**Building a BLAST database with your (local) sequences**

Standalone BLAST Setup for Windows PC

<https://www.ncbi.nlm.nih.gov/books/NBK52637/>

<http://barcwiki.wi.mit.edu/wiki/blastTips>

<http://medialab.egloos.com/1273794>

<https://rdrr.io/github/mhahsler/rBLAST/man/makeblastdb.html>

<https://m.blog.naver.com/PostView.naver?isHttpsRedirect=true&blogId=enthusiasm&logNo=20123301075>

<https://varamos.tistory.com/7>

E:\NCBI\blast-BLAST\_VERSION+\db

위의 경로에서 cmd.을 실행함.

열리는 커맨드 창에서 아래의 명령문을 실행함.

만들고자 하는 폴더에서 명령문을 실행해야 함.

$ makeblastdb -in dna-sequences.fasta -parse\_seqids -blastdb\_version 5 -taxid\_map taxonomy.tsv -title "silva-138-99-tax-515-806\_2" -dbtype nucl -out silva-138-99-tax-515-806\_2

* 실바 데이터베이스 313,735개에서 312,201개의 서열을 가진 데이터베이스로 줄어듦. 하지만, 어차피 geneious에서 ‘nt’ 데이터베이스로 다시금 blast 할거면 굳이 원본 데이터베이스에서 서열을 줄이는 작업을 진행하지 않아도 될 것 같음.

$ makeblastdb -in dna-sequences.fasta -parse\_seqids -blastdb\_version 5 -taxid\_map taxonomy.tsv -title "silva-138-99-tax-515-806\_1" -dbtype nucl -out silva-138-99-tax-515-806\_1

* 여전히 Description 부분에 아무런 내용이 기입되지 않음. 따라서 ‘2. Batch BLAST OTU consensus sequences’ 과정이 필요하기 때문에 O/N으로 제작한 sequence 파일을 사용한다고 해도 어차피 blast를 다시 해야함. 그래서 <SequenceId>를 제대로 인식시켜 주기 위해 실바 데이터베이스 서열 아이디 “.”뒤에 붙은 서열 위치 정보를 엑셀로 전부 제거함.

$ makeblastdb -in dna-sequences.fasta -parse\_seqids -blastdb\_version 5 -taxid\_map taxonomy.tsv -title "silva-138-99-tax-515-806" -dbtype nucl -out silva-138-99-tax-515-806

* Taxon이 sequence에 연결되지 않음. 그래서 " -taxid\_map”에서 지정하는 “Format:<SequenceId> <TaxonomyId><newline>” 파일 형태로 맞추기 위해 제일 첫 라인을 삭제함!! 만약 이 시도가 실패하면, O/N으로 제작한 sequence 파일에 taxon이 연결된 fasta 파일로 작업에 들어갈 예정임.

$ blastdbcmd -db silva-138-99-tax-515-806 -info

BLAST® Command Line Applications User Manual [Internet].

<https://www.ncbi.nlm.nih.gov/books/NBK569841/>