Phylogenetics

SPECTRE: a suite of phylogenetic tools for reticulate evolution

Sarah Bastkowski1,*,†, Daniel Mapleson1,†, Andreas Spillner2, Taoyang Wu3, Monika Balvočiūtė4 and Vincent Moulton3

1Earlham Institute, Norwich Research Park, Norwich NR4 7UZ, UK, 2Merseburg University of Applied Sciences, 06217 Merseburg, Germany, 3School of Computing Sciences, University of East Anglia, Norwich NR4 7TJ, UK and 4Department of Computer Science, University of Tübingen, Sand 14, 72076 Tübingen, Germany

*To whom correspondence should be addressed.
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.
Associate Editor: Alfonso Valencia

Received on June 6, 2017; revised on October 13, 2017; editorial decision on November 11, 2017; accepted on November 23, 2017

Abstract

Summary: Split-networks are a generalization of phylogenetic trees that have proven to be a powerful tool in phylogenetics. Various ways have been developed for computing such networks, including split-decomposition, NeighborNet, QNet and FlatNJ. Some of these approaches are implemented in the user-friendly SplitsTree software package. However, to give the user the option to adjust and extend these approaches and to facilitate their integration into analysis pipelines, there is a need for robust, open-source implementations of associated data structures and algorithms. Here, we present SPECTRE, a readily available, open-source library of data structures written in Java, that comes complete with new implementations of several pre-published algorithms and a basic interactive graphical interface for visualizing planar split networks. SPECTRE also supports the use of longer running algorithms by providing command line interfaces, which can be executed on servers or in High Performance Computing environments.

Availability and implementation: Full source code is available under the GPLv3 license at: https://github.com/maplesond/SPECTRE. SPECTRE’s core library is available from Maven Central at: https://mvnrepository.com/artifact/uk.ac.uea.cmp.spectre/core. Documentation is available at: http://spectre-suite-of-phylogenetic-tools-for-reticulate-evolution.readthedocs.io/en/latest/

Contact: sarah.bastkowski@earlham.ac.uk

Supplementary information: Supplementary data are available at Bioinformatics online.

1 Introduction

Split-networks are a generalization of phylogenetic trees that are commonly used to analyze reticulate evolution in organisms such as plants, bacteria and viruses (see Fig. 1 for an example). They provide a snapshot of the data and can be used to display conflicting signals. Examples of algorithms for computing such networks include split-decomposition (Bandelt and Dress, 1992), NeighborNet (Bryant and Moulton, 2004), QNet (Gruenewald et al., 2007), SuperQ (Gruenewald et al., 2013) and FlatNJ (Balvočiūtė et al., 2014). A comprehensive overview of split-networks can be found in (Huson and Bryant, 2006). Currently, the main program available for computing split-networks is the user-friendly SplitsTree program (Huson and Bryant, 2006). In addition, various methods for computing split-networks such as some of those mentioned above have been implemented and released as standalone applications. Implementing data structures capable of representing the mathematical structures used to describe and compute split networks is not a trivial undertaking and existing software either is closed source or have their data structures and algorithms tightly integrated with their host tool, so are not easily reusable. There are, therefore, currently few options for developers wishing to create or extend their own tools based on these concepts other than to start from scratch. Hence, there is a need for a robust and flexible open-source library that...
NetME produces a minimum evolution tree compatible with an
FlatNJ constructs a flat split network from a multiple sequence
SuperQ constructs a circular split network from a set of (partial)
NeighborNet rapidly constructs a circular split network from a
•
details see Section 1 of Supplementary Material):
- NeighborNet rapidly constructs a circular split network from a
distance matrix or a sequence alignment (Bryant and Moulton,
2004). NetMake implements variants of NeighborNet as
described in (Levy and Pachter, 2011).
- SuperQ constructs a circular split network from a set of (partial)
input trees (Gruenwald et al., 2013).
- FlatNJ constructs a flat split network from a multiple sequence
alignment, weighted quartet data or location data (Balvočiutė
et al., 2014).
- NetME produces a minimum evolution tree compatible with an
existing circular split network (Bastkowski et al., 2014).

These tools are accessible to the user via graphical and command
line interfaces. Apart from driving the tools, the interactive graphical
interface can visualize planar split networks using the drawing algo-
rimth in (Spillner et al., 2012). The interface offers a number of
basic functions for orientating the canvas (e.g. zoom, pan, flip and
rotate), manipulating labels (size, color, location) and creating
image files (PDF, EPS, SVG, PNG). The command line implement-
tion enables bioinformaticians to integrate tools into pipelines. This
works on desktop PCs, like SplitsTree, but is designed so long
running tools are executable on servers or high performance comput-
ing environments where displays are not available. For de-
velopers wishing to reuse code and develop their own tools, SPECTRE
provides a core library containing common data structures (e.g.
splits, trees, networks, distances, quartets and multiple sequence
alignments), algorithms (e.g. NeighborNet) and robust file parsers
to process a range of input files (e.g. NEXUS, PHYLIP, Newick,
Emboss, FastA); see Section 2 of Supplementary Material for more
details. The library is available directly from Maven Central, giving
developers direct access to the most recent version of the library and
and providing a convenient way to integrate it into the processes for
building their own projects.

2 SPECTRE
Here, we present SPECTRE, a suite of tools for computing, mod-
eling and visualizing reticulate evolution based on split-networks.
SPECTRE builds in part on existing open-source implementations of
some of these tools, in particular for QNet, SuperQ and FlatNJ,
integrating them into a unified and extendible library. The main
tools available through SPECTRE are summarized below (for more
details see Section 1 of Supplementary Material):

- NeighborNet rapidly constructs a circular split network from a
distance matrix or a sequence alignment (Bryant and Moulton,
2004). NetMake implements variants of NeighborNet as
described in (Levy and Pachter, 2011).
- SuperQ constructs a circular split network from a set of (partial)
input trees (Gruenwald et al., 2013).
- FlatNJ constructs a flat split network from a multiple sequence
alignment, weighted quartet data or location data (Balvočiutė
et al., 2014).
- NetME produces a minimum evolution tree compatible with an
existing circular split network (Bastkowski et al., 2014).

These tools are accessible to the user via graphical and command
line interfaces. Apart from driving the tools, the interactive graphical
interface can visualize planar split networks using the drawing algo-

3 Concluding remarks
SPECTRE provides a collection of open-source tools and resources
for modelling, understanding and visualizing reticulate evolution
based on split networks. We believe that our software will both en-
able bioinformaticians to easily test and compare methods for infer-
ring planar split networks and help computer scientists build their
own methods for inferring phylogenetic networks by reusing our
existing data structures and algorithms via the open-source library.
Moreover, this also provides the option to easily add such new tools
to the library making them readily available to other users.

Acknowledgement
The authors would like to thank Stephan Gruenewald.

Conflict of Interest: none declared.

References
Balvočiutė, M. et al. (2014) FlatNJ: a novel network-based approach to visualize
evolutionary and biogeographical relationships. Syst. Biol., 63, 383–396.
 Bandelt, H.-J. and Dress, A. (1992) Split decomposition: a new and useful approach
to phylogenetic analysis of distance data. Mol. Phylogen. Evol., 1, 242–252.
Bastkowski, S. et al. (2014) Fishing for minimum evolution trees with
NeighborNets. Inform. Process. Lett., 114, 13–18.
Bollyky, P. L. et al. (1996) Recombination between sequences of hepatitis B
virus from different genotypes. J. Mol. Evol., 42, 97–102.
Bryant, D. and Moulton, V. (2004) NeighborNet: an agglomerative method for
the construction of phylogenetic networks. Mol. Biol. Evol., 21, 255–265.
Gruenwald, S. et al. (2007) QNet: an agglomerative method for the construction of
phylogenetic networks from weighted quartets. Mol. Biol. Evol., 24, 532–538.
Gruenwald, S. et al. (2013) SuperQ: computing supernetworks from quartets.
IEEE/ACM Trans. Comput. Biol. Bioinform., 10, 151–160.
Huson, D. and Bryant, D. (2006) Application of phylogenetic networks in evo-
lutionary studies. Mol. Biol. Evol., 23, 254–267.
Levy, A. and Pachter, L. (2011) The Neighbor-Net algorithm. Adv. Appl.
Math., 47, 240–258.
Spillner, A. et al. (2012) Constructing and drawing regular planar split net-
works. IEEE/ACM Trans. Comput. Biol. Bioinform., 9, 395–407.