생물정보학 Report

NCBI에서 목련과(Magnoliaceae) 3종, 녹나무과(Lauraceae) 3종, 미나리아재비과(Ranunculaceae) 3종, 소나무과(Pinaceae) 1종에서 rbcL유전자를 검색하여 총 10개의 유전자에 대한 data matrix를 만들.

tip: 한개의 유전자를 NCBI에서 찾아서 Mega로 옮긴 후 Mega의 search 기능을 사용하여 다른 과의 유전자들을 포함시킬 수도 있음.

1. 이들 10종에 대하여 DNA data를 FASTA form을 만들어
 2. Mega에서 **coding frame을 맞추어** amino-acid 로 변환 후 탑재된 clustalX 로 정렬 (default option 사용) 후 align된 amino-acid sequence를 interleaved form으로 변환하여 **제출**
 3. DNA와 amino acid matrix 각각에 대하여 “각 종들간의 일대일 distance”, “전체 평균 distance”, “각 그룹 내의 평균 distance”, “그룹 간의 distance”, “그룹 간의 평균 distance”를 구하여 **제출**.
 - 이때 그룹은 matrix내의 각 과(family)를 의미함.
 - outgroup 하나도 하나의 group으로 사용하여 그룹간의 distance에 포함시킬것
 - Kimura's two-parameter method를 사용
 4. DNA와 amino acid matrix 각각에 대하여 neighbor-joining tree를 그리고 (DNA model은 Kimura's two parameter method, amino-acid는 default option), 100번의 bootstrapping을 하여 각 node에 값을 표시하여 **제출**
 5. DNA와 amino acid matrix 각각에 대하여 maximum parsimony tree를 그리고 (모두 default option), 100번의 bootstrapping을 하여 각 node에 값을 표시하여 **제출**.
 - 동일한 최소길이의 tree가 여러 개 생겼을 경우에는 strict consensus tree를 그리고 이 위에 bootstrap value들을 기입.
- 모든 tree는 포함된 나자식물(소나무과) outgroup을 이용하여 rooting 할 것.

Final Report는 기말고사 성적의 20%에 해당함.