
phylogenetics Documentation

Release 0.4.2

Zach Sailer

Oct 03, 2019

Contents

1	Development goals	3
2	Basic Example	5
3	Installation	7
4	Dependencies	9
5	Table of Contents	11

Phylogenetics is a minimal Python API for doing phylogenetics. It manages the annoying aspects of phylogenetics (i.e. file conversion) for you and lets you focus on exploring and interpreting the data.

CHAPTER 1

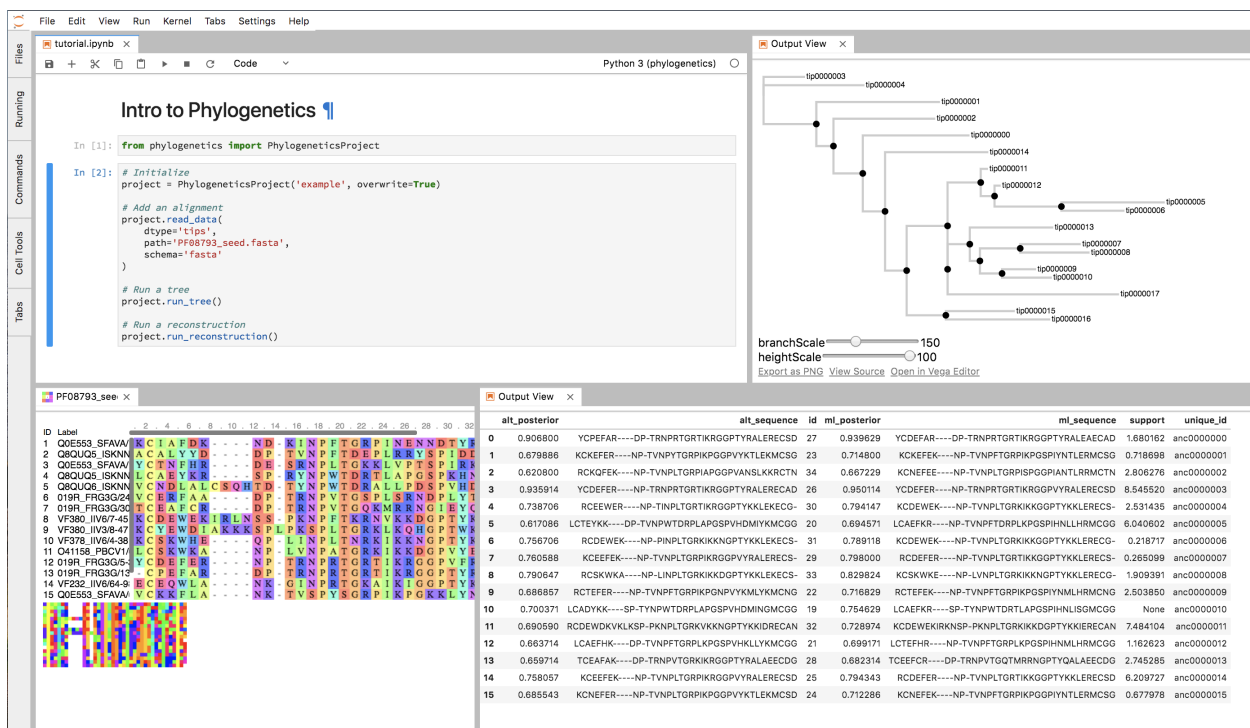
Development goals

- **docs:**
 - installation
 - **specific cases**
 - * long branches
 - * what does branch support mean?
 - * sequence quality control
 - * alignment quality
 - * ancestors, posterior probabilities
 - * how do I come up with an answerable evolutionary question?
 - **design philosophy**
 - * why do we need another package?
 - * as little low-level crap as possible (use Biopython, dendropy, etc.), let users interact simply via familiar csv and pandas df
- **to implement:**
 - **phylopandas (critical):**
 - * uid/csv awareness
 - * tree/align integration (Zach)
 - species correction
 - datatype awareness (dna, rna, protein, codon)
 - phylobot mk 2?

CHAPTER 2

Basic Example

TO DO



CHAPTER 3

Installation

Install from PyPi:

```
pip install phylogenetics
```

To install a development version:

```
git clone https://github.com/Zsailer/phylogenetics
cd phylogenetics
pip install -e .
```


CHAPTER 4

Dependencies

phylogenetics manages phylogenetics data. Currently, it doesn't do any of the phylogenetic calculations itself. For this, use external tools like:

1. **PhyML**: building maximum likelihood trees.
2. **PAML**: reconstructing ancestors.

phylogenetics is built on top of following python stack:

1. **DendroPy**: A Python library for phylogenetic scripting, simulation, data processing and manipulation.
2. **ToyTree**: A minimalist tree plotting library using toyplot graphs
3. **PhyloPandas**: Pandas for phylogenetics
4. **PyASR**: Ancestral Sequence Reconstruction in Python

5.1 Installing PhyML and PAML

It's important that the following executables are callable from any location for `phylogenetics` to work properly with external tools:

1. **PhyML_**: building maximum likelihood trees.
2. **PAML**: reconstructing ancestors.

5.1.1 PhyML

Get the latest, stable release of '**PhyML_**'

Download, unpack, and install PhymL. Export path to the `src` file in `bashrc` file.

5.2 TreeProject API

The following pages is the TreeProject API documentation. The TreeProject is the main class you'll use in the `phylogenetics` package.

5.2.1 TreeProject Class