**Update ProPIP**

**ProPIP** Alignment gave me the following responses to INDELible, PIP, real data inputs.

**INDELible data**

* 16 missing alignments
* **Reason/Error:**
  + - 13 sequences with complete gap leaves.
    - Tree branch length is too small.
    - Something wrong with TR value. TR is neither match, not GAPX, nor GAPY.
    - Unknown error (no error file provided at the output)

**PIP data**

* 1 missing alignment
* **Reason/Error:**
  + - Tree branch length is too small.

**Real data**

* 1 missing alignment- Solved
* **Reason/Error:**
  + - Handling ‘**?’, ‘X’** in the sequence.
    - **Solution:** contained one ‘?’ in the sequence. PRANK and MAFFT modified it as ‘-‘ , So I changed it manually and run the modified version in ProPIP also.

Dear all,

I re-simulated all with complete branch lengths (previously I rounded it to avoid exponential representation error). This was something making error in ProPIP. After making these changes I aligned the sequence using PRANK, MAFFT and now running ProPIP.

**Meeting 07.04.2020**

1. Do we have to generate more data to do the statistics? (We decided to exclude the sequences with blank leaves from INDELible).
2. Steps in Statistical analysis.

I would like to mention some comparisons we planned before here:

**C1 Does PIP adopt to long Indels?**

C1.1 ProPIP (INDELible data) vs ProPIP (PIP data)

C1.2 ProPIP (real data) vs ProPIP (PIP data)

C1.3 ProPIP (INDELible data) vs ProPIP (real data)

**C2 How do the Indel statistics differ?**

C2.1 ProPIP (real data) vs PRANK (real data)

C2.2 ProPIP (real data) vs MAFFT (real data)

C2.3 MAFFT (real data) vs PRANK (real data)

C2.4 ProPIP (INDELible data) vs PRANK (INDELible data)

C2.5 ProPIP (INDELible data vs MAFFT (INDELible data)

C2.6 MAFFT (INDELible data) vs PRANK (INDELible data)

A copy of Manuel’s note is provided for reference below.

A close up of text on a whiteboard

Description automatically generated