

# Benchmark Experiment

We will compare our method with some state-of-the-art Multiple Sequence Alignment (MSA) methods.

Methods	URL
T-Coffee	<a href="https://tcoffee.readthedocs.io/en/latest/tcoffee_installation.html">https://tcoffee.readthedocs.io/en/latest/tcoffee_installation.html</a>
Muscle	<a href="https://www.drive5.com/muscle/">https://www.drive5.com/muscle/</a>
MAFFT	<a href="https://mafft.cbrc.jp/alignment/software/">https://mafft.cbrc.jp/alignment/software/</a>
Clustal Omega	<a href="http://www.clustal.org/omega/">http://www.clustal.org/omega/</a>
Kalign	<a href="https://github.com/timolassmann/kalign">https://github.com/timolassmann/kalign</a>

## Datasets:

Datasets	URL	Paper	Length
Greengenes	<a href="#">Data</a>	<a href="#">Paper</a>	1325-1500
Zymo	<a href="#">Data</a>	<a href="#">Paper</a>	1325-1500
Silva_23S	<a href="#">Data</a>	<a href="#">Paper</a>	2750-3150
BRAlIbase	<a href="#">Data</a>	<a href="#">Paper</a>	????

### Greengenes:

16S rRNA sequence database. High diversity.

### Zymo:

16S rRNA sequence dataset. Only 8 different species. Low diversity.

### Silva\_23S:

23S rRNA sequence database. High diversity. Longer than 16S.

### BRAlIbase:

I have never used this. It is a dataset for benchmark MSA methods. There are 3 datasets on their website. I think the BRAlIbase III might be useful.