

Global transcriptome analysis of *Sabina chinensis* (Cupressaceae), a valuable reforestation conifer

Xian-Ge Hu · Hui Liu · Jia-Qing Zhang · Yan-Qiang Sun ·
YuQing Jin · Wei Zhao · Yousry A. El-Kassaby · Xiao-Ru Wang ·
Jian-Feng Mao

Received: 3 October 2015 / Accepted: 1 July 2016
© Springer Science+Business Media Dordrecht 2016

Abstract *Sabina chinensis* has broad distribution in China and is widely used in the reforestation and as an urban tree. The species is frost resistant and grows well on contaminated soils and is becoming valuable for soil remediation and protection against air pollution. Breeding programs aimed at exploiting the species' unique properties were handicapped by the lack of basic genetic information. Here, we established a transcriptomic profiling study from five different tissues using RNA-Seq to gain insight on the functional genes and the development of molecular markers for breeding and conservation purposes. In

total 90,382,108 high-quality sequence reads (~9.0 bp) were obtained, and 116,814 unigenes (≥ 200 nt) were assembled. Of which, 45,026 and 15,589 unigenes were mapped to the Nr and KOG databases, 31,288 (26.78 %) and 17,596 (15.06 %) were annotated to GO and KEGG database, respectively. Additionally, 28,843 (24.68 %) and 43,033 (36.84 %) *S. chinensis* unigenes were aligned to the *Pinus taeda* draft genome and PLAZA2.5 database, respectively. A total of 4570 simple sequence repeat (SSR) motifs were identified in the unigenes. Furthermore, we obtained 6 (12.5 %) polymorphic and 21 (43.75 %) monomorphic loci in the verification of 48 randomly selected SSR loci. This study represents the first transcriptome data of *S. chinensis* and confirms that the transcriptome assembly data of *S. chinensis* are a useful resource for EST-SSR loci development. The substantial number of transcripts obtained will aid our understanding of the species adaptation mechanisms and provide valuable genomic information for conservation and breeding applications.

Data achieving statement The original RNA-Seq data from this study can be found in the NCBI Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/Traces/sra>) under accession number SRR3596662.

Xian-Ge Hu, Hui Liu, Jia-Qing Zhang have equally contributed to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s11032-016-0526-3) contains supplementary material, which is available to authorized users.

X.-G. Hu · H. Liu · J.-Q. Zhang · Y.-Q. Sun ·
Y. Jin · W. Zhao · X.-R. Wang · J.-F. Mao (✉)
National Engineering Laboratory for Tree Breeding, Key
Laboratory of Genetics and Breeding in Forest Trees and
Ornamental Plants, Ministry of Education, College of
Biological Sciences and Technology, Beijing Forestry
University, Beijing, China
e-mail: jianfeng.mao@bjfu.edu.cn

W. Zhao · X.-R. Wang
Department of Ecology and Environmental Science,
Umeå University, Umeå, Sweden

Y. A. El-Kassaby
Department of Forest and Conservation Sciences, Faculty
of Forestry, University of British Columbia, Vancouver,
BC, Canada