Bangladesh J. Bot. **52**(2): 243-252, 2023 (June) DOI: https://doi.org/10.3329/bjb.v52i2.66945

ANALYSIS OF DIFFERENTIALLY EXPRESSED UNIGENES INVOLVED IN THE DEVELOPMENT OF KOREAN PINE (*PINUS KORAIENSIS* SIEB. ET ZUCC.)

DAN HOU, LI ZHANG[1], JINNING WANG[2], JINQUAN LI[2], LEI ZHANG[*] AND HANGUO ZHANG[*]

State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin 150040, China

Keywords: Korean pine, Transcriptome, GO enrichment, KEGG enrichment, Differentially expressed unigenes

Abstract

Korean pine is a monoecism gymnosperm that predominantly exists as both female and male. Due to the absence of a model organism reference genome, high-throughput sequencing was used to analyze the candidate unigenes of tissue and organ development. This was performed in the tissues of microstrobilli (Ms), needles (Ne), and mixed buds (Mb) of the Korean pine to mine for genes related to bud differentiation and flowering. Overall, 26.32 Gb of clean data comprising 26972 annotated unigenes were obtained. Out of these, 26323 (97.59%) were annotated in the NR database based on sequence homology. Twenty-four GO terms and 30 candidate unigenes related to the formation or development of flowers, needles, buds, and pollens were found by GO enrichment analysis.

Introduction

Pinus koraiensis, commonly known as the Korean pine, an evergreen conifer belonging to Pinaceae, is a native tree species in the forests of Northeast China, and a grade II national key preserved wild plant of China (Lim 2013). Because of its great ecological and economic value, *P. koraiensis* has gained more attention through the years. In particular, Korean pine seeds have become more popular in pine seed products due to their medical uses and nutritional value (Liang *et al*. 2019). Pine nuts not only contain large amounts of crude protein, crude fat, total sugar, and crude fiber, but also some vitamins, minerals, and trace elements (Zadernowski *et al*. 2009). However, Korean pine is a monoecious gymnosperm that predominantly exists as both female and male, resulting in varying strobilus production among different varieties. Currently, research on particular methods to promote seed yields of this monoecious gymnosperm is mostly macroscopic. Some seed orchards did successfully construct clones with relatively high seed production (Cui *et al.* 2005, Zhang 2015). However, the best methods for promoting flowers and seed yields for different materials are not unified, which indicates that the most scientific and effective solution has not been found because people do not understand the mechanism of flower bud differentiation or flowering in Korean pine. In order to develop a more scientific and reasonable plan, one should understand the molecular mechanism of flower bud differentiation, which can avoid the influence of external environment and genetic background, and truly clarify which genes play a role in the process of flower bud differentiation, and which factors regulate these genes. This provides a theoretical basis for the subsequent promotion of flowers and fruits in seed orchards and scientific management.

*Author for correspondence: <hanguozhang1@sina.com>, <zhanglei@nefu.edu.cn>. [1]Medical School, Shandong Xiandai University, Jinan 250000, China. [2]Bohai Forest Seed orchard in Ning'an City, Mudanjiang 157400, China.

RNA-seq is a recently developed approach for profiling transcriptomes and is now widely used to analyze gene expression, discover novel transcripts, and decipher the molecular mechanisms for growth and development. It is cost-effective, highly accurate, and has a large dynamic range (Hou *et al*. 2018). In particular, it has been a powerful tool for species that lack reference genome information (Collins *et al*. 2008). Until recently, transcriptome sequencing has been used to study the unigenes related to cold stresses (Wang *et al*. 2020) and fatty acid synthesis regulation (Du *et al*. 2017) at *P. koraiensis*. However, only a few reports on unigenes related to flower bud and foliar bud differentiation in the Korean pine and other gymnosperms are available.

In the present work, high-throughput sequencing approaches were used to obtain transcriptomic data from the Korean pine. Differential gene expression among the tissues of microstrobilli (Ms), needles (Ne), and buds (Mb) were investigated to reveal differential regulation of key gene ontology (GO) terms and pathways.

Materials and Methods

Fresh samples were collected from the clone (NB66) in the Bohai Korean pine seed orchard (44°01'50"-44°04'02" N; 128°50'23"–128°54'41" E) on the east slope of Zhangguangcai Ridge, southwest of Ning'an City, Heilongjiang Province, Northeast China. The altitude, average slope, average frost-free season, mean annual precipitation, mean annual temperature, active accumulated temperature, and extreme low temperature of the test site are 370–490 m, 15°, 134 days, 670 mm, 3.5, 2400 and -40.1℃, respectively. The seed orchard was grafted and built in 1979, with plant spacing at a dislocated 4 x 6 sq.m arrangement. In 2018, three fresh tissue samples were collected from microstrobilli (Ms), needles (Ne), and mixed buds (Mb), and Ms and Ne were sampled in mid-June, Mb was sampled in mid-September. These samples were immediately frozen in liquid nitrogen and stored at -80 ℃ for RNA extraction. RNA from each of the three materials (Ms, Ne, Mb) according to the manufacturer's instructions (Trizol Reagent, Invitrogen, USA) was extracted. The purity, concentration, and integrity of RNA samples were tested to ensure good quality. Sequencing libraries were generated using the NEBNext®Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following the manufacturer's index. PCR was performed and clusters were created following the library fragments purified with the AMPure XP system (Beckman Coulter, Beverly, USA). Sequencing was processed by an Illumina Hiseq 2000 platform using paired-end reads.

Cleaned data (reads) were obtained by removing reads containing adapter sequences, poly-N, and low-quality reads from the raw data. Trinity software was used for the mixed assembly of sample data to obtain transcripts (Grabherr *et al.* 2011). The longest transcripts were selected as the unigene sequence for downstream analyses. Unigene sequences were compared among the following databases: NR (NCBI non-redundant protein sequences), Swiss-Prot, COG (Clusters of Orthologous Groups of proteins), euKaryotic Orthologous Groups (KOG), GO, eggNOG and Kyoto Encyclopedia of Unigenes and Genomes (KEGG) using BLAST.

Gene expression levels were estimated by RSEM (Li and Dewey 2011). According to comparison information, Fragments Per Kilobase of transcript per Million mapped reads (FPKM) was used to reflect the abundance of unigene expression. Differential expression analysis was performed using the EBSeq R package. FDR < 0.01 and $|\log_2(\text{fold change})| > 1$ as the threshold for significantly differential expression were selected. Gene ontology (GO) enrichment analysis of the differentially expressed unigenes (DEGs) was implemented using the topGO R package.

Results and Discussion

A total of 26.32 Gb of cleaned data were obtained. The number of reads from Ne and Mb was approximately equal, but less than that of Ms. However, the number of annotated unigenes in the study was lower than in some other studies (Du *et al.* 2017, Wang *et al.* 2020), which might be due to the materials used and sample volume. The ratio of mapped reads, GC content, and Q30 were around 83, 45, and 94%, respectively (Table 1). In addition, 138112 contigs were further assembled into 48156 unigenes, of which the average length was 1099.97 nt, with an N50 of 1802 nt. Unigenes with a sequence length greater than 2000 nt accounted for 15.84% of the sample, while those with a sequence length between 1000 and 2000 nt accounted for 19.32%.

Table 1. Quality of cleaned data.

Sample	Reads number	Mapped reads	Unique mapped reads	Multi-mapped reads	Base number	GC content	Q30
Ms	24,180,792	19,628,340	7,620,410	12,007,930	7,206,008,582	45.69	94.30
Ne	21,069,234	17,359,238	7,083,641	10,275,597	6,284,003,858	45.92	94.92
Mb	21,455,946	17,924,033	7,105,629	10,818,404	6,414,568,887	45.08	94.55

Reads number Annotated: the total number of paired-end reads; Mapped reads Annotated: the total number of mapped reads; Unique mapped reads Annotated: the number of uniquely mapped reads; Multi-mapped reads Annotated: the number of multi-mapped reads; Base number Annotated: total base number of cleaned data; GC content Annotated: percentage of G and C bases in the cleaned data; Q30 Annotated: percentage of bases with a quality value greater than or equal to 30.

In the present study, 26972 unigenes with annotation information were obtained. Approximately 26323 (97.59%) were annotated in the NR database based on sequence homology, while annotations in other databases ranged from 32.77 to 87.55%. Detailed annotations for the Korean pine unigenes are summarized in Table 2.

Table 2. Annotations of Korean pine unigenes against public databases.

Database	Number	Percentage (%)	300 <= length < 1000	length >= 1000
Nr	26323	97.59	11669	14654
Swissprot	17859	66.21	7077	10782
COG	8840	32.77	3019	5924
KOG	14855	55.08	6226	8629
GO	15735	58.34	6621	9114
eggNOG	23613	87.55	9709	13904
KEGG	9528	35.33	3667	5861
Pfam	19542	72.45	7271	12271
All	26972	100	12222	14750

According to the E-value distribution of top hits from the NR database, 60.61% of the matched sequences showed strong homology (<1e-50) (Fig. 1A), while 60.79% of the sequences had a similarity rate > 60% (Fig. 1B). Species distribution of the top BLAST hits showed that these unigenes had the greatest number of matches with unigenes of *Picea sitchensis* (35.58%; Fig. 1C), indicating that the transcriptome of Korean pine was more closely related to that of *P. sitchensis* than to other plant genomes that are currently present in public databases.

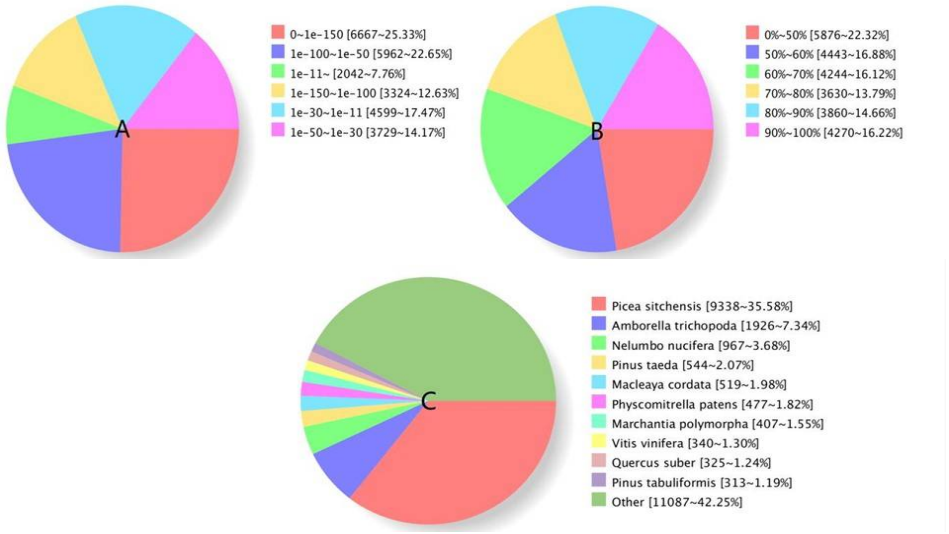

Fig. 1. Nr Evalue, Identity, and Homologous Species Distribution. (A) Evalue Distribution. (B) Identity Distribution. (C) Homologous Species Distribution.

The GO database is a structured biological annotation system for genes and proteins and provides detailed properties of unigenes and their products. Gene ontology system alignment showed that 15735 unigenes were classified into 49 main functional groups, including cellular component (15), molecular function (14), and biological process (20) (Fig. 2).

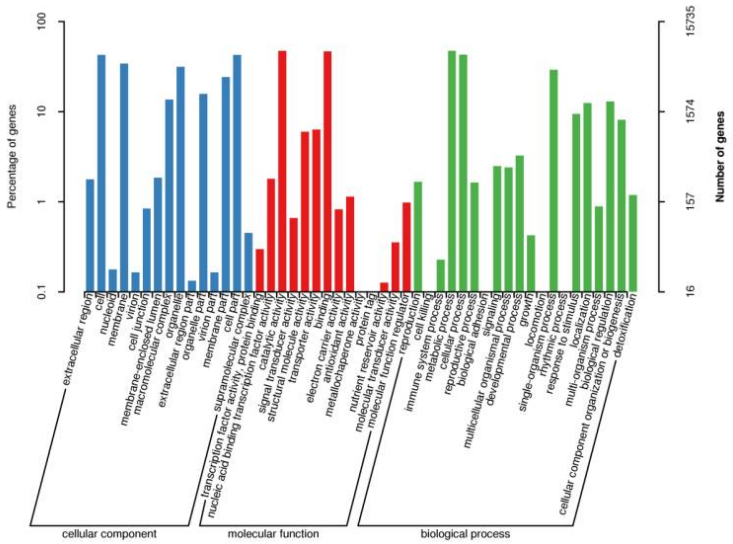

Fig. 2. Gene ontology functional classification of the Korean pine transcriptome.

Kyoto Encyclopedia of Unigenes and Genomes (KEGG) analysis was used to understand the specific processes, gene functions, and gene interactions involved at the transcriptome level. In the present study, 9528 unigenes were mapped to 129 predicted metabolic pathways through the

KEGG database, with 27 pathways having more than 100 identified genes. The pathways that ranked among the top 50 are listed in Fig. 3, all of which were classified into five physiological and biochemical processes, with metabolism being dominant.

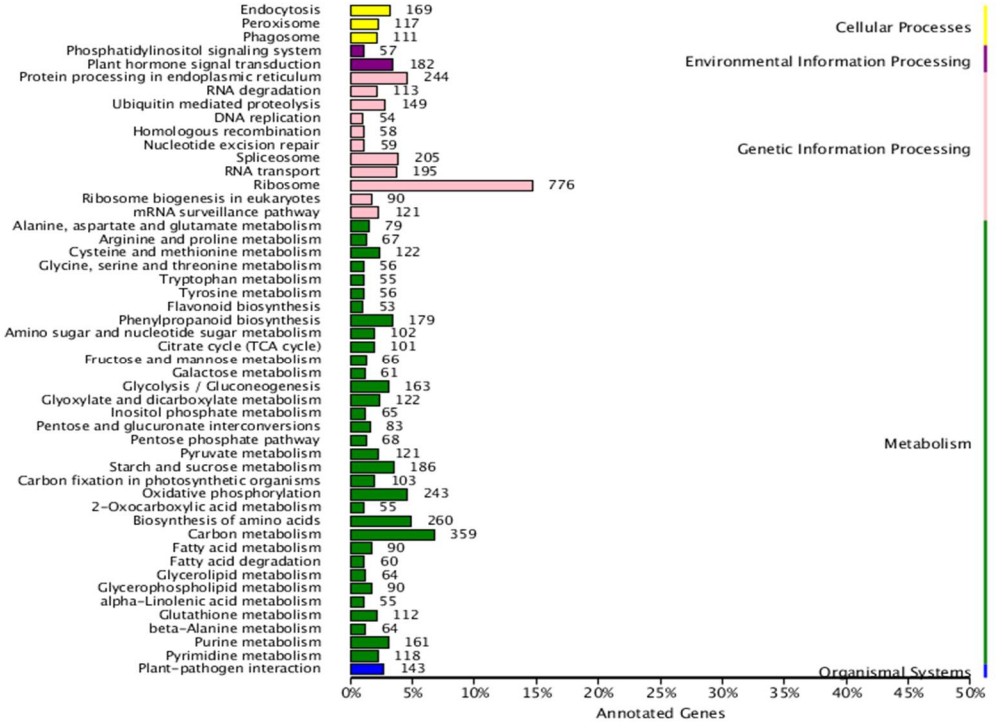

Fig. 3. Kyoto Encyclopedia of Unigenes and Genomes classification of the Korean pine transcriptome.

Expression levels of unigenes were determined by aligning the RNA-seq reads from each library to the assembly. Comparisons among unigene expression identified 7373, 8320, and 3741 unigenes that were differentially expressed in Ms vs. Mb, Ne vs. Mb , and Ms vs. Ne, respectively. These suggest that unigenes involved in bud development were significantly different from those of microstrobilli and needles.

Difference was observed in the number of unigenes annotated against different databases in the three groups: Ms vs. Mb, Ne vs. Mb, and Ms vs. Ne. The numbers of annotated unigenes with up-regulated expression among the three groups were 2303, 2744, and 1691, respectively, while they were 3241, 3658, and 1300, respectively, for those that were down-regulated. The NR database contained the most numbers of annotations for all three groups, and the numbers of annotated unigenes in all datasets combined were 558, 643, and 351, respectively. Detailed annotation results for the Korean pine unigenes are summarized in Table 3.

In order to study the overlap among DEGs in different samples, a Venn diagram of DEGs identified from Ms vs. Mb, Ne vs. Mb , and Ms vs. Ne was made. There were 3224 DEGs that differed only in Ms vs. Mb and Ne vs. Mb, which potentially play important roles in bud development and formation. Meanwhile, there were 799 shared DEGs among the three groups, the functions for which also require further research and verification.

Table 3. Differentially expressed unigenes annotations from different databases among the three groups.

DEG set	Annotated	COG	GO	KEGG	KOG	Pfam	Swiss-Prot	eggNOG	Nr
Ms vs. Mb	5544	1887	3244	1802	2224	4149	3957	4681	5459
Ne vs. Mb	6402	2195	3743	2078	2582	4804	4543	5432	6310
Ms vs. Ne	2991	1121	1839	1036	1126	2336	2190	2573	2928

To further investigate the function of these DEGs, the 799 shared DEGs among the three groups and the 3224 DEGs identified in Ms vs. Mb and Ne vs. Mb were analyzed by GO enrichment. Twenty-four GO terms related to the formation or development of flowers, needles, buds, and pollens are listed in Table 4. From this analysis, it can be observed that some DEGs fell under one specific GO term, while some were classified under several GO terms. Meanwhile, some unigenes demonstrated significant differences among Ms, Mb , and Ne, suggesting that these unigenes may be related to the formation or development of flowers, needles, buds, and pollens.

Results from a Blast sequence alignment of 30 candidate unigenes obtained from GO enrichment are listed in Table 5. The annotation information for some unigenes were clearly defined, while some were unknown or simply predicted. Therefore, the functions of these unigenes require further investigations.

Twenty-four GO terms and 30 candidate unigenes related to the formation or development of flowers, needles, buds, and pollens were identified by GO enrichment analysis. Some DEGs were clearly annotated by the NR database, including "protein arginine N-methyltransferase", "*B3* domain-containing protein", "*LOB* domain-containing protein", "*GA20ox1*", "*SCL3*", "cytochrome *P450*", "*MADS*", "*SNF2*", "expansin *B1*", "*LRR* receptor-like serine/threonine-protein kinase *ERL1*", "*FT/TFL1-like 2*", and "*ABC* transporter B family member 19-like". Wang *et al.* (2012) determined that *B3* domain-containing proteins played a central role in plant life, from embryogenesis to seed maturation, and even in dormancy. However, the gene involved in *B3* domain-containing proteins in this study expressed the highest in needles but the lowest in buds. The study on *Taxus* by Li *et al*. (2015) emphasized that lateral organ boundaries (*LOB*) domain transcription factors played a potential role in plant secondary growth. Meanwhile, the gene involved in *LOB* had the lowest expression level in buds in this study, which may be beneficial to leaf morphogenesis. Previous studied on the Arabidopsis life cycle suggested that *GA20ox1* and *GA20ox2* act redundantly to promote hypocotyl and internode elongation, flowering time, elongation of anther filaments, the number of seeds that develop per silique, as well as the elongation of siliques (Rieu *et al.* 2008). Similar to obtained results by Rieu's study, the expression level of gene related to *GA20ox1* was the highest in buds, which may be related to flower development. Weng *et al.* (2020) found that protein synthesis and flavonoid biosynthesis were regulated by *SCL3* in the *Arabidopsis thaliana* root system. However, in the present study it was found that the expression of gene like *SCL3* was the highest in buds, which may be involved in leaf development. There were many studies on the role for "*MADs*", which involved flower development (Wang 2019), fruit ripening (Li 2017), and cold resistance (Zhang 2017). The expression of gene like *MADS* was high in conelet and buds, but low in needle, which indicated that the gene might be related to flower development. Arabidopsis contained 41 *SNF2* family proteins that fall into 18 subfamilies (Knizewski *et al.* 2008), who found a gene may also belong to *SNF2* family, which was involved in regulating flower development. Muthusamy *et al.* (2020) determined that expansin-like *B1* in *Brassica rapa* is involved in root development, drought stress

Table 4. Gene Ontology enrichment analysis of DEGs.

GO term	GO description	Gene ID
GO:0009908	flower development	c65084.graph_c0, c66595.graph_c0, c74336.graph_c0, c75466.graph_c0, c76576.graph_c0
GO:0009909	Regulation of flower development	c63914.graph_c0, c81308.graph_c0
GO:0009911	Positive regulation of flower development	c80688.graph_c1
GO:0048451	Petal formation	c69669.graph_c0
GO:0048281	Inflorescence morphogenesis	c64220.graph_c0
GO:0048497	Maintenance of floral organ identity	c71026.graph_c0
GO:0048573	Photoperiodism, flowering	c81308.graph_c0
GO:0009965	Leaf morphogenesis	c64220.graph_c0, c67160.graph_c0
GO:0048366	Leaf development	c65084.graph_c0, c69669.graph_c0, c74875.graph_c0, c79094.graph_c0
GO:0010016	Shoot system morphogenesis	c50674.graph_c0, c79103.graph_c2
GO:0048856	Anatomical structure development	c76996.graph_c0, c77632.graph_c1, c82232.graph_c0
GO:0048827	Phyllome development	c74957.graph_c0
GO:0003006	Developmental process involved in reproduction	c77632.graph_c1
GO:0007126	Meiotic cell cycle	c79398.graph_c0
GO:0090698	Post-embryonic plant morphogenesis	c79103.graph_c2
GO:0009791	Post-embryonic development	c74957.graph_c0, c79103.graph_c2, c80265.graph_c0
GO:0019953	Sexual reproduction	c74820.graph_c0, c78890.graph_c0
GO:0009846	Pollen germination	c69673.graph_c0, c76560.graph_c0
GO:0009860	Pollen tube growth	c43601.graph_c1, c69673.graph_c0, c76560.graph_c0
GO:0048544	Recognition of pollen	c78280.graph_c1
GO:0010584	Pollen exine formation	c80688.graph_c1
GO:0048232	Male gamete generation	c72674.graph_c0
GO:0048443	Stamen development	c82278.graph_c0
GO:0048657	Anther wall tapetum cell differentiation	c80688.graph_c1

response, and seed germination, while it was found that the gene may be related to sexual reproduction in the present study. *FT/TFL1* was directly or indirectly related to seed germination (Yu *et al.* 2019), bud burst (Carneros *et al.* 2017), and flower bud formation (Avia *et al.* 2014). More or less similar results showed that *FT/TFL1* was the highest expressed in buds, which might be involved in regulation of flower development, photoperiodism, and flowering.

Table 5. NR annotation of the 30 DEGs.

Gene ID	Nr Annotation	Reference species
c43601.graph_c1	Unknown	*Picea sitchensis*
c50674.graph_c0	Unknown	*Picea sitchensis*
c63914.graph_c0	PREDICTED: protein arginine N-methyltransferase 1.5	Nelumbo nucifera
c64220.graph_c0	Unknown	Picea sitchensis
c65084.graph_c0	PREDICTED: B3 domain-containing protein Os03g0120900-like	Nelumbo nucifera
c66595.graph_c0	Unknown	Picea sitchensis
c67160.graph_c0	LOB domain-containing protein 12-like	Amborella trichopoda
c69669.graph_c0	Unknown	Picea sitchensis
c69673.graph_c0	Unknown	Picea sitchensis
c71026.graph_c0	Homeobox 1	Picea abies
c72674.graph_c0	Unknown	Picea sitchensis
c74336.graph_c0	GA20ox1, partial	Pinus tabuliformis
c74820.graph_c0	Unnamed protein product	Coffea canephora
c74875.graph_c0	SCL3, partial	Pinus radiata
c74957.graph_c0	Cytochrome P450 85A1	Setaria italica
c75466.graph_c0	MADS-box transcription factor GbMADS5	Ginkgo biloba
c76560.graph_c0	Unknown	Picea sitchensis
c76576.graph_c0	SNF2-related	Macleaya cordata
c76996.graph_c0	Protein kinase domain	Macleaya cordata
c77632.graph_c1	Predicted protein	Physcomitrella patens
c78280.graph_c1	Unnamed protein product, partial	Vitis vinifera
c78890.graph_c0	Expansin B1	Pinus radiata
c79094.graph_c0	Protein kinase domain	Macleaya cordata
c79103.graph_c2	PREDICTED: LRR receptor-like serine/threonine-protein kinase ERL1	Phoenix dactylifera
c79398.graph_c0	Serine/threonine-protein kinase ATR isoform X1	Amborella trichopoda
c80265.graph_c0	Hypothetical protein AXG93_1130s1520	Marchantia polymorpha subsp. *Ruderalis*
c80688.graph_c1	Cytochrome P450 90B1	Apostasia shenzhenica
c81308.graph_c0	FT/TFL1-like 2	Pinus sylvestris
c82232.graph_c0	Unknown	Picea sitchensis
c82278.graph_c0	PREDICTED: ABC transporter B family member 19-like	Malus domestica

In the present study 26.32 Gb of clean data comprising 26972 annotated unigenes were obtained. Twenty-four GO terms and 30 candidate unigenes related to the formation or development of flowers, needles, buds, and pollens were found. These sequence information was deposited to the NCBI SRA database (accession no. BankIt2412979). The study aided in a better understanding of the molecular mechanisms for flower bud and foliar bud differentiation, and can provide a theoretical basis for the molecular breeding and functional genomic studies on the Korean pine.

Acknowledgments

The authors are thankful to Applied Technology Research and Development Program of Heilongjiang Province (GA19B201-3); Heilongjiang Touyan Innovation Team Program for financial assistance.

References

Avia K, Karkkainen K, Lagercrantz U and Savolainen O 2014. Association of *flowering locus t/terminal flower 1*-like gene *FTL2* expression with growth rhythm in Scots pine (*Pinus sylvestris*). New Phytol. **204**: 159-170.

Carneros E, Yakovlev I, Viejo M, Olsen JE and Fossdal CG 2017. The epigenetic memory of temperature during embryogenesis modifies the expression of bud burst-related genes in Norway spruce epitypes. Planta. **246**: 553-566.

Collins LJ, Biggs PJ, Voelckel C and Joly S 2008. An approach to transcriptome analysis of non-model organisms using short-read sequences. Genome Inform. **21**: 3-14.

Cui BL, Yang JM, Zheng H and Yang MS 2005. Advances on the yield in seed orchard of Chinese conifers. Hebei J. For. Orchard Res. **20**: 120-123+137.

Du J, Zhang Z, Zhang HG and Tang JH 2017. EST–SSR marker development and transcriptome sequencing analysis of different tissues of Korean pine (*Pinus koraiensis Sieb. et Zucc*.), BIOTECHNOL. BIOTEC. EQ. **31**: 679-689.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, Chen ZH, Mauceli E, Hacohen N, Gnirke A, Rhind N, Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N and Regev A 2011. Full-length transcriptome assembly from RNA Seq data without a reference genome. Nat. Biotechnol. **29**: 644-652.

Hou XG, Guo Q, Wei WQ, Guo LL, Guo DL and Zhang L 2018. Screening of genes related to early and late flowering in tree peony based on bulked segregant RNA sequencing and verification by quantitative real-time PCR. Molecules. **23**: 689.

Knizewski L, Ginalski K and Jerzmanowski A 2008. Snf2 proteins in plants: gene silencing and beyond. Trends Plant Sci. **13**: 557-565.

Li B and Dewey CN 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. Bmc Bioinform. **12**: 323.

Li S, 2017. Functional identification of the key transcriptional factor MADS-RIN in tomato fruit. PhD dissertation, Food Biotechnology, China Agricul. Univ., Beijing, 122p.

Li YY, Yang YF, Wang JQ, Wang S, Liu HW and Qiu DY 2015. Cloning and Expression of Lateral Organ Boundaries Domain Genes (*TcLBDs*) in *Taxus chinensis*. Scien. Silvae Sinic. **51**: 126-133.

Liang DY, Wang BY, Song SL, Wang JY, Wang LF, Wang Q, Ren XB and Zhao XY 2019. Analysis of genetic effects on a complete diallel cross test of *Pinus koraiensis*. Euphytica. **5**: 92-103.

Lim TK, 2013. Edible medicinal and non-medicinal plants. Springer, Netherlands, 297-303 p.

Muthusamy M, Kim JY, Yoon EK, Kim JA and Lee SI 2020. *BrEXLB1*, a *Brassica rapa* Expansin-Like B1 Gene Is Associated with Root Development, Drought Stress Response, and Seed Germination. Genes. **11**: 404.

Rieu I, Ruiz-Rivero O, Fernandez-Garcia N, Griffiths J, Powers SJ, Gong F, Linhartova T, Eriksson S, Nilsson O, Thomas SG, Phillips AL and Hedden P 2008. The gibberellin biosynthetic genes At*GA20ox1* and At*GA20ox2* act, partially redundantly, to promote growth and development throughout the Arabidopsis life cycle. Plant J. **53**: 488-504.

Wang F, Chen S, Liang DY, Qu GZ, Chen S and Zhao XY 2020. Transcriptomic analyses of *Pinus koraiensis* under different cold stresses. Bmc Genomics. **21**: 10.

Wang X, 2019. Identification and evolution studies of *MADS*-box gene family related to Jute, Yam and Narrow Lupin Flowering. MA thesis, Bioengineering, South China Univ. Technol., Guangzhou, 66 p.

Wang YJ, Deng DX, Zhang R, Wang SX, Bian YL and Yin ZT 2012. Systematic analysis of plant-specific B3 domain-containing proteins based on the genome resources of 11 sequenced species. Mol. Biol. Rep. **39**: 6267-6282.

Weng CY, Zhu MH, Liu ZQ and Zheng YG 2020. Integrated bioinformatics analyses identified *SCL3*-induced regulatory network in *Arabidopsis thaliana* roots. Biotechnol. Lett. **42**: 1019-1033.

Yu XL, Liu H, Sang N, Li YF, Zhang TT, Sun J and Huang XZ 2019. Identification of cotton *MOTHER OF FT AND TFL1* homologs, *GhMFT1* and *GhMFT2*, involved in seed germination. PLOS ONE. **14**: e0215771.

Zadernowski R, Naczk M and Czaplicki S 2009. Chemical composition of *Pinus sibirica* nut oils. Eur. J. Lipid Sci. Tech. **111**: 698-704.

Zhang Q, 2017. Cloning and functional characterization of *MsMADs1* gene from ALFALFA (*Medicago sativa L.*). MA thesis, Genetics, Harbin Normal University, Harbin, 77 p.

Zhang Z, 2015. Seed morphology and nutrient composition variation in Korean pine clonal. PhD dissertation, Forest Tree Genetic and Breeding, Northeast For. Univ., Harbin, 132 p.

(Manuscript received on 4 November, 2021; revised on 20 May, 2023)