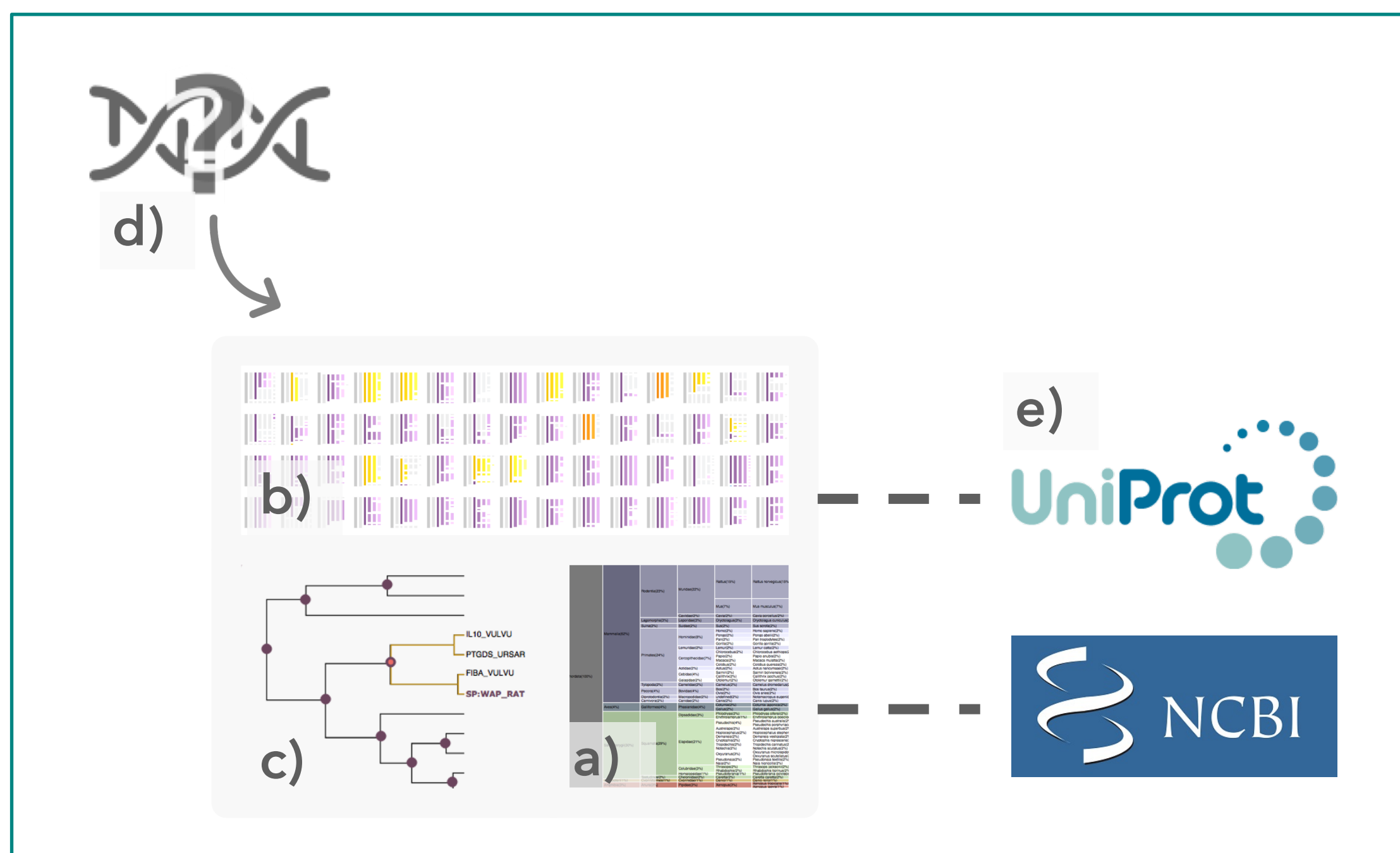


BioCicle

A visual analytics tool for summarize
and compare biological sequence
alignments



WHY BioCicle?

Biological Sequence Alignments detect regions of similarity that can lead to evolutionary relationships between sequences

Sequence Comparisons
produce myriads of results



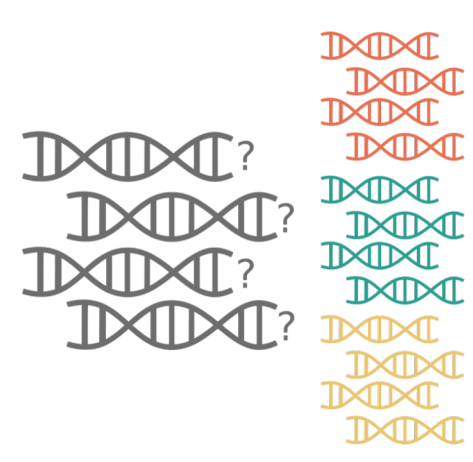
Misclassifications may lead to a widespread of erroneous data stored in sequence databases

Analysis Tasks

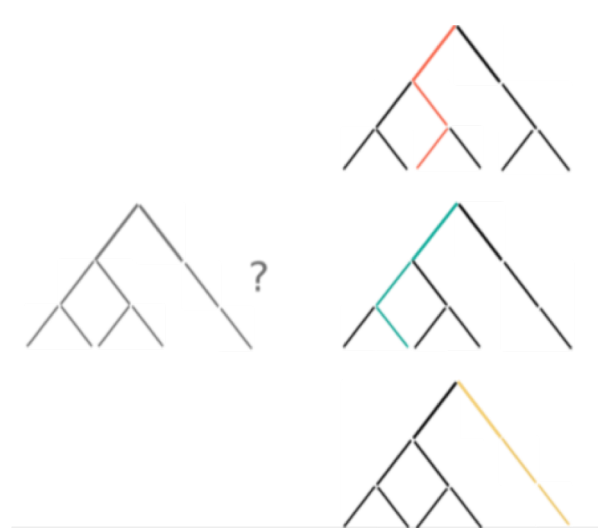
ATa
Summarize
single
alignment
results



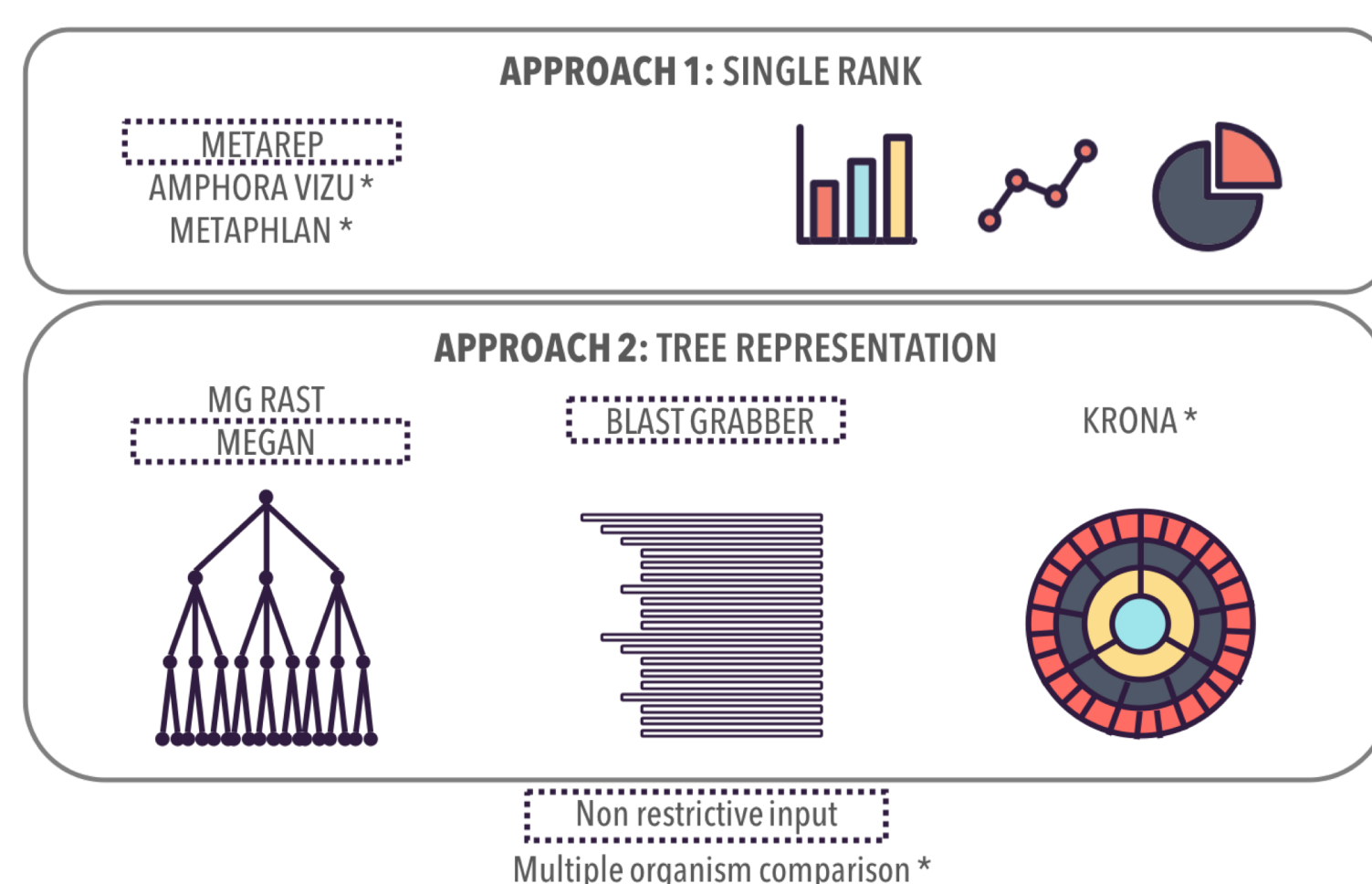
ATb
Compare
multiple
alignment
results



Taxonomic report analysis
help to trace species
carrying such sequence



Current approaches



IMPLEMENTATION

a) Icicle: Size to represent tree depth, and horizontal spatial position to show link relationships and sibling

b) Small Multiples: (ATb) Present an overview of the entire result set. Preserve colors in taxonomies.

c) Dendrogram: Filter results by taxonomic rank

d) Input formats: Supports usual input formats as FASTA, accession id or pregenerated results

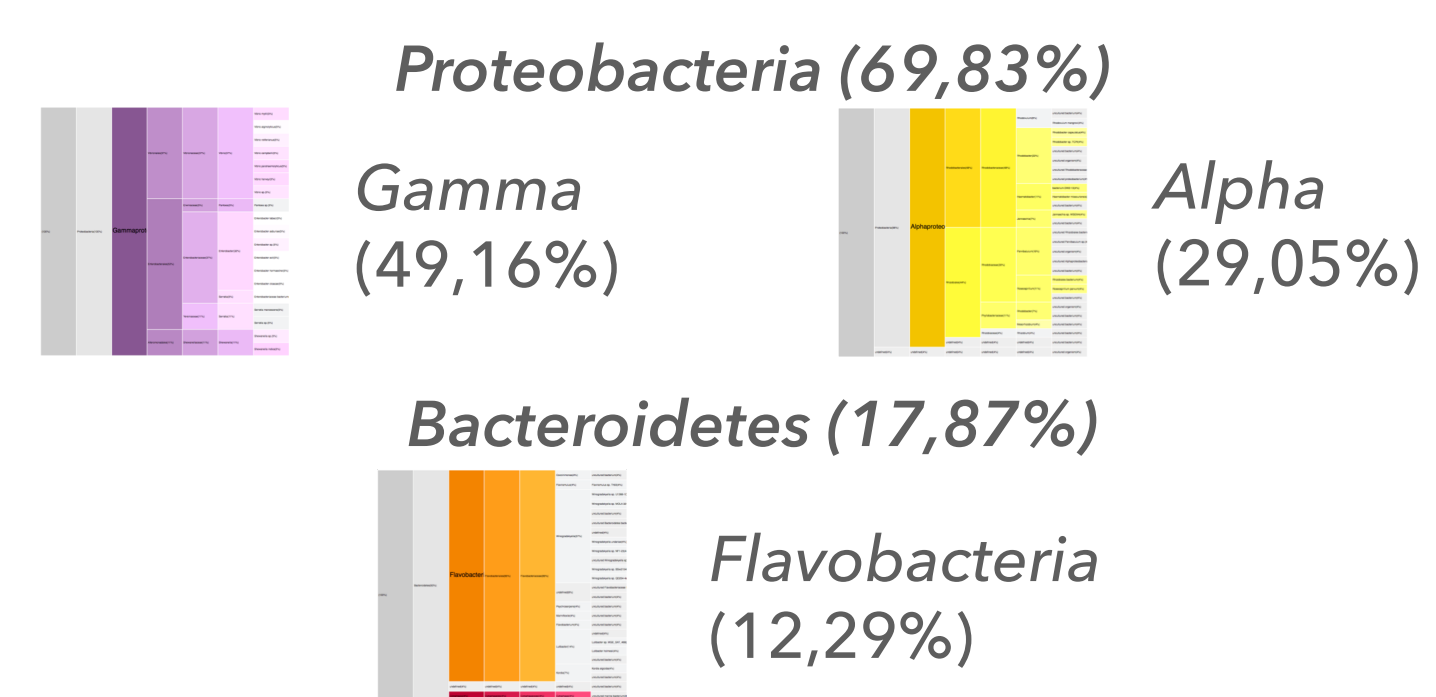
e) NCBI/EBI and UniProt API's: Directly connected with biological alignment frameworks and databses

RESULTS



179 nucleotide sequences
extracted from a single
sample

How can we describe the results?



CONCLUSIONS AND FUTURE WORK

We present a prototype as a proof of concept that:

- Supports taxonomic profile analysis (AT a,b)
- Web-based and open source
- Connects with the NCBI and the UniProt API
- Supports only protein sequence analysis

Future work should be focused in:

- Tackling sequence's description analysis as an extended framework
- Extending the use for nucleotide sequences



Meili Vanegas Hernández (m.vanegas10)

Advisors:

John Alexis Guerra Gómez (ja.guerrag),

José Tiberio Hernández Peñalosa (jhernand)

<https://github.com/mvanegas10/BioCicle>



<http://sistemas.uniandes.edu.co>



/DISCUniandes



/DISCUniandes