Data in Brief 69 (2026) 113347



Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib



Data Article

Dataset of full-length transcriptome assembly and annotation of a two-year-old *Bupleurum chinense* DC



Yanxia Gao[a,1], Binbin Wang[a,1], Jing Wang[a], Yan Xie[a], Jiahui Wang[b], Lan Yang[c], Gaixia Zhang[d], Hui Wang[a,*]

[a] *Sericulture Research Institute, Chengde Medical University, Chengde, 067000, China*
[b] *School of Basic Medical Sciences, Chengde Medical University, Chengde, 067000, China*
[c] *Department of Traditional Chinese Medicine, Chengde Medical University, Chengde, 067000, China*
[d] *College of Pharmacy, Shandong University of Traditional Chinese Medicine, Jinan, 250355, China*

ARTICLE INFO

Article history:
Received 14 July 2026
Revised 18 August 2026
Accepted 8 October 2026
Available online 13 October 2026

Dataset link: Full-length transcriptome sequencing of Bupleurum chinense DC. (Original data)

Keywords:
Bupleurum Chinense DC
Full-length transcriptome
Gene functional annotation
SSR

ABSTRACT

Transcriptome sequencing has become an important tool for gene discovery in medicinal plants. Compared with next-generation sequencing (NGS), third-generation sequencing, also known as single-molecule real-time (SMRT) sequencing, provides longer reads and avoids PCR amplification, making it particularly suitable for full-length transcriptome analysis. *Bupleurum chinense* DC. is an important medicinal herb in China. However, genome information for *B. chinense* remains limited. In this study, total RNA was separately extracted from the root, stem, and leaf of a two-year-old *B. chinense* plant, and equal amounts of RNA from the three tissues were pooled for PacBio SMRT full-length transcriptome sequencing. This study generated the first full-length transcriptome dataset for *B. chinense*. The dataset includes 11,448 annotated unigenes, 690 transcription factors, and 6178 SSR markers, with many unigenes associated with terpenoid biosynthesis and other secondary metabolism pathways. The raw data were deposited in the SRA database under BioProject accession number PRJNA1494021. This dataset will facilitate gene discovery, functional characterization, saikosaponin

* Corresponding author.
E-mail address: wanghui@cdmc.edu.cn (H. Wang).
[1] Authors contributed equally to the manuscript.

https://doi.org/10.1016/j.dib.2026.113347
2352-3409/© 2026 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

biosynthesis research, marker development, and molecular breeding in *B. chinense* and related species.

© 2026 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Specifications Table

Subject	Biology
Specific subject area	Medicinal Plant Transcriptome Analysis
Type of data	Table, Figure, Raw, Processed
Data collection	The samples of root, stem and leaf of *B. chinense* were processed using E.Z.N.A.® HP Total RNA Kit (Omega Bio-tek, USA). Only RNA samples with RIN $\geq$ 7.0 were retained. RNA purity and concentration were determined by a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, USA) and a Qubit 3 Fluorometer (Thermo Fisher Scientific, USA), respectively. RNA was sequenced on a PacBio Sequel II platform using Single Molecule Real-Time(SMRT)technology. Raw PacBio sequencing data were processed using SMRT Link (v11.0) and the Iso-Seq3 pipeline to obtain full-length transcript sequences.
Data source location	Root, stem and leaf of *B. chinense* DC. were collected at: • Chengde Medical University • City/Town/Region: Anyuan Road, Shuangqiao District, Chengde • Country: China
Data accessibility	Repository name: National Center for Biotechnology Information (NCBI) Data identification number: https://www.ncbi.nlm.nih.gov/ under the accession number PRJNA1494021 Direct URL to data: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1494021 The BioSample and Sequence Read Archive (SRA) numbers are SAMN61584846 and SRR39583740, respectively.
Related research article	'none'.

1. Value of the Data

- This dataset provides the first full-length transcriptome assembly of cultivated *Bupleurum chinense* generated using single molecule real-time (SMRT) sequencing.
- This dataset facilitates the identification of genes involved in the biosynthesis of key medicinal components including saikosaponins, volatile oils, polysaccharides, and flavonoids. It also supports gene discovery and cloning, thereby laying a foundation for further functional validation and studies on biosynthesis mechanisms.
- This dataset offers SSR markers that may be valuable for breeding programs in *B. chinense* and related species.

2. Background

Bupleurum species are important medicinal plants. Their roots, known as Bupleuri Radix, have been used in traditional medicine for >2000 years [1]. Among the officially authenticated source species, *B. chinense* DC. is the principal cultivated species in China and the primary commercial source of Bupleuri Radix [2]. Saikosaponins (SS) are the major bioactive constituents of *B. chinense* and are responsible for a wide range of pharmacological activities. In addition, volatile oils, polysaccharides, flavonoids, and other compounds from Bupleuri Radix have been reported to exhibit antimicrobial, antioxidant, hepatoprotective, and other beneficial effects [3].

The biosynthesis of these compounds is regulated by multiple genes involved in the corresponding metabolic pathways. Therefore, obtaining sequence information for functional genes is essential for gene characterization and pathway elucidation.

Previous studies have identified several full-length cDNA sequences encoding *β-amyrin synthase* (*β-AS*), *cytochrome P450 enzymes* (*P450s*), and *UDP-glycosyltransferases* (*UGTs*), which are key genes in the saikosaponin biosynthetic pathway. These sequences were initially obtained from next-generation sequencing (NGS) libraries and subsequently cloned and validated by PCR or RACE [4–6]. However, these approaches have inherent limitations when numerous full-length gene sequences are required. PCR and RACE are labor-intensive, time-consuming, and relatively costly. Although NGS has overcome some of these drawbacks and has been widely applied to generate transcriptomic datasets for *B. chinense* [7], its short-read output, typically <300 bp, makes it difficult to obtain complete transcript sequences directly. Third-generation sequencing technologies, such as single-molecule real-time (SMRT) sequencing by Pacific Biosciences (PacBio), overcomes these limitations by generating long reads, often exceeding 10 kb. Using third-generation sequencing, Li et al. profiled 5- and 15-day-old sprouts of *B. falcatum* L. and identified 130 genes closely involved in saikosaponin biosynthesis [8]. More recently, Wang et al. used the PacBio Sequel platform to obtain full-length cDNAs of ABA-induced bHLH transcription factors from ABA-treated hairy roots of *B. chinense* [9]. Although the genome of *B. chinense*. has been reported, detailed data have not been made publicly available [10]. As a result, a complete and comprehensive set of genes involved in saikosaponin biosynthesis and the biosynthesis of other medicinal compounds has not yet been fully characterized, and full-length sequences for many pathway genes remain unavailable in *B. chinense*.

In this study, we performed full-length transcriptome sequencing of mixed root, stem, and leaf tissues from cultivated *B. chinense* using PacBio SMRT sequencing. The resulting full-length transcriptomic dataset may serve as a valuable reference resource in the absence of a publicly available genome and may facilitate gene function characterization, molecular marker development, and the identification of full-length genes in *B. chinense*.

3. Data Description

To characterize the full-length transcriptome profile of *B. chinense*, total RNA was isolated from the mixed root, stem and leaf samples. RNA-seq library was sequenced on the PacBio Sequel platform. The sequencing generated 1.61 GB of circular consensus sequences (CCSs) reads, which had been deposited in the SRA database (PRJNA1494021). Detailed statistics of transcripts and unigenes in the full-length transcriptome of *B. chinense* are shown in Table 1. After clustering and correction, 25,016 high-quality FLNC reads were retained. The longest sequence in each cluster was designated as a unigene, yielding 11,878 unigenes. This dataset provides a valuable resources for obtaining full-length unigene sequences in *B. chinense*.

Table 1
Statistics of transcripts and unigenes for the full-length transcriptome of *B. chinense*.

	Transcript subreads	Circular consensus sequence (CCS)	Full-length non-concatemer (FLNC)	Unigene
Total Number	49,403,610	543,207	268,078	11,878
Sequence Base (bp)	78,851,851,671	856,295,942	397,587,759	19,355,819
N50 Length (bp)	1683	1718	1640	1825
Max Length (bp)	362,271	19,555	6964	5735
Average Length (bp)	1596.07	1576.37	1483.1	1629.55
Mean Length (bp)	1380	1512	1421	1570
Min Length (bp)	51	91	51	309

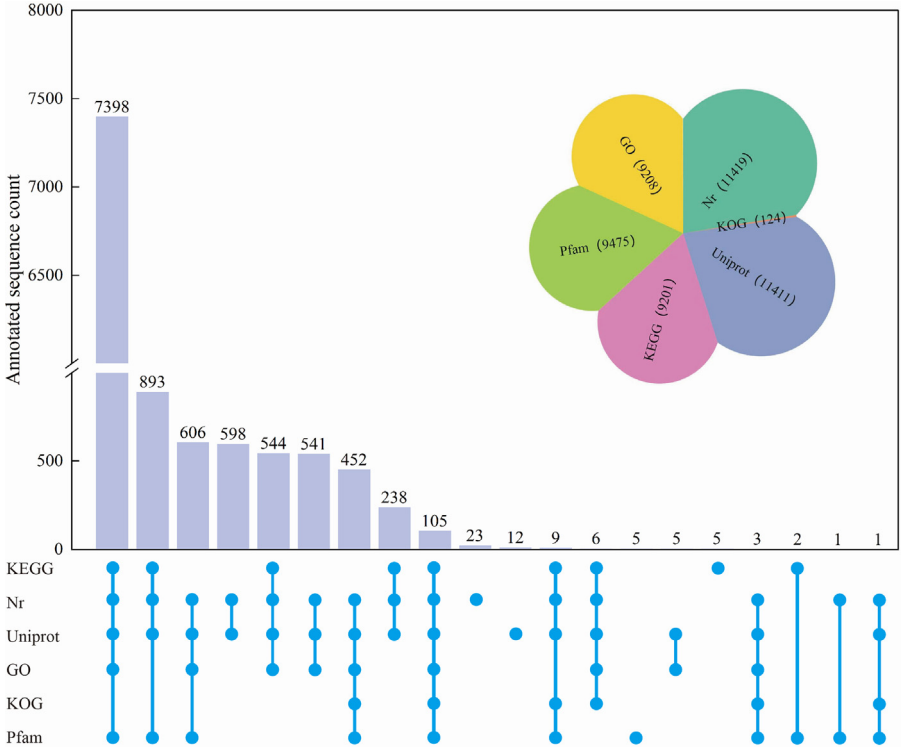


Fig. 1. Summary of the functional annotaion of the *B. chinense* transcriptome.

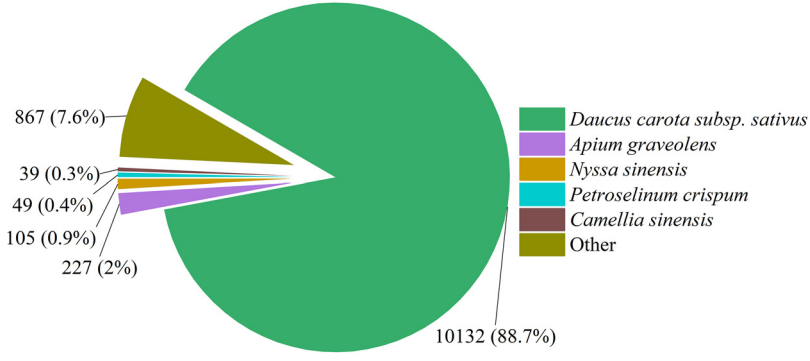


Fig. 2. Distribution of homologous species in *B. chinense* with NR annotation.

To obtain functional insights into the unigenes, annotation was carried out against six public databases, including the Non-Redundant Protein (NR) database, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), Eukaryotic Orthologous Groups (KOG), Protein families (Pfam) database, and UniProt. In total, 11,448 unigenes were annotated in at least one database (Fig. 1), providing a broad functional overview of the *B. chinense* transcriptome. BLAST searches against the NR database indicated that *Daucus carota* subsp. *sativus* showed the closest sequence similarity to *B. chinense* (Fig. 2), suggesting evolutionary relationships with other species in Apiaceae. Transcription factor analysis identified 690 unigenes distributed across 48 families, among

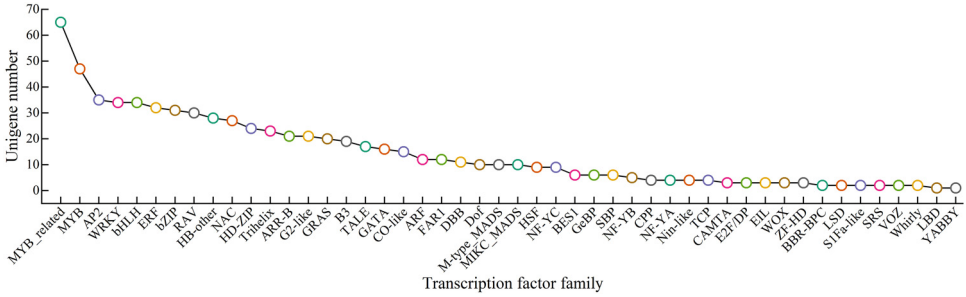


Fig. 3. Classification of unigenes encoding transcription factors.

Table 2
SSR description.

Item	Number
Total length of Sequence examined (bp)	19,355,819
Total number of SSRs identified	6178
Number of compound microsatellites	900
Total length of SSR	14,660
Relative abundance (SSR/Mb)	377.61
Relative density (bp/Mb)	759.39
Number of SSR containing sequences	4976
Number of sequences containing >1 SSR	1576
Mono-nucleotide	1092
Di-nucleotide	3247
Tri-nucleotide	888
Tetra-nucleotide	33
Penta-nucleotide	4
Hexa-nucleotide	14

which MYB-related, MYB, AP2, WRKY, bHLH, ERF, and bZIP were the most abundant (Fig. 3). These transcription factors are known to be involved in plant development, stress responses, and secondary metabolism, and may contribute to the regulation of bioactive compound accumulation in *B. chinense*.

GO annotation showed that the unigenes were mainly associated with fundamental biological processes, membrane-related cellular components, and binding activities (Fig. 4). Unigenes involved in terpenoid biosynthesis were also detected. KEGG analysis further highlighted pathways related to metabolism and secondary metabolite biosynthesis (Fig. 5). In particular, unigenes associated with terpenoid backbone biosynthesis and sesquiterpenoid and triterpenoid biosynthesis were identified, providing valuable candidate genes for elucidating the molecular basis of saikosaponin biosynthesis and for future functional studies, molecular breeding, and metabolic engineering in *Bupleurum* species.

To date, no SSR analysis based on full-length transcriptome has been reported for any species in the genus *Bupleurum*. Using MISA software, we indentified a comprehensive set of SSRs from the full-length transcriptome of *B. chinense*, which were distributed across 4976 unigenes. The SSRs were mainly composed of di-nucleotide repeats, followed by mono- and tri-nucleotide repeats, and compound SSRs were also detected (Table 2). A summary of SSR distribution by repeat number is provided in Table 3. The most frequent motifs was AT/AT (Fig. 6). The SSR markers identified from full-length unigenes provide a valuable molecular resource for genetic diversity analysis, marker-assisted breeding, and germplasm identification in *B. chinense* and related species.

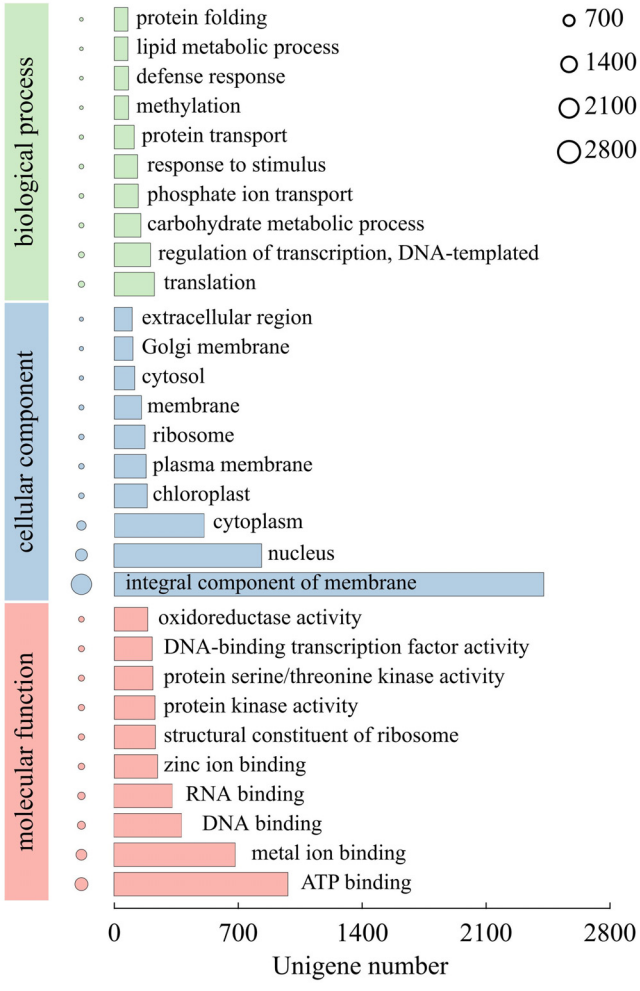


Fig. 4. GO category distribution of blast annotated *B. chinense* unigenes.

Table 3
The distribution of SSRs based on the number of repeat units.

Repeat number	Mono	Di	Tri	Tetra	Penta	Hexa	Total	Percentage(%)
5	–	–	529	19	3	9	560	10.61
6	–	758	188	8	1	2	957	18.13
7	–	524	75	2	–	1	602	11.41
8	–	529	30	2	–	–	561	10.63
9	–	422	28	1	–	–	451	8.54
10	440	451	19	–	–	1	911	17.26
11	223	324	4	1	–	–	552	10.46
12	116	122	3	–	–	–	241	4.57
13	83	50	3	–	–	–	136	2.58
14	65	24	2	–	–	–	91	1.72
≥15	165	43	7	–	–	1	216	4.09

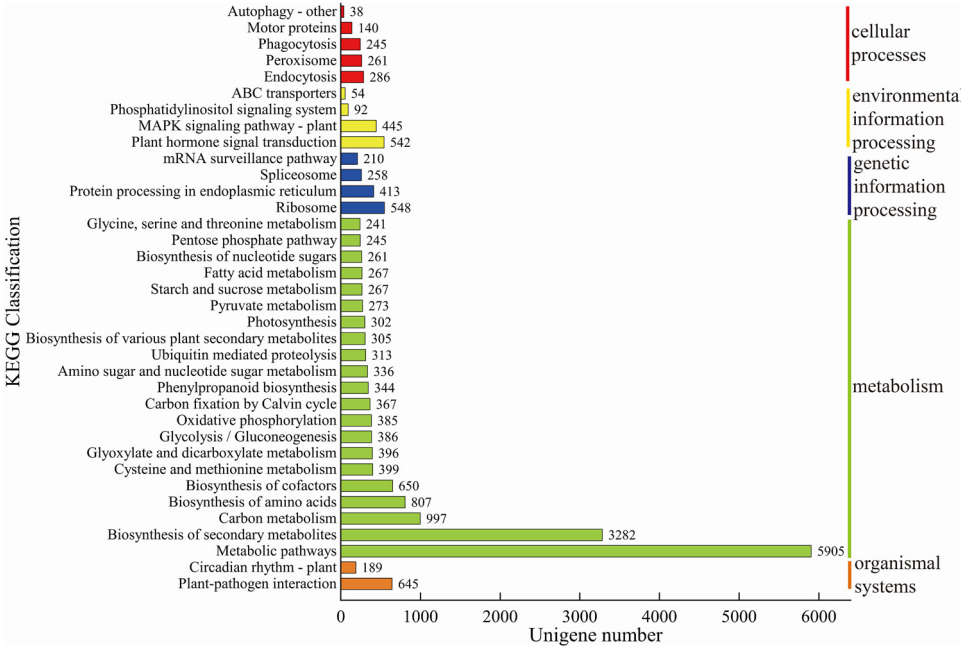


Fig. 5. Functional classification of KEGG identified in *B. chinense* transcriptome.

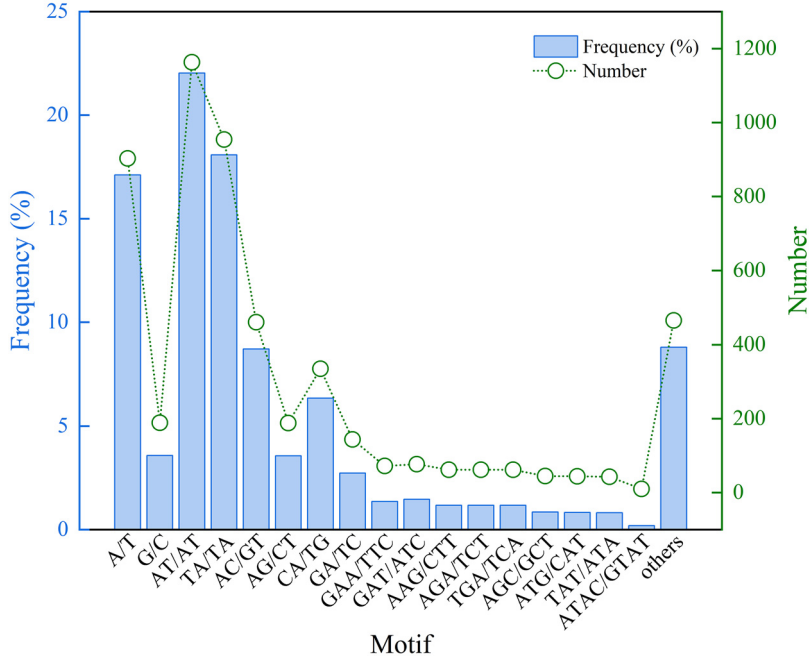


Fig. 6. Distribution of SSRs across different repeat motifs in the *B. chinense* transcriptome.

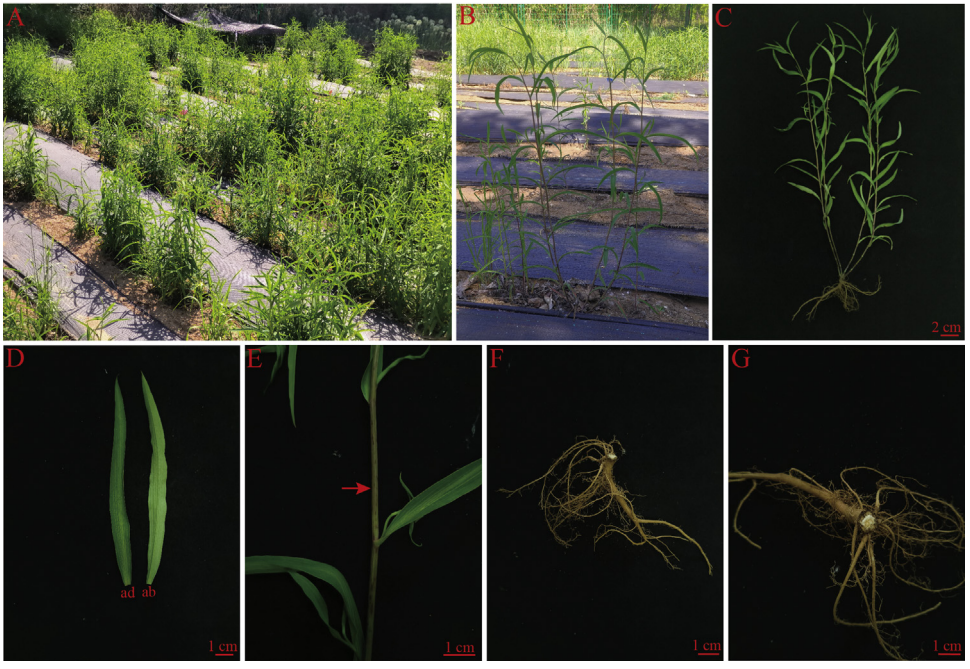


Fig. 7. Images of a two-year-old *B.* Chinese. A. Plantation of *B. chinense*. B, C. Whole plant of Zhongchai No. 2. D. Leaf; ad: adaxial surface; ab: abaxial surface. E. Stem. F, G. Root.

4. Experimental Design, Materials and Methods

4.1. Plant materials

B. chinense cultivar Zhongchai No.2 was used in this study. Seeds were sown on May 1, 2022 in the experimental cultivation field of Chengde Medical University, Chengde, China (41°2′8″ N and 117°56′48″ E, 218 m above sea level), with a row spacing of 40 cm and a plant spacing of 10 cm (Fig. 7). Seedlings were covered with a shading net (four-needle mesh) and the net was removed after the six-leaf stage. Field management of *B. chinense* was conducted in accordance with the group standard of the China Association of Chinense Medicine, namely the "Code of Practice for Good Agricultural Practice of Bupleuri Radix" (T / CACM 1374.128 – 2021).

Irrigation frequency was determined according to soil moisture conditions and weather conditions, and watering was performed in the evening. During periods of high summer temperatures, irrigation frequency was increased appropriately. In rainy summer periods, drainage was strengthened to prevent root rot caused by the combined stress of high temperature and waterlogging. Weeds were removed regularly. On June 1, 2024, a healthy *B. chinense* plant was randomly selected. Root, stem, and leaf tissues were separated, rapidly frozen in liquid nitrogen, and stored at – 80 °C until further use.

4.2. Total RNA extraction, library preparation, and sequencing

Total RNA was extracted from root, stem and leaf tissues using an E.Z.N.A.® HP Total RNA Kit (Omega Bio-tek, USA) according to the manufacturer's instructions. Only RNA samples with RIN $\geq$ 7.0 were retained. RNA purity and concentration were determined by a Nanodrop 2000

spectrophotometer (Thermo Fisher Scientific, USA) and a Qubit 3 Fluorometer (Thermo Fisher Scientific, USA), respectively.

Equal amounts of high-quality total RNA from the three tissues were pooled prior to cDNA library preparation to maximize transcript diversity and increase the likelihood of capturing low-abundance full-length isoforms. This strategy was chosen because the objective of the study was to generate a comprehensive reference transcriptome rather than tissue-specific expression profiles.

Full-length cDNA was synthesized using the Clontech SMARTer^{TM} PCR cDNA Synthesis Kit (Clontech Laboratories, Palo Alto, CA, USA) following the manufacturer's instructions. Briefly, the oligo(dT) primer annealed to the poly(A) tail of eukaryotic mRNA during first-strand synthesis, and the SMARTer template-switching mechanism enabled the addition of defined sequences to both ends of the cDNA, thereby enriching for full-length transcripts. The resulting full-length cDNA was amplified by PCR, purified with PB magnetic beads, and quantified using a Qubit 3.0 fluorometer. To improve library quality, cDNA fragments smaller than 1 kb were removed by size selection with PB bead. The size-selected cDNA was then subjected to end repair and ligated to SMRTbell hairpin adapters to generate SMRTbell templates. Unligated DNA molecules and other non-circularized products were removed by exonuclease digestion, followed by an additional PB bead purification step. Library quality was assessed by Qubit 3.0 quantification and fragment size analysis using an Agilent 2100 Bioanalyzer. Only libraries meeting the required quality criteria were used for sequencing. SMRT sequencing was performed on a PacBio Sequel II platform. The library was sequenced using 1 SMRT cells, with a Sequel II Binding Kit 2.0 chemistry.

4.3. Transcriptome assembly

Raw PacBio sequencing data were processed using SMRT Link (v11.0) and the Iso-Seq3 pipeline. Polymerase reads were split into subreads, which were used to generate circular consensus sequences (CCSs). CCS reads were screened for the presence of 5′ and 3′ primers and poly(A) tails. Reads containing both primers and lacking concatemeric structures were classified as full-length non-concatemeric (FLNC) reads. FLNC reads were clustered and polished using the Iso-Seq3 cluster module. Redundant sequences were removed using CD-HIT (v4.8.1) with a 99% similarity threshold [11], and the resulting high-quality non-redundant full-length transcripts were retained for downstream analyses.

4.4. Functional annotation

To annotate the structural and functional features of the newly identified transcripts, coding sequences (CDSs) were predicted using TransDecoder (v5.5.0) with the parameters -m 50 and –single_best_only (http://transdecoder.github.io). Functional annotation was then performed against the NR (downloaded on 2024–02–07), Pfam (release 30.0), UniProt (release 2021_01), KEGG (release 97.0), GO (release 2021–01–01), and KOG (last updated in 2003) databases using both sequence similarity- and domain-based approaches. Predicted protein sequences were aligned against the NR and UniProt databases using DIAMOND blastp (v2.0.6.144) with an E-value cutoff of 1e-5 [12]. Conserved domains were identified using hmmscan (v3.3.2) against Pfam with an E-value threshold of 0.01 [13]. KEGG Orthologs were assigned using kofam_scan (v1.3.0) with parameter -E 1e-5. GO annotations were summarized into biological process, cellular component, and molecular function categories. Species distribution of NR hits was also analyzed.

4.5. Analysis of transcription factors and simple sequence repeats (SSRs)

Transcription factors were identified in B. *chinense* unigenes using PlantTFDB (v5.0) [14]. MISA (v.1.01) [15] was used to retrieve the SSR loci from the full-length transcriptome of *B. chinense*. The search criteria were as follows: the minimum repeat numbers for mono-, di-, tri-, tetra-, penta-, and hexanucleotide SSRs were 10, 6, 5, 5, 5 and 5, respectively.

Limitations

None.

Ethics Statement

The authors have read and followed the ethical requirements for publication in Data in Brief and confirmed that the current work does not involve human subjects, animal experiments, or any data collected from social media platforms.

CRediT Author Statement

Yanxia Gao: Methodology, Investigation, Writing - Original draft preparation. **Binbin Wang:** Data curation, Software. **Jing Wang:** Visualization. **Yan Xie:** Methodology, Visualization. **Jiahui Wang:** Visualization. **Lan Yang:** Visualization. **Gaixia Zhang:** Data curation, Investigation. **Hui Wang:** Conceptualization, Supervision, Data curation, Writing-Reviewing and Editing.

Data Availability

Full-length transcriptome sequencing of Bupleurum chinense DC. (Original data) (National Center for Biotechnology Information)

Acknowledgements

This work was financially supported by the Natural Science Foundation of Hebei Province [grant numbers H2022406031]; the Natural Science Foundation of Shandong Province [grant numbers ZR2022QH160]; and Scientific Research Program of Traditional Chinese Medicine in Hebei Province [grant numbers 2023138].

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Z.H. Wang, L. Tian, Y.S. Xiao, M.Y. Zhao, Y.Y. Chang, Y.J. Zhou, S.Y. Liu, H.X. Zhao, Y. Xiu, Quantitative and differential analysis between *Bupleurum chinense* DC. and *Bupleurum scorzonerifolium* Willd. Using HPLC-MS and GCMS– coupled with multivariate statistical analysis, Molecules 28 (15) (2023) 5360, doi:10.3390/molecules28155630.

[2] S.W Zhou, L.L. Yang, J. Wan, L. Chen, X.P. Gu, L. Qiao, L.X. Chu, N. Dong, C.M. Dong, W.S. Feng, Suitable concentrations of brassinolide application enhances growth performance and saikosaponin biosynthesis by transcriptional activation of triterpenoid pathway genes in *Bupleurum chinense* DC, Front. Plant Sci. 16 (2025) 1517434, doi:10.3389/fpls.2025.1517434.
[3] C.H. Zeng, J. Zhao, H. Chen, C. Xin, B. Wang, M. Yu, J.H. Wei, Traditional use, germplasm identification, phytochemistry, pharmacology of Bupleuri Radix: a review, Med. Plant Biol. 2 (2023), doi:10.48130/MPB-2023-0018.
[4] Y.P. Mao, H. Chen, J. Zhao, Y.C. Li, L. Feng, Y.P. Yang, Y.G. Zhang, P. Wei, D.B. Hou, Molecular cloning, functional characterization and expression of the β-amyrin synthase gene involved in saikosaponin biosynthesis in *Bupleurum chinense* DC, J. Plant Biochem. Biotechnol. 32 (2023) 284–295, doi:10.1007/s13562-022-00804-2.
[5] M.Y. K.M. Dong, H. Chen, S.Y. Wang, X. Yin, C.X. Mo, M. Zhang, Y.C. Li, Y. Luo, Z.R. Yu, Z.S. Yang, J. Zhao, C. Xin, W.S. Chen, Z.Y. Xue, J.H. Wei, Differing functions of two CYP716 genes for saikosaponin biosynthesis in *Bupleurum chinense*, Crop. J. 14 (1) (2026) 224–234, doi:10.1016/j.cj.2025.12.007.
[6] C. Sui, J. Zhang, J.H. Wei, S.L. Chen, Y. Li, J.S. Xu, Y. Jin, C.X. Xie, Z.H. Gao, H.J. Chen, C.M. Yang, Z. Zhang, Y.H. Xu, Transcriptome analysis of *Bupleurum chinense* focusing on genes involved in the biosynthesis of saikosaponins, BMC Genom. 12 (1) (2011) 539, doi:10.1186/1471-2164-12-539.
[7] X.H Feng, Y. Sun, Y. Fan, Q.F. Zhang, X. Bu, D.M. Gao, Integrated physiological, transcriptomics and metabolomics analysis revealed the molecular mechanism of *Bupleurum chinense* seedlings to drought stress, PLoS One 19 (6) (2024) e0304503, doi:10.1371/journal.pone.0304503.
[8] Y.C. Li, J. Zhao, Chen Hua, Y.P. Mao, Y.P. Yang, L. Feng, C.X. Mo, L. Huang, D.B. Hou, M. Yu, Transcriptome level reveals the triterpenoid saponin biosynthesis pathway of *Bupleurum falcatum* L, Genes (Basel) 13 (12) (2022) 2237, doi:10.3390/genes13122237.
[9] H. Wang, S.Q. Hu, T. Li, X.J. Qu, J.Q. Zhang, B.S. Wang, Y.X. Sun, R. Cao, Y.T. Yan, Z. Song, X.N. Zhang, R. Luo, Y.R. Tong, C.L. Liu, Comparative transcriptome analysis reveals abscisic acid-induced bHLH transcription factors involved in saikosaponin biosynthesis in *Bupleurum chinense* DC, Plant Signal Behav. 20 (1) (2025) e2495301, doi:10.1080/15592324.2025.2495301.
[10] Q.F. Zhang, M. Li, X.Y. Chen, G.X. Liu, Z.e Zhang, Q.Q. Tan, Y. Hu, Y.Y. Fan, Y.Y. Liu, T.S. Zhu, X. Yang, M.M. Yue, X. Bu, Y.Q. Zhang, Chromosome-level genome assembly of *Bupleurum chinense* DC provides insights into the saikosaponin biosynthesis, Front. Genet. 13 (13) (2022) 878431, doi:10.3389/fgene.2022.878431.
[11] L.M. Fu, B.F. Niu, Z.W. Zhu, S.T. Wu, W.Z. Li, CD-HIT: accelerated for clustering the next-generation sequencing data, Bioinformatics 28 (23) (2012) 3150–3152, doi:10.1093/bioinformatics/bts565.
[12] B. Buchfink, C. Xie, D.H. Huson, Fast and sensitive protein alignment using DIAMOND, Nat. Methods 12 (1) (2015) 59–60, doi:10.1038/nmeth.3176.
[13] A. Prakash, M. Jeffryes, A. Bateman, R.D. Finn, The HMMER webserver for protein sequence similarity search, Curr. Protoc. Bioinform. 60 (1) (2017) 3.15.11–13.15.23, doi:10.1002/cpbi.40.
[14] F. Tian, D.C. Yang, Y.Q. Men, J.P. Jin, G. Gao, PlantRegMap: charting functional regulatory maps in plants, Nucleic. Acids Res. 48 (D1) (2019) D1104–D1113, doi:10.1093/nar/gkz1020.
[15] S. Beier, T. Thiel, T. Münch, U. Scholz, M. Mascher, MISA-web: a web server for microsatellite prediction, Bioinformatics 33 (16) (2017) 2583–2585, doi:10.1093/bioinformatics/btx198.