



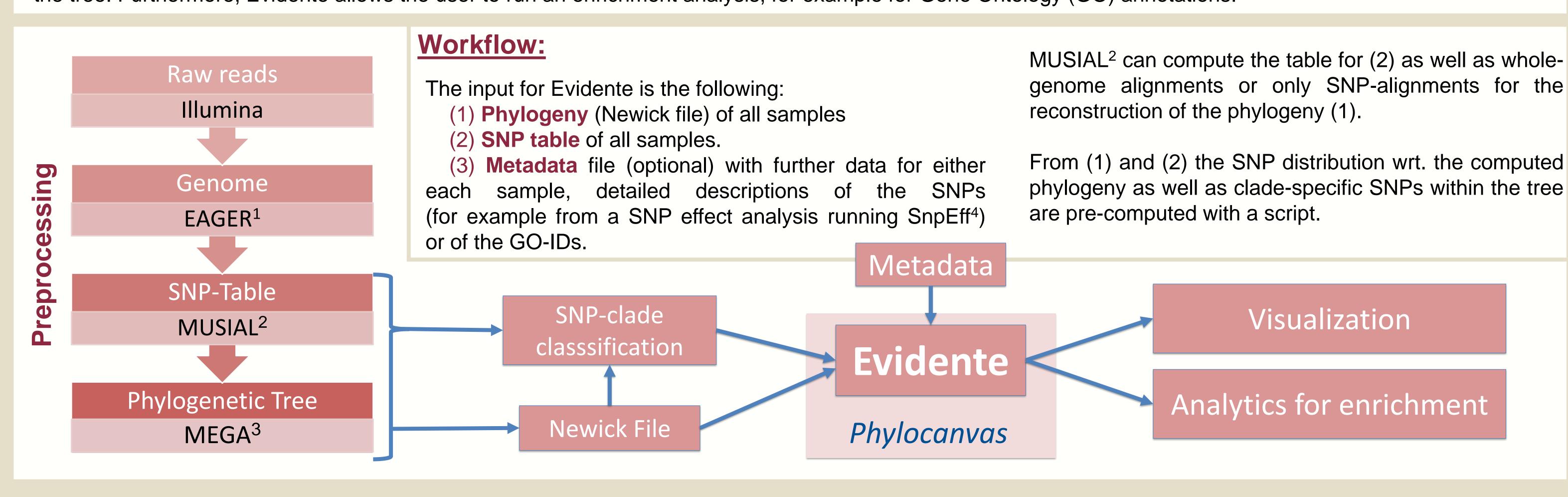
EVIDENTE facilitates the identification of enriched characteristics in SNP-based phylogenetic trees

Mathias Witte Paz, Alexander Seitz, Kay Nieselt

Institute for Bioinformatics and Medical Informatics University of Tübingen Sand 14, 72076 Tübingen



Motivation for EVIDENTE: In recent years the developments of the next-generation sequencing technologies have enabled genome resequencing projects of many individuals within one species. The genomes are often analyzed with respect to single-nucleotide polymorphisms (SNPs) or small indels. This gives the possibility of reconstructing a phylogenetic tree of all individuals based on the detected mutations. From such a phylogenetic tree, a common question is to identify clade-specific SNPs within the reconstructed phylogeny, i.e. that support the computed topology. Then one also often wishes to analyze these mutations in more detail to retrieve for example functional consequences that the SNP may have on the organism or to compute enrichment of certain features within the phylogenetic tree. Here, we present on-going work in developing the visual analytics tool Evidente for annotation and analysis of metadata in SNP-based phylogenetic trees. It enables the user to get a visual overview of distribution of SNPs across all samples as well as clade-specific SNPs within the tree. Furthermore, Evidente allows the user to run an enrichment analysis, for example for Gene Ontology (GO) annotations.



Use case: MI10-99_Mali_2016 Node Search MI10-95_Mali_2016 Mycobacterium leprae⁵ MI10-97_Mali_2016 MI10-93_Mali_2016 Origin Place: Ng12-33_Niger_2015 Mali, Nigeria Ng17-39_Niger_2015 STATE OF THE PROPERTY OF THE P MI9-86_Mali_2014 Ng15-37_Niger_2015 MI10-91_Mali_2016 S13_Mali_2012 Ng13-32_Niger_2015 Ng15-36_Niger_2015 No Of Infections Ng16-38_Niger_2015 MI10-94_Mali_2016 Filter Nodes found: 17 Export to file. MI9-79_Mali_2016 MI9-81_Mali_2014 SNP Search Terral parity and Supporting SNPs US57_Marshall_Islands_2000 Non-Supporting SNPs Ryukyu-2_Japan_2000 Zensho-9_Japan_2000 CM1_Philippines_1994 Apaper, agai, min man, agai, min man, agai, min S9 New Caledonia 1996 1987172 Jorgen_507_Denmark_1058-1253 SK11_Hungary_600-800 S10_China_2006 Files Upload Izumi_Japan_2000 Tree Options Node Search Kanazawa_Japan_2000 Oku-4_Japan_2000 Kyoto-1_Japan_2000 Kyoto-2_Japan_1991 Korea-3-2_Korea_2000 Tsukuba-1_Japan_2000 2_Japan_2000 3. SNP3: 1987172 😵 **Enrichment Analysis on SNPs** Subtree ID: 335 Significance level SNP Characteristic Gene Name Enrichment namena marina and program or more agreement **Enrichment** - Angalianja amanjaran ar - Angalianja amang jawa ar - Angalianjaran - Maranjaran - Maranjaran - Maranjaran The following p-values have been corrected Fischer's Exact Test runned as logarithms. Visualization of Comparisons made: 205, corrected significance level: 0.000244 analysis enriched gene on genes Printer Marches Printer Marches Printer Marches Printer Marches Marche

Features:

Filtering:

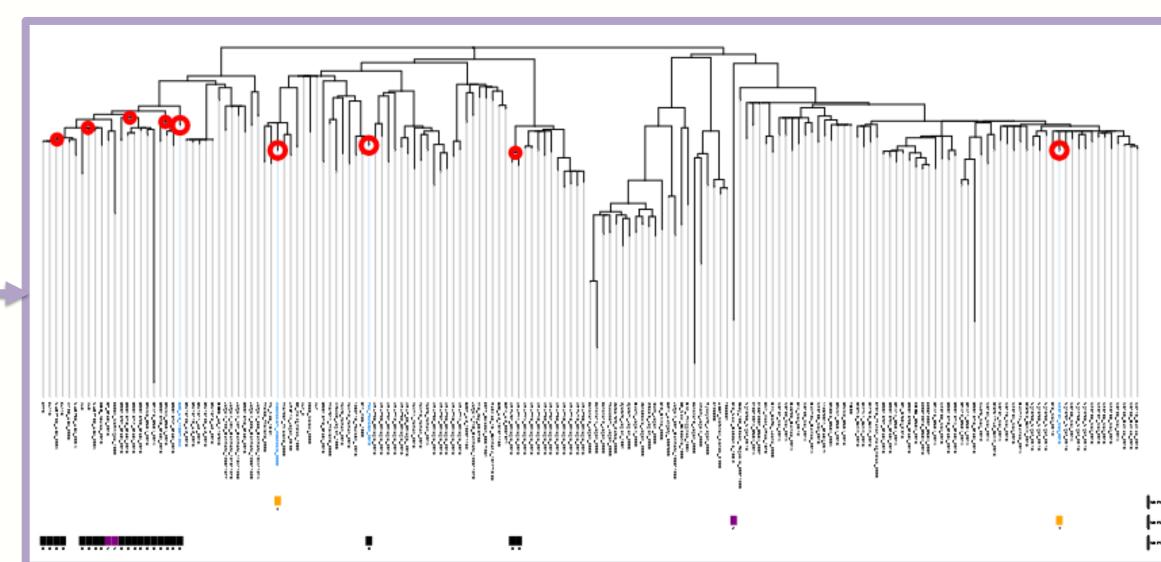
Through metadata of the samples

SNP Visualization:

- Clade-specific SNPs
- Non-supporting SNPs
- Common SNPs between subtrees

Enrichment Analysis:

- Fisher's Exact Test
- Enrichment of SNP characteristics
- Enrichment of taxonomic information
- Visualization of enriched features



References

- 1) Peltzer, A. et al. EAGER: efficient ancient genome reconstruction. Genome Biol 17:60 (2016).
- 2) https://github.com/Integrative-Transcriptomics/MUSIAL
- 3) Kumar, S. et al. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets, Molecular Biology and Evolution, 33: 1870-1874 (2016).

A CONTRACTOR

4) Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly 6:80-92 (2012). 5) Schuenemann, VJ. et al. Ancient genomes reveal a high diversity of Mycobacterium leprae in medieval Europe. PLOS Pathogens 4: e1006997(2018).

Availability:



https://lambda.informatik.uni-tuebingen.de/gitlab/paz/evidente

Poster download!