Bio-informatica 14-5-2018

Expect 1\*10-10 , 1\*10-4  , 1 \*10-5 , 1\*10-20 , 1\*10-10 🡪 in relatie met de ASCI score.

1\*10-20 🡪 minder resultaten die betrouwbaarder zijn.

1\*10-4  🡪 meer resultaten die minder betrouwbaar zijn.

Word size 🡪 5 is nauwkeuriger dan 6, maar duurt ook langer. Blijft op 6 staan.

Low complexity filter 🡪 f = uit, t= normaal aan, l = aan+ mask

Repeads niet meenemen in het blasten, bacterien hebben geen repeads, win je veel tijd mee.

Organise 🡪 exclude eukaryoten of zoogdieren (mammals)?

Een kleinere database 🡪 betere E-value.

MATRIX 🡪 blosum, beter in het zoeken van uit elkaar gelegen sequenties, lokal alignment

Pam, glabal alignment

GAP cost 🡪 -10 & -1 bacterien muteren snel dus ze zullen meer van elkaar verschillen dus iets soepeler tellen(hogere E-value, meer resultaten in de database).

Database 🡪 NR database