BMC Genomic Data

DATA NOTE

Open Access

Transcriptome profiling across 11 different tissues in *Pisum sativum*

Ru-Zhi Li[1,2†], Guan Li[2†], Xiao-Yan Zhang[3], Lei-Lei Li[2], Zhi-Wei Wang[2], Feng-Jing Song[3], Kai-Hua Jia[2], Na-Na Li[2*] and Peng-Fei Chu[1*]

Abstract

Objectives Peas (*Pisum sativum*) are vital for human nutrition and have significantly contributed to the understanding of Mendelian inheritance laws. In this study, we performed transcriptome sequencing on 11 distinct pea tissues, offering an extensive gene expression dataset. This study not only provides valuable insights into the gene expression patterns across various pea tissues but also lays the foundation for future research aimed at understanding the molecular mechanisms underlying pea growth, development, and response to environmental stimuli.

Data description In this study, we conducted transcriptome sequencing on tissue samples from 11 distinct pea plants, each with three biological replicates. This approach yielded a comprehensive RNA-seq dataset, abundant in transcriptomic information. Through principal component analysis (PCA) and gene ontology (GO) enrichment analysis, we identified significant variations in gene expression across different samples. This valuable transcriptomic resource enhances our understanding of gene expression in diverse pea tissues and provides new strategies, along with potential candidate genes, for the genetic improvement of peas.

Keywords *Pisum sativum*, Transcriptomics, Gene expression, Pea tissues

Objective

Pea (*Pisum sativum*), an annual cool-season crop [1], has long been an essential part of the human diet due to its high starch, protein, and nutrients content [2]. Recent research has primarily focused on the nutritional value of peas, emphasizing their health benefits as a high-quality plant protein source rich in fiber, vitamins, and minerals [2–4]. In this study, we present transcriptomic data from 11 pea samples, and through pairwise tissue comparisons, provide an in-depth analysis of the tissue-specific nature of differentially expressed genes.

In this study, we utilized the DNBSEQ platform to perform comprehensive transcriptome sequencing across various pea tissues. Through RNA sequencing of 11 pea samples (Fig. 1), we generated a substantial dataset and identified differentially expressed genes (DEGs).

†Ru-Zhi Li and Guan Li contributed equally to this work.

*Correspondence:
Na-Na Li
nanali_saas@163.com
Peng-Fei Chu
chupengfei@lcu.edu.cn
[1]College of Agriculture and Biology, Liaocheng University, Liaocheng 252000, China
[2]Shandong International Joint Laboratory of Agricultural Germplasm Resources Innovation, Institute of Crop Germplasm Resources, National Saline-Alkali Tolerant Crop Germplasm Resources Nursery (Dongying), Shandong Academy of Agricultural Sciences, Jinan 250100, China
[3]Qingdao Academy of Agricultural Sciences, Qingdao 266100, China

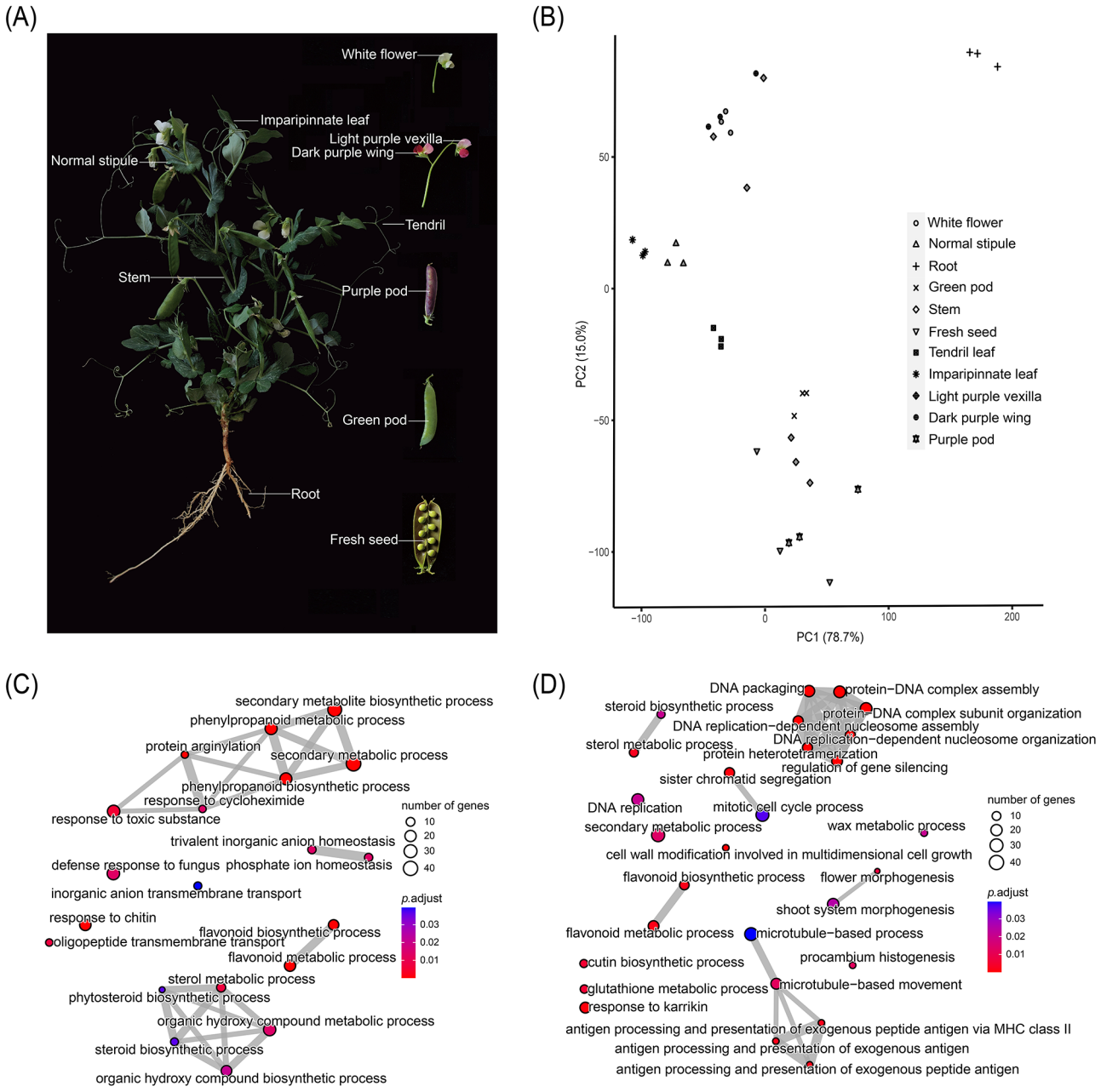

Fig. 1 Transcriptomic landscape of 11 distinct pea samples. (**A**) Image of the entire pea plant along with the specific tissues collected in this study. (**B**) Principal component analysis of gene expression from the 11 samples. (**C**) Enrichment analysis of upregulated expressed genes comparing light purple vexilla with white flowers. (**D**) Enrichment analysis of upregulated expressed genes comparing purple pods with green pods

This analysis revealed distinct gene expression patterns among the tissues, offering fundamental insights into their unique biological functions.

Data description

Sample collection

In this study, purple flowers and purple pods were collected from the cultivar Yunwan 127, while all other tissues were obtained from the cultivar Zhongwan 6. Prior to sample collection, the plants were not exposed to significant artificial or natural stress, irrigation was managed through natural rainfall, supplemented by manual watering as needed. Samples were collected two days after flowering, at 10:00 AM on April 16, 2024, under clear skies and a temperature of 21 °C, from a field located at 36°42'12.45" N, 117°04'49.39" E. A total of 11 distinct tissue samples were collected in triplicate, with each replicate consisting of a single plant. The samples included white flower, dark purple wing, light purple vexilla, normal stipule, root, stem, fresh seed, tendril leaf,

Table 1 Overview of data files/data sets

Label	Name of data file/data set	File types (file extension)	Data repository and identifier (DOI or accession number)
Data file 1	Transcriptomic Profile of 11 Pea Tissues	Portable document format (.pdf)	Figshare (https://doi.org/10.6084/m9.figshare.27313236) [5]
Data file 2	Primer sequence (qPCR)	Excel file (.xlsx)	Figshare (https://doi.org/10.6084/m9.figshare.27313236) [5]
Data file 3	GO analysis	ZIP file (.zip)	Figshare (https://doi.org/10.6084/m9.figshare.27313236) [5]
Data file 4	Results of qPCR	Portable document format (.pdf)	Figshare (https://doi.org/10.6084/m9.figshare.27313236) [5]
Data set 1	Transcriptome sequencing in pea	Fastq file (fastq.gz)	CNGB (https://doi.org/10.26036/CNP0006001) [14]

imparipinnate leaf, green pod, and purple pod. Immediately after collection, the tissue samples were frozen in liquid nitrogen and stored at -80℃.

To aeesee the expression levels of target genes across various tissues, a qRT-PCR experiment was performed using the Roche Light Cycler 480 platform with a 96-well plate format. The reaction volume was set to 20 µL, with each qRT-PCR reaction performed in three independent biological replicates. The housekeeping gene Actin [6] served as an internal reference. Specific primer sequences are provided in Primer Sequence (Table 1).

RNA extraction, library construction and transcriptome sequencing

Total RNA from each tissue sample was extracted using the NucleoSpin RNA extraction kit (Macherey-Nagel, Germany) according to the manufacturer's instructions, with DNase I treatment applied during extraction to eliminate residual genomic DNA. The concentration of total RNA was measured with a Nanodrop 2000 (Thermo Fisher Scientific), and RNA integrity was assessed using an Agilent 2100 Bioanalyzer (LabChip GX). RNA-seq libraries were prepared using the MGIEasy Fast RNA library kit (DNBSEQ) following the manufacturer's guidelines. Sequencing was performed on the DNBSEQ platform, generating paired-end reads with a length of 150 base pairs.

Gene expression analysis

Quality assessment of the raw sequence data was conducted using FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/). Adaptor sequences, low-quality bases, and poly-N sequences were trimmed using fastp v0.12.4 [7] to obtain clean data. After post-processing, transcript expression levels were quantified in transcripts per million (TPM) using salmon v1.6.0 [8], with advanced options such as --validateMappings and --numBootstraps 100 to enhance mapping accuracy and estimate expression variability. The processed transcript-level data were aggregated to gene-level counts using tximport v1.22.0 [9]. PCA of the TPM values indicated that biological replicates were well clustered (Table 1).

Differential expression analysis was performed using DESeq2 v1.4.5 [10], identifying DEGs based on an adjusted p-value threshold of ≤ 0.05 and $|\log_2(\text{fold-change})| \geq 1$. GO enrichment analysis was conducted with clusterProfiler v4.0 [11]. Comparing different tissues, we identified several notable pathways with significant differences (Table 1). For example, genes upregulated in pea green pods, compared to normal stipules, were significantly enriched in the carbohydrate derivative biosynthetic process and polysaccharide metabolic process, suggesting substantial starch synthesis in pea seeds (Table 1). Similarly, comparisons between purple and green pods and between light purple vexilla and white flowers showed that upregulated genes were significantly enriched in flavonoid metabolic and biosynthetic processes (Table 1). Generally, anthocyanins, produced via the flavonoid pathway [12], are major pigmentation compounds in flowering plants and primarily contribute to the red, orange, blue, and purple flower colors [13].

The qRT-PCR results confirmed that the expression trends of selected DEGs across different tissues were consistent with the transcriptome sequencing results (Table 1). These findings underscore the differential gene expression among pea tissues and their involvement in specific biological pathways.

Limitations

- We conducted transcriptome sequencing on only 11 different tissues, which may not fully capture the gene expression profile across all tissues.
- Additionally, this study was based on only two pea varieties, which may limit the generalizability of the results due to genetic diversity. Future research should incorporate a broader range of varieties to validate these findings and increase the applicability of the conclusions.

Abbreviations

PCA Principal Component Analysis
GO Gene Ontology
DEG Differentially Expressed Gene

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12863-024-01287-x.

Supplementary Material 1

Author contributions
KHJ, GL, PFC, NNL, XYZ, and FJS conceptualized and designed the study. LLL and ZWW performed the data analysis. ZWW was responsible for tissue sampling. RZL drafted the initial manuscript. GL and KHJ provided critical revisions and finalized the manuscript. All authors reviewed and approved the final version of the manuscript.

Funding
We extend our gratitude to the Modern Agriculture Coarse Grain Industry Technology System of Shandong Province for their valuable support. This research was supported by the Natural Science Foundation of Shandong Province for Young Scholars (ZR2021QC230), the Key R&D Program of Shandong Province (2021LZGC025, 2022LZGC022, and 2023LZGC001), and the Agricultural Science and Technology Innovation Project of SAAS (CXGC2024F13).

Data availability
The data that support the findings of this study have been deposited into CNGB Sequence Archive (CNSA) of China National GeneBank DataBase (CNGBdb) with accession number CNP0006001.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 28 August 2024 / Accepted: 26 November 2024
Published online: 20 February 2025

References

1. Li G, Liu R, Xu R, Varshney RK, Ding H, Li M, Yan X, Huang S, Li J, Wang D. Development of an Agrobacterium-mediated CRISPR/Cas9 system in pea (*Pisum sativum* L). Crop J. 2023;11:132–9.
2. Dahl WJ, Foster LM, Tyler RT. Review of the health benefits of peas (*Pisum sativum* L). Br J Nutr. 2012;108:S3–10.
3. Savage G, Deo S. The nutritional value of peas (*Pisum sativum*). Literature Rev. 1989;59:66–83.
4. Parihar A, Bohra A, Dixit G. Nutritional benefits of winter pulses with special emphasis on peas and rajmash. Biofortification Food Crops 2016;61–71.
5. Li RZ. Transcriptomic analysis and validation data support. figshare. Figure 2024. https://doi.org/10.6084/m9.figshare.27313236.
6. Foo E, Bullier E, Goussot M, Foucher F, Rameau C. The branching gene RAMOSUS1 mediates interactions among two novel signals and auxin in pea. Plant Cell. 2005;17:464–74.
7. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:i884–90.
8. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14:417–9.
9. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research. 2015;4:1521.
10. Love MI, Huber W, Anders S. Moderated estimation of Fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:1–21.
11. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16:284–7.
12. Pervaiz T, Songtao J, Faghihi F, Muhammad Salman H, Fang J. Naturally occurring anthocyanin, structure, functions and biosynthetic pathway in fruit plants. J Plant Biochem Physiol. 2017;05:1–9.
13. Castañeda-Ovando A, de Lourdes Pacheco-Hernández M, Páez-Hernández ME, Rodríguez JA, Galán-Vidal CA. Chemical studies of anthocyanins: a review. Food Chem. 2009;113:859–71.
14. Li LL. Transcriptome sequencing in pea. CNGBdb Dataset. 2024. https://doi.org/10.26036/CNP0006001.

Publisher's note