Phylogenies derived from somatic mutations agree with physical topologies in *Eucalyptus*

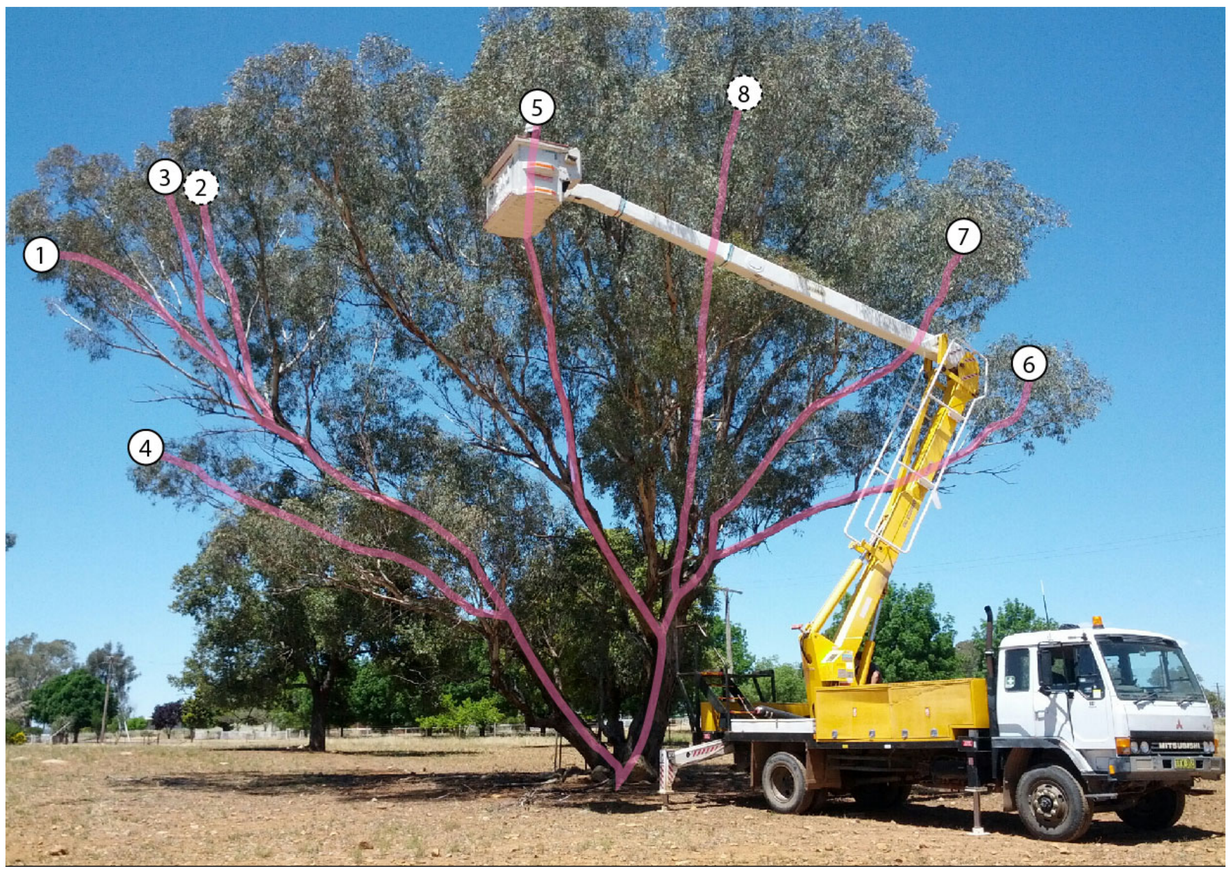
Adam J Orr [1,2] Robert Lanfear [3] Reed Cartwright [1,2]

[1]School of Life Sciences, Arizona State University
[2]Biodesign Institute, Arizona State University
[3]College of Medicine, Biology and Environment, Australian National University

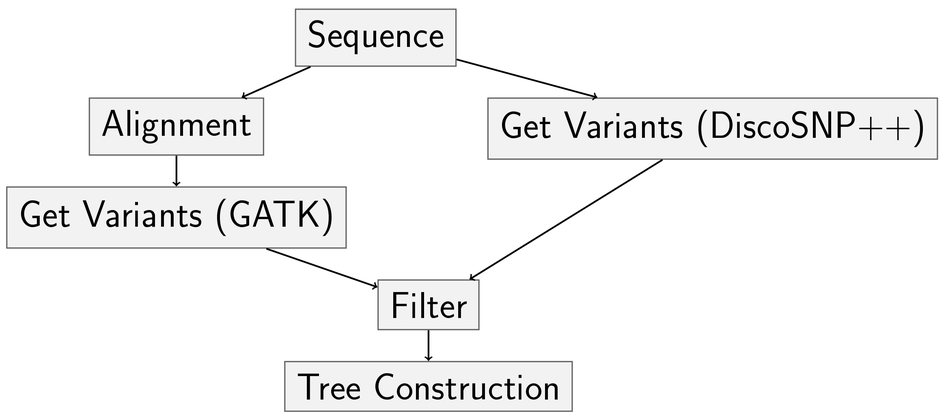
Introduction

- Somatic mutations are rare but sequencing errors are common, making somatic mutations difficult to detect.
- Little is known about the spread of somatic mutations, despite the key role they play in cancer development.
- If the pattern of mutations in the phylogeny matches the branching pattern of the plant, then plants can be used to easily validate somatic mutations.

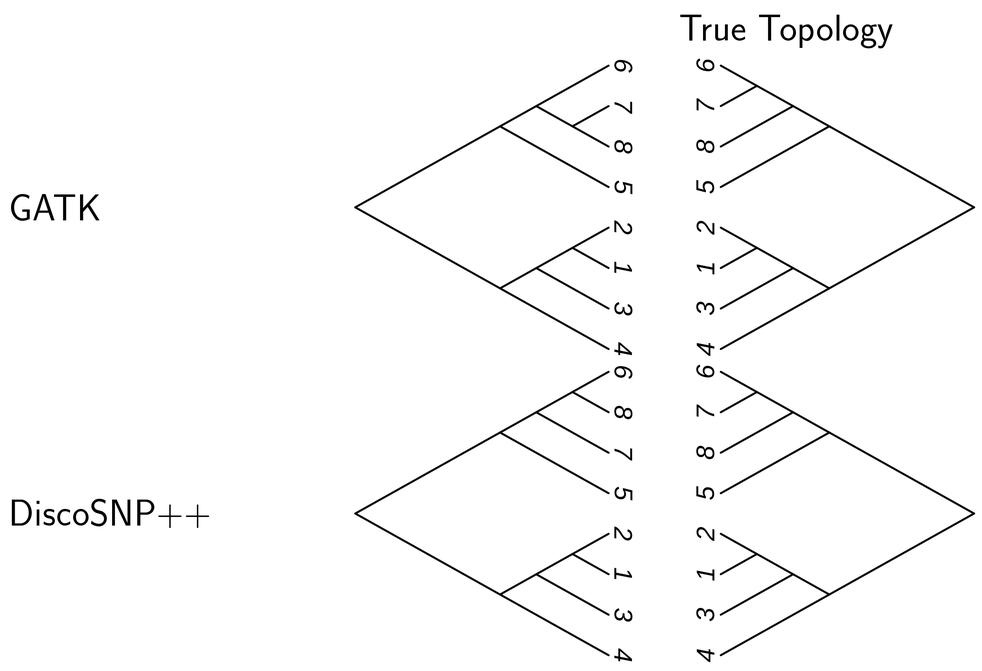
Methods: Variant Detection



- 8 samples collected in triplicate
- Variants were removed if the genotypes of all replicates of a sample were not identical



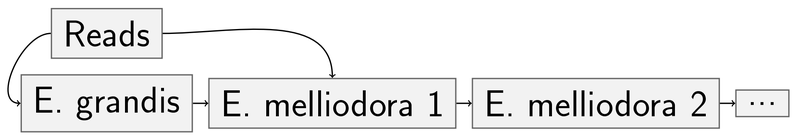
Results: Variant Detection



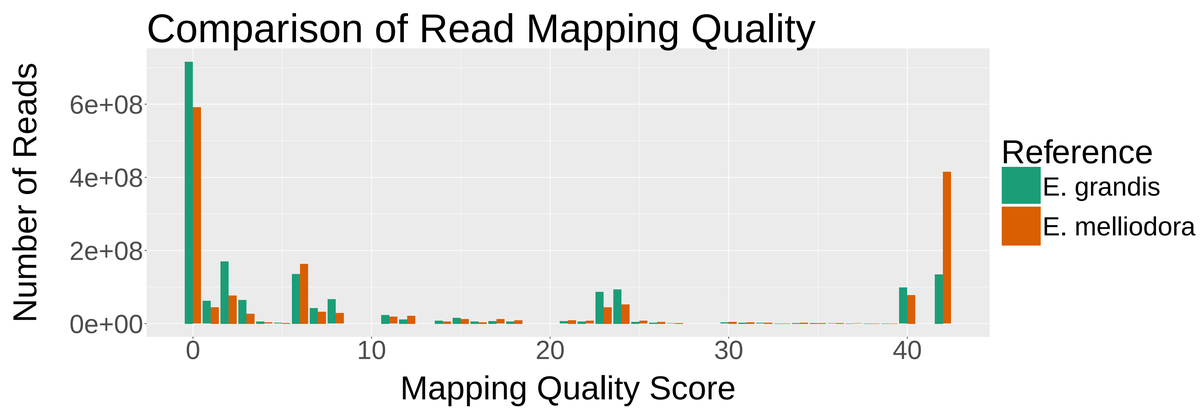
Next steps: Reference Improvement

To improve resolution of short internodes, we attempt to modify the *E. grandis* genome to make it more suitable for *E. melliodora* by:

- Aligning the reads to the *E. grandis* genome
- Creating a consensus sequence from this alignment
- Aligning the reads to this consensus to create a draft *E. melliodora* genome

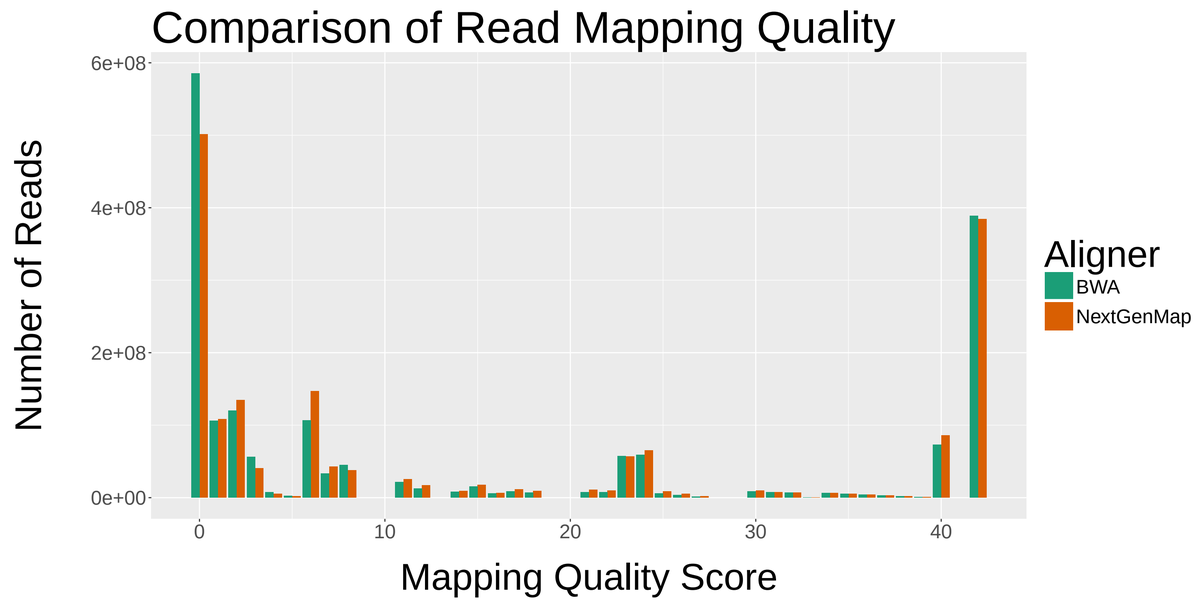


Results: Reference Improvement



Alignment to the consensus sequence produces an alignment with higher overall quality scores than alignment to the *Eucalyptus grandis* reference.

The Choice of Mapper Affects Generated Reference Quality



Conclusions

- Phylogenies of somatic mutations within a *Eucalyptus* tree match the branching patterns of the tree using both a reference-based and a reference-free variant caller.
- Aligning reads to a close relative, obtaining a consensus sequence, then realigning to that consensus seems to improve alignment quality.
- The choice of aligner makes a difference.

Acknowledgements

This work is supported by grant NIH R01-HG007178.