



Aug 16,
2019

SOAP-denovo-trans assembly

Eric J. Carpenter¹, Naim Matasci^{2,3}, Shuangxiu Wu⁴, Jing Sun⁴, Jun Yu⁴, Fabio Rocha Jimenez Vieira⁵, Chris Bowler⁵, Richard G. Dorrell⁵, Matt Gitzendanner⁶, Ling Li⁷, Wensi Du⁷, Kristian Ullrich⁸, Michael S. Barker⁹, James H. Leebens-Mack¹⁰, Gane Ka-Shu Wong¹¹

¹Department of Biological Sciences, University of Alberta, Edmonton, Alberta, T6G 2E9, Canada., ²CyVerse, University of Arizona, Arizona, U.S.A., ³Current address: Lawrence J. Ellison Institute for Transformative Medicine, University of Southern California, Los Angeles, CA 90033, U.S.A., ⁴CAS Key Laboratory of Genome Sciences and Information, Beijing, Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China., ⁵École Normale Supérieure, Paris., ⁶Department of Biology, University of Florida, Gainesville, Florida 32611, USA., ⁷BGI-Shenzhen, Beishan Industrial Zone, Yantian District, Shenzhen 518083, People's Republic of China., ⁸Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Biology, Plön, Germany., ⁹Department of Ecology & Evolutionary Biology, University of Arizona, Tucson, AZ 85721 USA., ¹⁰Department of Plant Biology, University of Georgia, Athens, GA 30602, USA., ¹¹Department of Biological Sciences, University of Alberta, Edmonton, Alberta, T6G 2E9, Canada. BGI-Shenzhen, Beishan Industrial Zone, Yantian District, Shenzhen 518083, People's Republic of China. Department of Medicine, University of Alberta, Edmonton, Alberta, T6G 2E1, Canada.

1 Works for me dx.doi.org/10.17504/protocols.io.38pgvrn



ABSTRACT

SOAP-denovo-trans assembly protocol used to assemble plant transcriptomes for the 1KP project.

THIS PROTOCOL ACCOMPANIES THE FOLLOWING PUBLICATION

<https://academic.oup.com/bioinformatics/article/30/12/1660/380938>

- 1 SOAPdenovo2 based DBG construction and Trinity based low frequency k-mer removal.
- 2 Map reads to the contigs and build linkages.
 - a) Utilizing single-end read information
 - b) Utilizing paired-end read information
 - c) Merge the linkages
- 3 Remove erroneous and redundant linkages
 - a) Manage ambiguous contigs
 - b) Linearize contigs into scaffolds
- 4 Divide graph into sub-graph. Oases based graph traversal.
- 5 Gap closure and output the transcripts.



This is an open access protocol distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited