estrogen-like activity. *Chemical & pharmaceutical bulletin* 48 (4):505–508. doi:10.1248/cpb.48.505
- Rai A, Yamazaki M, Takahashi H, Nakamura M, Kojoma M, Suzuki H, Saito K (2016) RNA-seq Transcriptome Analysis of *Panax japonicus*, and Its Comparison with Other *Panax* Species to Identify Potential Genes Involved in the Saponins Biosynthesis. *Frontiers in plant science* 7:481. doi:10.3389/fpls.2016.00481
- Rashed K, Said A, Abdo A, Selim S (2016) Antimicrobial activity and chemical composition of *Pistacia chinensis* Bunge leaves. *International Food Research Journal* 23(1): 316–321 (2016) 23 (1):316–321
- Rauf A, Patel S, Uddin G, Siddiqui BS, Ahmad B, Muhammad N, Mabkhot YN, Hadda TB (2017) Phytochemical, ethnomedicinal uses and pharmacological profile of genus *Pistacia*. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie* 86:393–404. doi:10.1016/j.biopha.2016.12.017
- Shi Q, Zuo C (1992) Chemical components of the leaves of *Pistacia chinensis* Bge. *Zhongguo Zhong Yao Za Zhi* 17:422–423
- Spinelli F, Speakman J-B, Rademacher W, Halbwirth H, Stich K, Costa G (2005) Luteoforol, a flavan 4-ol, is induced in pome fruits by prohexadione-calcium and shows phytoalexin-like properties against *Erwinia amylovora* and other plant pathogens. *European Journal of Plant Pathology* 112 (2):133–142. doi:10.1007/s10658-005-2192-x
- Tang M, Zhang P, Zhang L, Li M, Wu L (2012) A potential bioenergy tree: *Pistacia chinensis* Bunge. *Energy Procedia* 16:737–746
- Zaloglu S, Kafkas S, Doğan Y, Güney M (2015) Development and characterization of SSR markers from pistachio (*Pistacia vera* L.) and their transferability to eight *Pistacia* species. *Scientia Horticulturae* 189:94–103. doi:https://doi.org/10.1016/j.scienta.2015.04.006
- Zeng L, Tu X-L, Dai H, Han F-M, Lu B-S, Wang M-S, Nanaei HA, Tajabadi-pour A, Mansouri M, Li X-L, Ji L-L, Irwin DM, Zhou H, Liu M, Zheng H-K, Esmailzadeh A, Wu D-D (2019) Whole genomes and transcriptomes reveal adaptation and domestication of pistachio. *Genome Biology* 20 (1):79. doi:10.1186/s13059-019-1686-3
- Zhang Z, Xie W, Zhao Y, Zhang J, Wang N, Ntakirutimana F, Yan J, Wang Y (2019) EST-SSR marker development based on RNA-sequencing of *E. sibiricus* and its application for phylogenetic relationships analysis of seventeen *Elymus* species. *BMC Plant Biology* 19 (1):235. doi:10.1186/s12870-019-1825-8
- Ziya Motalebi-pour E, Kafkas S, Khodaei-aminjan M, Çoban N, Gözel H (2016) Genome survey of pistachio (*Pistacia vera* L.) by next generation sequencing: Development of novel SSR markers and genetic diversity in *Pistacia* species. *BMC genomics* 17 (1): 998–998. doi:10.1186/s12864-016-3359-x