



ELSEVIER

Contents lists available at ScienceDirect

## Data in Brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)

## Data Article

## Phylogeny of the plant 4/1 proteins

Sergey Y. Morozov\*, Andrey G. Solovyev, Alexey V. Troitsky

A. N. Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, 119992 Moscow, Russia

## ARTICLE INFO

## Article history:

Received 3 November 2015

Accepted 16 November 2015

Available online 25 November 2015

## ABSTRACT

The Nt-4/1 protein of unknown function has been shown to be alpha-helical and predominantly expressed in conductive tissues of tobacco plants. So far, obvious Nt-4/1 orthologs were found only in flowering plants. We report the analysis of 4/1 genes and the encoded proteins of lower land plants (Morozov et al., 2015) [1]. In this data article, we present two phylogenetic trees of angiosperm 4/1 proteins together with orthologs from liverworts, lycophytes, ferns and gymnosperms.

© 2015 The Authors. 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 area	Biology
More specific subject area	Phylogenetics
Type of data	Figure (Phylogenetic trees)
How data was acquired	Phylogenies were acquired using Fast Minimum Evolution and Neighbour-joining methods at NCBI (COBALT) and TREECON packages
Data format	Analyzed
Experimental factors	Amino acid sequences were retrieved from NCBI and/or 1KP databases (see below)
Experimental features	Sequences were aligned using NCBI protein Multiple Alignment Tool (see below)

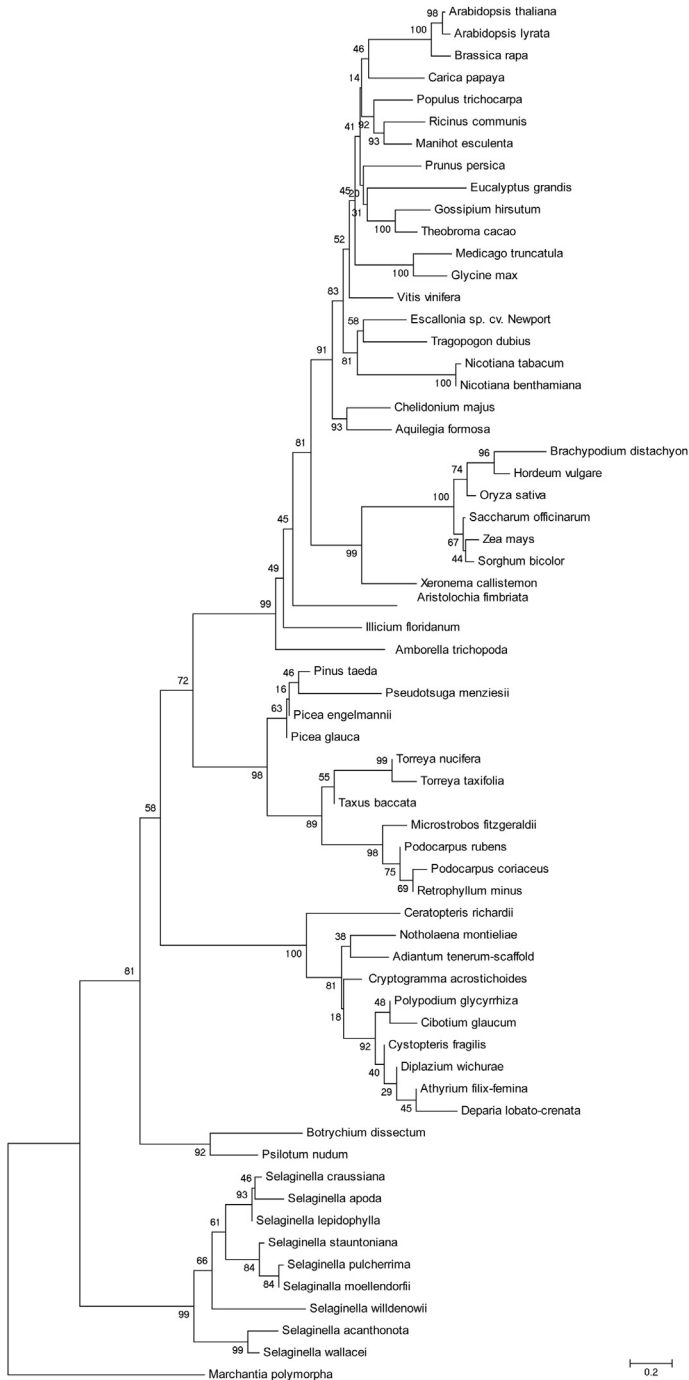
DOI of original article: <http://dx.doi.org/10.1016/j.biochi.2015.10.019>

\* Corresponding author.

E-mail address: [morozov@genebee.msu.su](mailto:morozov@genebee.msu.su) (S.Y. Morozov).

<http://dx.doi.org/10.1016/j.dib.2015.11.041>

2352-3409/© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



**Fig. 1.** Neighbour-joining (NJ) tree based on the 4/1 protein sequences showing the phylogenetic relationship between 4/1 polypeptides from land plants. The numbers indicate the NJ bootstrap values for 1000 replicates.

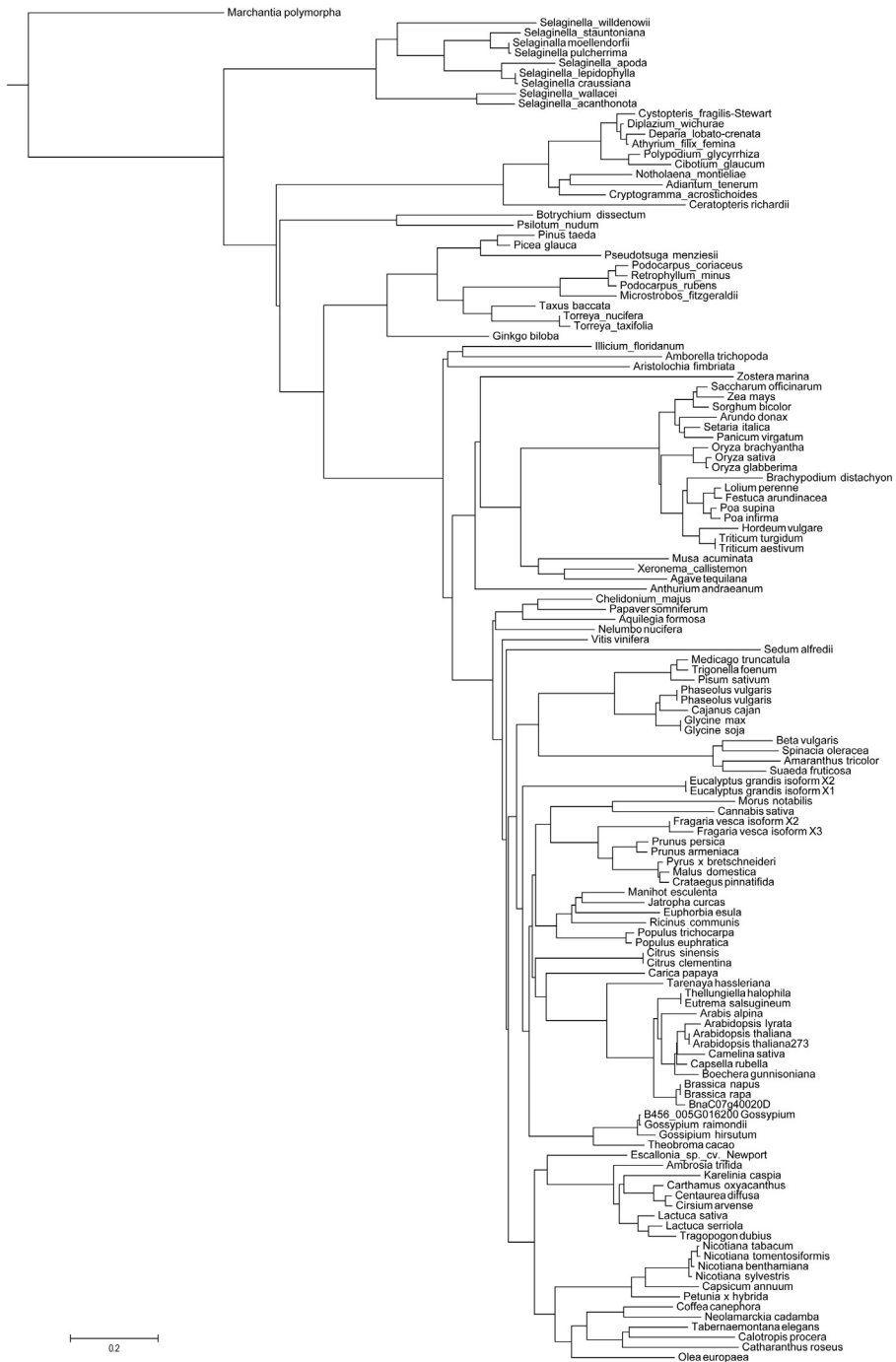


Fig. 2. The phylogenetic tree based on analysis of the 134 aligned 4/1 proteins from land plants. Fast Minimum Evolution tree was obtained at <http://www.ncbi.nlm.nih.gov/tools/cobalt/> with the use of default parameters.

Data source location	NCBI: <a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a> TREECON: <a href="http://bioinformatics.psb.ugent.be/downloads/psb/Userman/treeconw.html">http://bioinformatics.psb.ugent.be/downloads/psb/Userman/treeconw.html</a> 1KP: <a href="http://www.onekp.com">http://www.onekp.com</a>
Data accessibility	<i>With this article</i>

---

### Value of the data

- The 4/1 genes are low-copy genes and its molecular evolution is intriguing to understand their phylogenies among plants.
- Data on phylogenies separately estimated using Fast Minimum Evolution and neighbour-joining methods enable researchers to examine how the topologies differ from each other.
- Data on phylogenies of 4/1 proteins is intriguing to understand their unique features.
- Data on phylogenies of 4/1 proteins enable researchers to infer the possible ranges of time frames in the divergence events of 4/1 low-copy genes and its molecular evolution in general.

### 1. Data, experimental design, materials and methods

The phylogenetic tree obtained using COBALT (<http://www.ncbi.nlm.nih.gov/tools/cobalt/>) Fast Minimum Evolution method for 4/1 proteins from 62 plant species was presented in [1]. The data shown here represent the phylogenetic tree of 62 sequences of 4/1 proteins separately reconstructed using a neighbour-joining method by TREECON 1.3b package (Fig. 1) and combined phylogenetic tree for 134 plant 4/1 proteins constructed by a COBALT Fast Minimum Evolution method (Fig. 2). All the sequence data used in this data article were retrieved from NCBI (<http://www.ncbi.nlm.nih.gov/>) and 1KP databases (<http://www.onekp.com>). These sequences were aligned by NCBI protein Multiple Alignment Tool software using default parameters.

### Acknowledgements

We thank the researchers who contributed samples used in this study to the 1KP initiative. Database searches and phylogenetic analysis by the Fast Minimum Evolution method was performed by A.S. and S.M. in Moscow State University with financial support of the Russian Science Foundation (Grant 14-14-00053). Phylogenetic analysis by neighbour-joining method was performed by A.T. in Moscow State University with support of the Russian Foundation for Basic Research (Grant 15-04-06027).

### Reference

- [1] S.Y. Morozov, I.A. Milyutina, V.K. Bobrova, D.Y. Ryazantsev, T.N. Erokhina, S.K. Zavriev, A.A. Agranovsky, A.G. Solovyev, A. V. Troitsky, Structural evolution of the 4/1 genes and proteins in non-vascular and lower vascular plants, *Biochimie* 119 (2015) 125–136.