"nearest neighbor" strategy:

new sequences:

} N seqs

trusted alignment:

M seqs

?
?
?
?

1. Find nearest neighbor

2. pairwise alignment to nearest neighbor

3. Add to new alignment

Profile-based alignment:

trusted alignment ("seed"):

construction procedure

profile

new sequences:

alignment
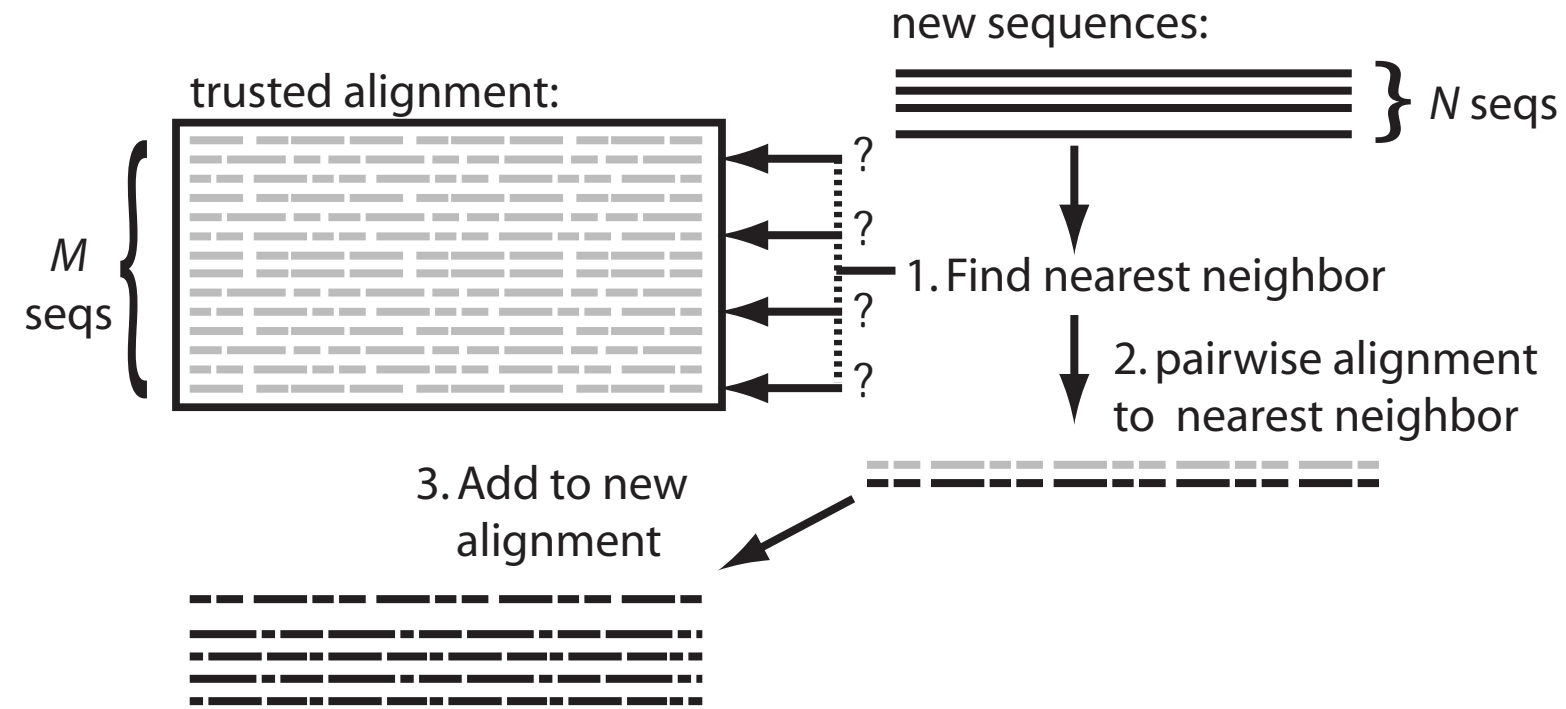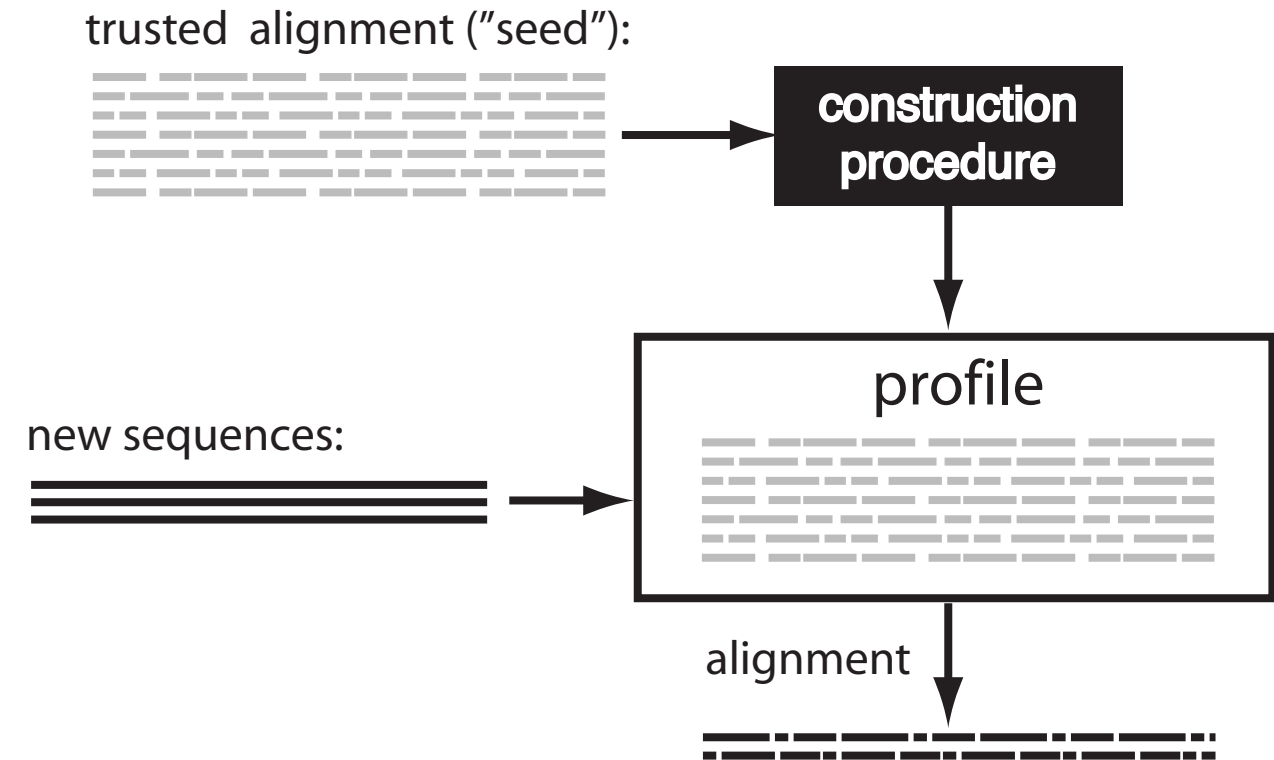
Potential drawbacks:

- *M* may need to be large (~5,000 SSU rRNA seqs for Greengenes(*))

- errors in reference alignment will be propagated

- pairwise alignment ignores varying conservation across alignment

Advantages versus nearest-neighbor:

- *M* can be smaller (~100), so reference is easier to create and trust

- profile-based alignment uses *position-specific* scores

Potential drawbacks:

- only *consensus* positions are aligned, other residues are *inserted*