

Gap (indel) models

Using gap characters

GARLI implements two models appropriate for use on gap characters in *fixed alignments*

DIMM model based on Rivas and Eddy (2008)

- Real model of insertion-deletion process
- Non-reversible, dollo (one insert, multiple delete)

Variant of Mk_v (Lewis, 2001) model (no full gap columns)

- Allows many gap-base transitions

gapped alignment



**ignore gaps, e.g. GTR
(the usual)**

gapped alignment



5-state approach (e.g., Rivas and Eddy)

**ignore gaps, e.g. GTR
(the usual)**

gapped alignment



5-state approach

(e.g., Rivas and Eddy 2008)



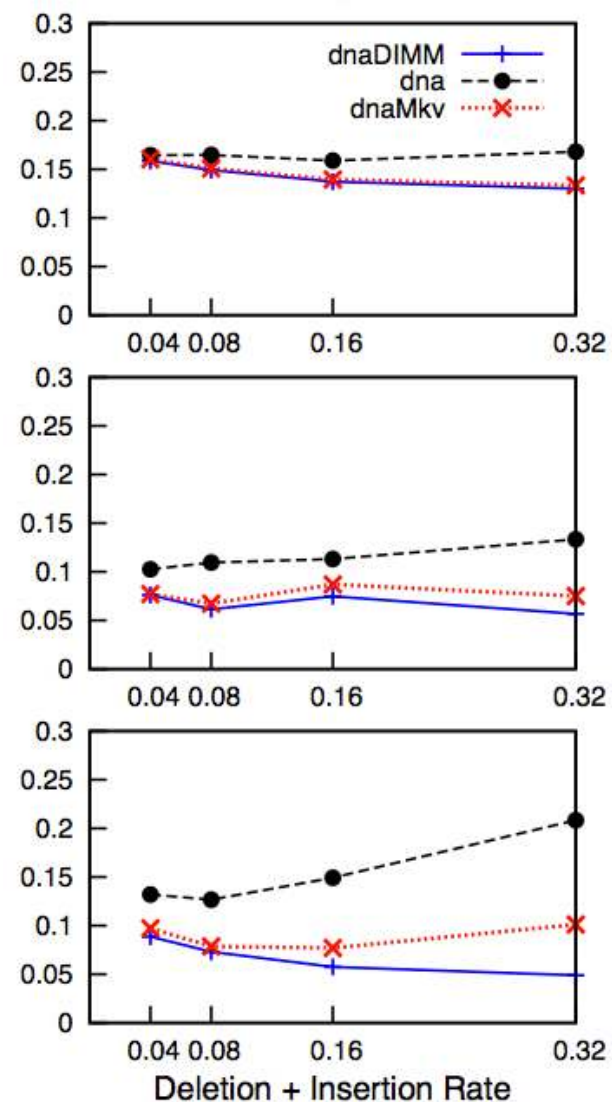
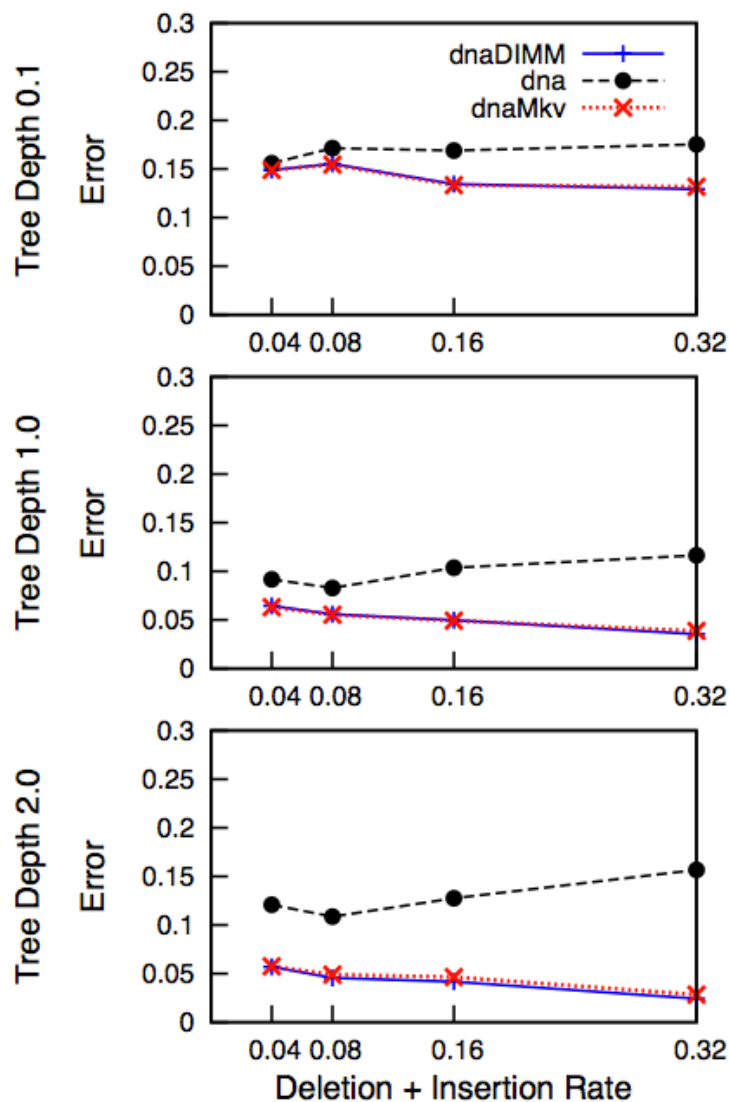
GTR

DIMM
(or Mk_v)

Gap model accuracy – 64 taxa, true alignments

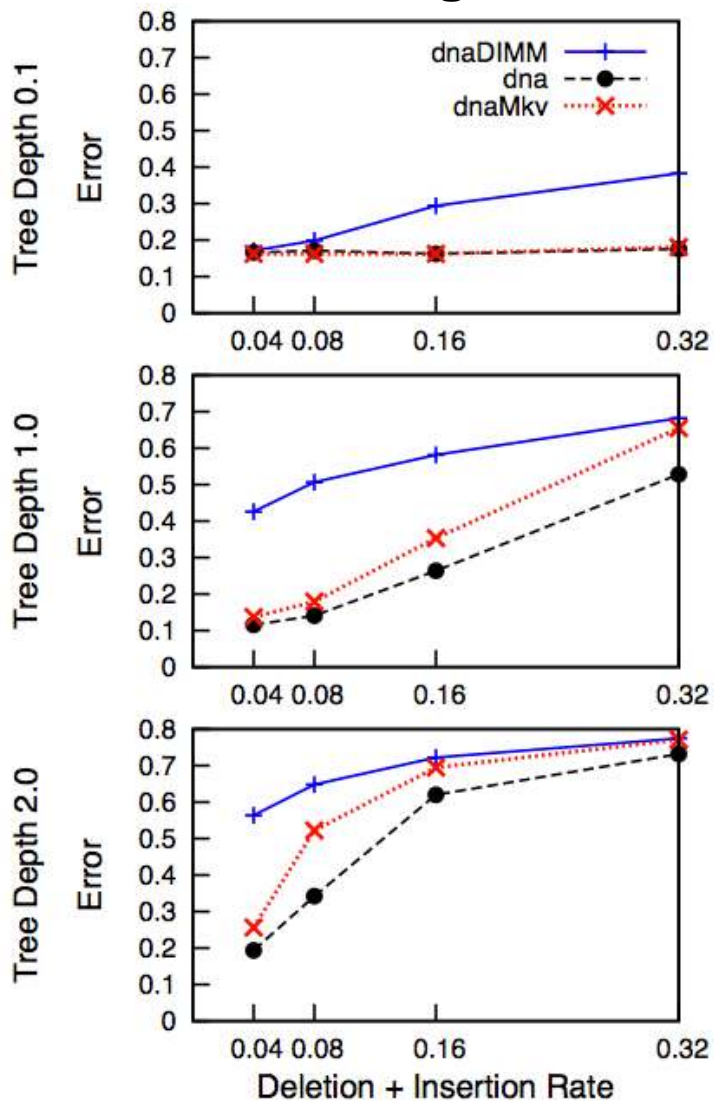
single site indels

multi-site indels

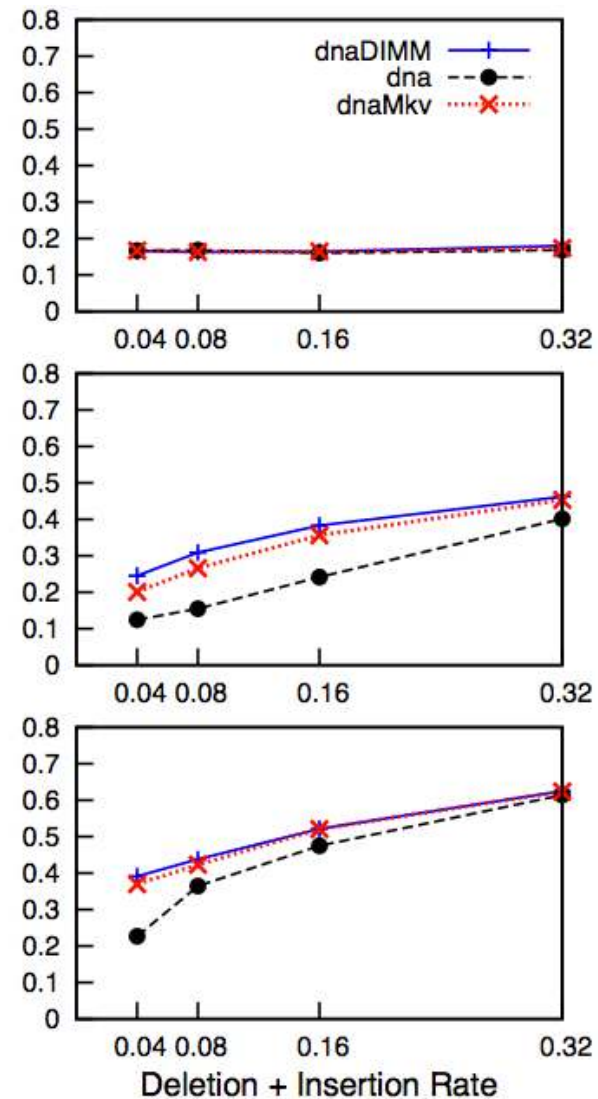


Gap model accuracy – 64 taxa, estimated alignments

MAFFT alignment



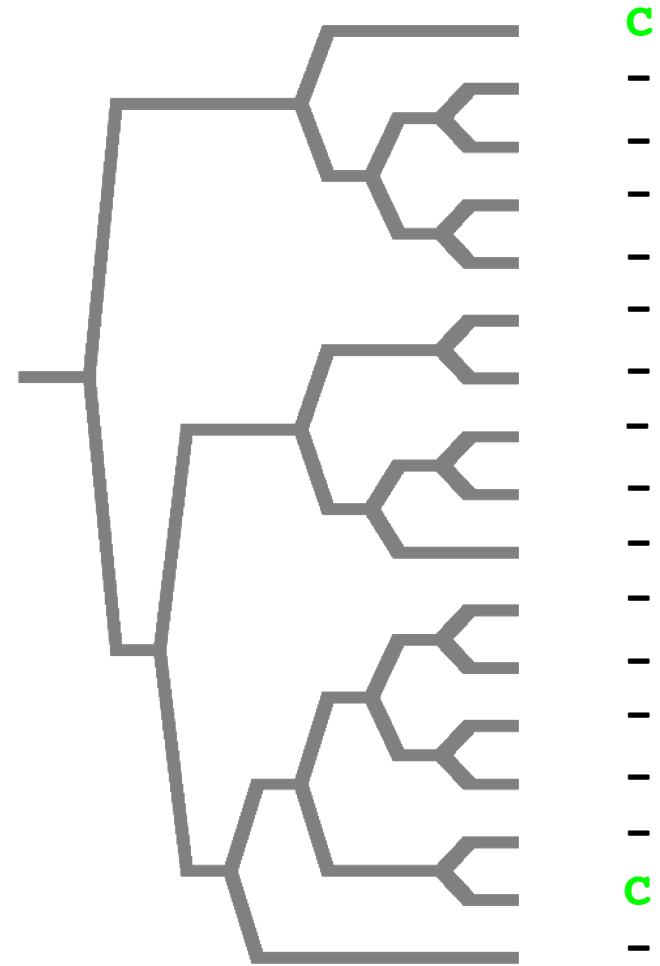
PRANK NJ-F Alignment



Part of the problem: innocuous alignment errors

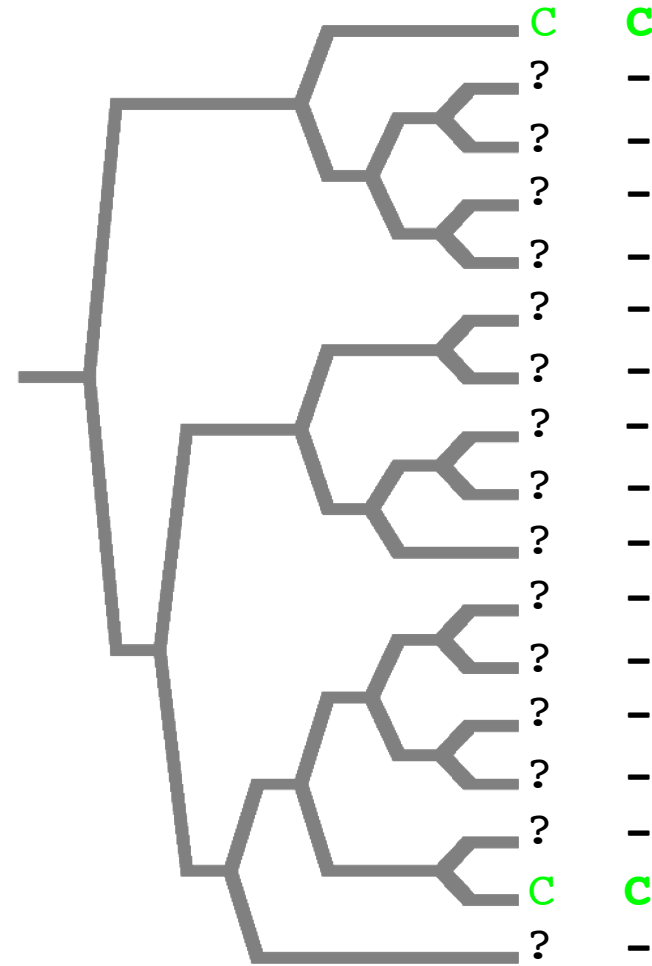
[illegible]

Innocuous alignment errors?



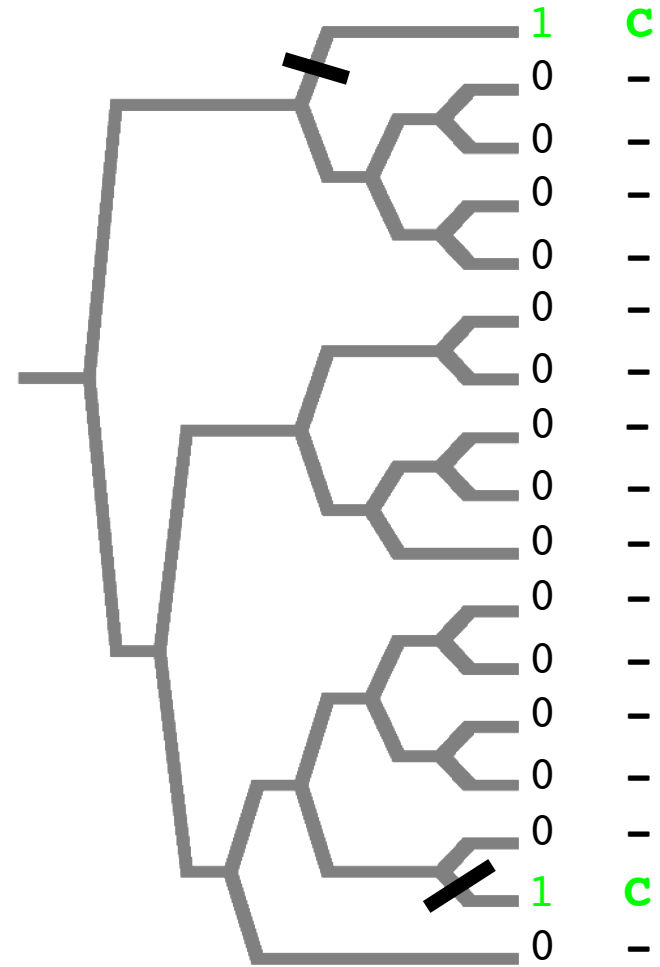
Innocuous alignment errors?

- GTR – no events, little effect



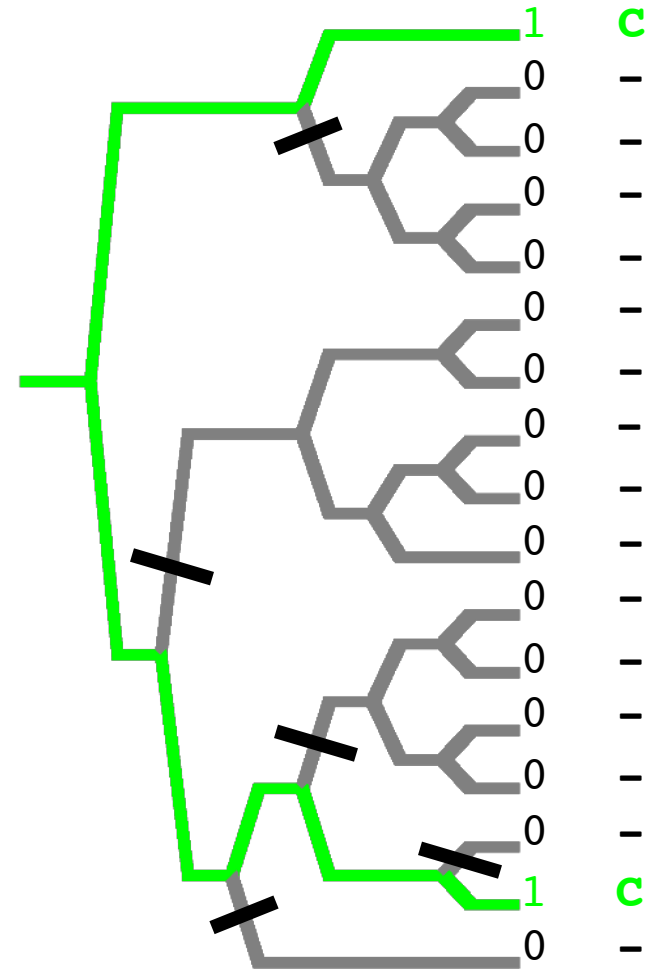
Innocuous alignment errors?

- GTR – no events, little effect
- $Mk_v - 2$ events, moderate effect



Innocuous alignment errors?

- GTR – no events, little effect
- Mk_v – 2 events, moderate effect
- DIMM – 5 deletions! strong effect (Dependent on tree size!!)



Conclusions: gap models

Indel events can provide useful signal *if* properly modeled

BUT, without correct column homology they can do more evil than good