Pplacer Pipeline

1. Prepare Your Reference Alignment

1.1 Alignment with MAFFT.

```
mafft ref_filename.fasta > ref_filename.aln.fasta
```

1.2 Remove duplicate with Seqmagick.

```
seqmagick convert --deduplicate-sequences ref_filename.aln.fasta
ref_filename.aln.dedup.fasta
```

2. When You Have Your Alignment Ready

2.1 Build Tree with FastTree (Model Selected) and create a log file.

```
FastTree -log ref_filename.tree.log -nt -gtr ref_filename.aln.dedup.fasta >
ref_filename.tree
```

2.2 Make Reference Package with Taxtastic.

```
taxit create -l nod -P ref_filename.refpkg --aln-fasta ref_filename.aln.dedup.fasta --
tree-stats ref_filename.tree.log --tree-file ref_filename.tree
```

2.3 Convert Alignment format from Fasta to Stockholm format.

```
seqmagick convert ref_filename.aln.dedup.fasta ref_filename.aln.dedup.sto
```

2.4 Build HMM Profile with HMMER.

```
hmmbuild ref_filename.aln.dedup.hmm ref_filename.aln.dedup.sto
```

3. Prepare Your Query Sequence

3.1 Use HMM Profile to do an HMM Search on the Metatranscriptomics file and get output in .sto format.

```
hmmsearch -A query_filename.query.sto ref_filename.aln.dedup.hmm
query_filename.fasta
```

3.2 Use HMM Align to align query hits to the Reference Alignment.

```
hmmalign -o filename.combo.sto --mapali ref_filename.aln.dedup.sto
ref_filename.aln.dedup.hmm query_filename.query.sto
```

Adapted by Chalermpong Intarat, Nov 20, 2019

NBT, NSTDA chalermpong.int@biotec.ot.th

4. Placement: (Run Pplacer at /smirarab-sepp-53242af/tools/bundled/Darwin/pplacer.)

4.1 Run Pplacer using .refpkg.

```
pplacer -c ref_filename.refpkg filename.combo.sto
```

4.2 Run Guppy fat to make a fat phyloXML.

```
guppy fat filename.combo.jplace
```

4.3 Run Guppy tog to make a tog.

```
guppy tog filename.combo.jplace
```

4.4 Run Guppy tog to make a tog phyloXML.

```
guppy tog --xml filename.combo.jplace
```

Adapted by Chalermpong Intarat, Nov 20, 2019

NBT, NSTDA chalermpong.int@biotec.ot.th