JIPB Journal of Integrative Plant Biology

Brief Communication
https://doi.org/10.1111/jipb.70401

T2T genome assembly and CRISPR/Cas9 editing of *Medicago lupulina*

Qing Yuan[1,2†], Cheng Li[1†], Guangyan Wang[1†], Wenfei Dai[1†], Eerdungaridi[3], Di Wu[3], Qingdun Fan[1], Yongping Yang[1], Yuanwen Duan[1*] and Xudong Sun[1*]

1. Yunnan Key Laboratory of Crop Wild Relatives Omics, The Germplasm Bank of Wild Species, Institute of Tibetan Plateau Research at Kunming, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China
2. School of Ecology and Environmental Science, Yunnan University, Kunming 650500, China
3. National Technology Innovation Center for Prataculture, Hohhot 010000, China
†These authors contributed equally to this work.
*Correspondences: Yuanwen Duan (duanyw@mail.kib.ac.cn, Dr. Duan is fully responsible for the distribution of all materials associated with this article); Xudong Sun (sunxudong@mail.kib.ac.cn)

Medicago lupulina L. (black medick) (Figure S1) is a leguminous forage species widely distributed in temperate and subtropical regions around the world. It grows across a broad altitudinal range from sea level to over 4,000 m, making it a potential model plant species for studying ecological adaptation mechanisms (Ye et al., 2025). Furthermore, due to the high crude protein (above 25.79%) and the low crude fiber (less than 19.13%), *M. lupulina* is a potential high-quality forage (Feng and Cao, 2009), which may also serve as an excellent germplasm resource for the genetic improvement of alfalfa. However, the lack of a high-quality reference genome and an efficient genetic transformation system have severely impeded functional genomic studies and the exploitation of its breeding potential. Here, we assembled a telomere-to-telomer (T2T) gap-free genome of *M. lupulina* and established CRISPR/Cas9-mediated gene editing targeting *PINNATE-LIKE PENTAFOLIATA1* (*PINNA1*), thereby providing essential genomic tools for genetic dissection and molecular breeding in this species.

The *M. lupulina* plant material used in this study was collected from Lasa (91°21′8.1″, 29°51′22.64″, altitude 3,800 m), Xizang, China. Survey analysis estimated its genome size as approximately 640.95 Mb, with a heterozygosity rate of 0.2% (Table S1). We assembled the T2T genome of *M. lupulina* using BGI-seq short reads (127×), PacBio HiFi reads (73×), ONT ultra-long reads (63×), and Hi-C data (153×) (Table S2). The final assembled genome size was 568,875,845 bp, successfully anchored onto eight pseudochromosomes (Figures 1A, S2), with a BUSCO completeness of 98.8% (Table S3), which was higher than the published genomes of *Medicago* species (Li et al., 2020; Yin et al., 2021; Hu et al., 2026). Telomere and centromere positions were identified (Figures S3, S4; Tables S4, S5). The high accuracy and completeness of the assembly were confirmed by multiple quality metrics, including read mapping rates exceeding 98%, a scaffold N50 of 70.2 Mb, an LAI of 13.77, and a QV of 57.14 (Figures S5, S6; Tables S6–S8). Repetitive sequences were 381,885,597 bp, accounting for 67.13% of the genome, with transposable elements (TEs) comprising 66.18% (Tables S9, S10). A total of 30,070 protein-coding genes were annotated in the *M. lupulina* genome, and 95.8% of these were functionally annotated using GO, KEGG, NR, KOG, and SwissProt databases (Figures S7, S8; Table S11). Additionally, 780 transfer RNAs (tRNAs), 2382 microRNAs (miRNAs), 1,175 ribosomal RNAs (rRNAs), and 935 small nuclear RNAs (snRNAs) were identified (Table S12).

Phylogenetic analysis revealed that *M. lupulina* diverged from the lineage comprising *M. polymorpha*, *M. truncatula,* and *M. sativa* subsp. *caerulea* approximately 8.5 million years ago (Mya) (Figure 1B) and shares with them the whole-genome duplication (WGD) event that occurred in the Papilionoideae (Figure S9; Young et al., 2011). Compared with the genomes of *M. truncatula* and *M. sativa* subsp. *caerulea*, we identified a pronounced chromosomal translocation between chromosomes 7 and 8 in *M. lupulina* (Figure S10). Additionally, a pair of large interchromosomal rearrangements was detected in the *M. truncatula* genome, which is consistent with previous studies (Li et al., 2020). Further comparative genomic analyses revealed 583 expanded and 636 contracted gene families in *M. lupulina*. The expanded gene families were significantly enriched in KEGG pathways associated with environmental adaptation, plant–pathogen interactions, and glycosyltransferases, while the contracted gene families were mainly enriched in protein kinases and cytochrome P450s (Figures S11–S14; Tables S14–S17). We further identified 210 *M. lupulina*-specific gene families that were significantly enriched in DNA repair and recombination proteins, as well as lysine degradation pathways (Figures S15–S17; Tables S18, S19). These expanded, contracted, and species-specific gene families of *M. lupulina* may collectively contribute to its adaptation to high altitudes. In addition, a total of 118 positively selected genes were identified (Tables S20–S22), including candidates functionally involved in environmental adaptation, such as *LEA14* (Xiong et al., 2002) and *ZAT1* (Feng et al., 2026).

Although genetic transformation and gene-editing systems have been established for *M. truncatula* and *M. sativa*, these protocols remained non-transferable to *M. lupulina*. To overcome this bottleneck, the same genotype was used to establish the gene-editing system, for which the culture conditions were systematically optimized, yielding a stable regeneration system with a plant regeneration efficiency of 28.25% (Figure 1C). *PINNA1* is a key regulator of compound leaf development in *M. truncatula*, and its mutation converts trifoliate leaves into pinnate pentafoliate ones (He et al., 2020). The *PINNA1* ortholog in *M. lupulina* was identified (Figures S18, S19) and we then generated the *MlPINNA1–CRISPR* construct according to a previously described method (Zhao et al., 2024). CRISPR/Cas9-mediated gene editing of *MlPINNA1* was subsequently performed and we obtained positive gene-edited plants (Figure S20). Three independent gene-edited plants were obtained, and insertions and deletions at the target sites were confirmed by sequencing, with a gene-editing efficiency of 2.83%. Notably, a pentafoliate leaf phenotype was observed in all the *Mlpinna1* lines (Figure 1D–F).

In summary, we generated a T2T gap-free genome assembly and established a CRISPR/Cas9-mediated gene-editing system for *M. lupulina*, a wild relative of cultivated alfalfa. The T2T genome provides a high-quality reference to resolve candidate genes, regulatory regions, repetitive sequences, and structural variations

© 2026 The Author(s). *Journal of Integrative Plant Biology* published by John Wiley & Sons Australia, Ltd on behalf of Institute of Botany, Chinese Academy of Sciences.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

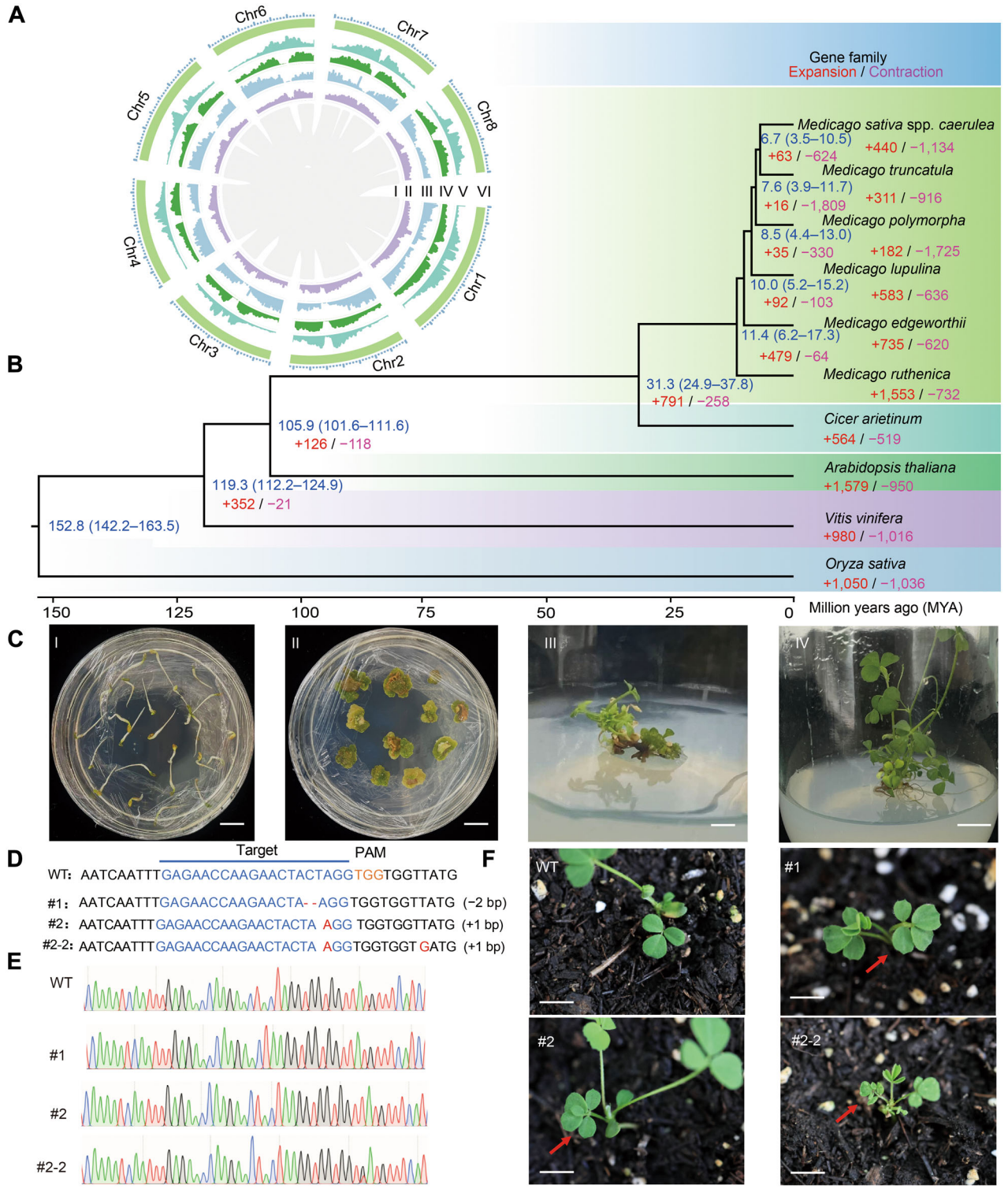


Figure 1. T2T Genome assembly and annotation, and establishment of regeneration and CRISPR-Cas9 gene-editing systems in *Medicago lupulina*

(A) Characterization and features of the *M. lupulina* chromosomes. (I) Synteny between chromosomes, (II) GC content, (III) Gene density, (IV) Copia transposon density, (V) Gypsy transposon density, and (VI) schematic representation of *M. lupulina* chromosome sequences. **(B)** Inferred maximum likelihood phylogenetic tree across 10 plant species based on 515 single-copy homologous genes. Gene family expansion (+) and contraction (–) are indicated in red and purple, respectively. The blue numbers represent the divergence time between species. **(C)** Establishment of regeneration systems for *M. lupulina*. (I) 7-d-old seedlings, (II) callus induction medium, (III) shoot induction and elongation medium, and (IV) root induction medium, Scale bar = 1 cm. **(D)** CRISPR-Cas9 mediated gene editing of *MlPINNA1*. Blue represents the target sequence and red marks the editing events at the target. **(E)** Sanger sequencing chromatograms of the wild type and representative mutant alleles at the target sites. **(F)** Comparison of leaf phenotypes between wild type (WT) and *Mlpinna1* lines. Red arrows highlight phenotypic variations in *Mlpinna1* mutant lines. Scale bar = 0.5 cm.

associated with forage quality traits, including high crude protein content. Together, this genomic and gene-editing system will facilitate functional genomics, molecular breeding, and identification of the candidate genes underlying desirable forage characteristics in *M. lupulina*. This work also lays a foundation for genetic improvement of alfalfa and potentially *de novo* domestication of new high-protein cultivars adapted to higher altitude environments.

DATA AVAILABILITY STATEMENT

The genome assembly data generated in this study have been uploaded to the National Genomic Data Center with the accession number PRJCA062977, and the annotations have been uploaded to Figshare (https://figshare.com/) under DOI: 10.6084/m9.figshare.32098651.

ACKNOWLEDGEMENTS

We gratefully thank Dr. Caspar C. C. Chater (Royal Botanic Gardens, UK) for revising the manuscript for writing style, English language, and syntax. This work was financially supported by the National Technology Innovation Center for Prataculture Special fund for innovation platform construction (CCPTZX2024GJ11), the Major Science and Technology Project of Qinghai Province (2025-NK-A1-01), the Major Program for Basic Research Project of Yunnan Province (202401BC070001), the Major Program of Yunnan Province (202602AE090029), and the National Natural Science Foundation of China (32371586, 32370405, and 32500309).

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

Y.D. and X.S conceived the study and revised the manuscript; G.W, E., D.W, Y.G., and Y.Y. collected the biological materials; C.L. and Q.F. established the genetic transformation system and performed gene editing; and Q.Y. and W.D. performed genome assembly, annotation, comparative genomic analyses, and drafted the manuscript. All authors have read and approved the final manuscript.

Edited by: Feifei Yu, China Agricultural University, China

Received Jun. 2, 2026; **Accepted** Sep. 7, 2026

OO: OnlineOpen

REFERENCES

Feng, M., Liu, W., Chen, X., Huang, L., Peng, X., Li, Y., Gao, S., Lin, Z., Xu, L., and Guo, J. (2026). A hierarchical ScMYB11L–ScZAT1–ScDREB1A module defines a derepression pathway for cold-response regulation in sugarcane. Plant Stress. **21:** 101340.

Feng, Y., and Cao, Z. (2009). Study on the nutrition quality of wild germplasm of *Medicago lupulina*. Pratacult. Sci. **26:** 80–84.

He, L., Liu, Y., He, H., Liu, Y., Qi, J., Zhang, X., Li, Y., Mao, Y., Zhou, S., Zheng, X., et al. (2020). A molecular framework underlying the compound leaf pattern of *Medicago truncatula*. Nat. Plants **6:** 511–521.

Hu, H., Wu, S., Zheng, Y., Li, A., Wang, Z., Qu, K., Yang, Y., Wang, N., Yang, X., Wan, Y., et al. (2026). Structural variation drives rhizome innovation and adaptive divergence in sister *Medicago* species. J. Integr. Plant Biol. **68:** 406–424.

Li, A., Liu, A., Du, X., Chen, J.-Y., Yin, M., Hu, H.-Y., Shrestha, N., Wu, S.-D., Wang, H.-Q., Dou, Q.-W., et al. (2020). A chromosome-scale genome assembly of a diploid alfalfa, the progenitor of autotetraploid alfalfa. Hortic. Res. **7:** 194.

Xiong, L., Schumaker, K.S., and Zhu, J.-K. (2002). Cell signaling during cold, drought, and salt stress. Plant Cell **14:** S165–S183.

Ye, Q., Zhou, C., Lin, H., Luo, D., Jain, D., Chai, M., Lu, Z., Liu, Z., Roy, S., Dong, J., et al. (2025). Medicago2035: Genomes, functional genomics, and molecular breeding. Mol. Plant **18:** 219–244.

Yin, M., Zhang, S., Du, X., Mateo, R.G., Guo, W., Li, A., Wang, Z., Wu, S., Chen, J., and Liu, J. (2021). Genomic analysis of *Medicago ruthenica* provides insights into its tolerance to abiotic stress and demographic history. Mol. Ecol. Resour. **21:** 1641–1657.

Young, N.D., Debellé, F., Oldroyd, G.E.D., Geurts, R., Cannon, S.B., Udvardi, M.K., Benedito, V.A., Mayer, K.F.X., Gouzy, J., Schoof, H., et al. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. Nature **480:** 520–524.

Zhao, H., Zhao, S., Cao, Y., Jiang, X., Zhao, L., Li, Z., Wang, M., Yang, R., Zhou, C., Wang, Z., et al. (2024). Development of a single transcript CRISPR/Cas9 toolkit for efficient genome editing in autotetraploid alfalfa. Crop J. **12:** 788–795.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article: http://onlinelibrary.wiley.com/doi/10.1111/jipb.70401/suppinfo