

Technische documentatie

```
if (top !== self) {  
  function calcWidth() {  
    var wW = 0;  
    if (typeof window.innerWidth == 'number') {  
      wW = window.innerWidth;  
    } else if (document.documentElement && documentElement.clientWidth) {  
      wW = document.documentElement.clientWidth;  
    } else if (document.body && document.body.clientWidth) {  
      wW = document.body.clientWidth;  
    }  
    if (sH = document.documentElement.scrollHeight || document.documentElement.clientHeight || document.body.scrollHeight) {  
      var wH = window.innerHeight || document.documentElement.clientHeight || document.body.clientHeight;  
      if (wH < sH) {  
        wW = !document.all && (sH > wH) ? sH : wW;  
      }  
    }  
  }  
  calcWidth();  
}
```

Door Dyogo Borst, Richard Jansen, LysanneRosaria en Roel van de Wiel

Projectgroep 1

Klas Bi1A

Docent E. Kok

Inhoud

- Inleiding
- Werking van de applicatie
- Software architectuur
- Textscripts en resultaten
- Referenties

Inleiding

Het doel van onze web-applicatie is dat je aan de hand van een van de sequenties het desbetreffende BLAST-resultaat kan op halen uit onze database. Dit is zeer handig om snel van een sequentie de desbetreffende resultaten op te zoeken.

Het verslag is geschreven voor mensen die niet zo veel van informatica snappen en dus m.b.v. deze handleiding de database zouden kunnen gebruiken. Ook zit er een stuk bij in het verslag wat advies geeft hoe de server onderhouden moet worden en wat de systeemeisen zijn.

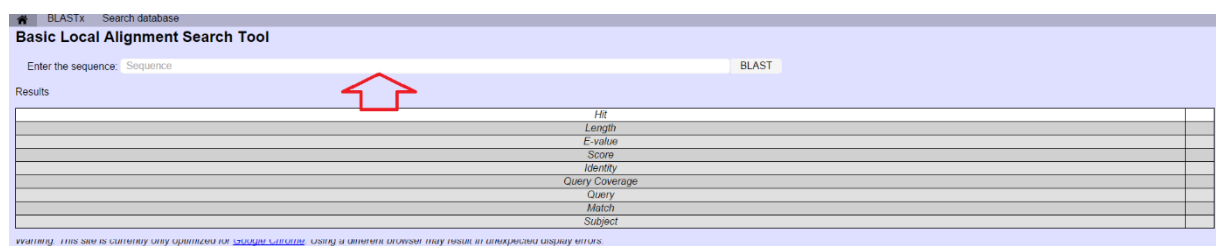
Werking van de applicatie

Voer onderstaande URL in:

<http://ithurtswhenip.nl:666/pg1/blast.psp>



Als u op het kopje BLASTx drukt komt u op onderstaande pagina waarin je een BLAST uit kunt voeren zoals u hieronder kunt zien.



Voer in de balk de sequentie in, hierna zult u een poosje moeten wachten en krijgt u het volgende resultaat

BLASTx Search database

Basic Local Alignment Search Tool

Enter the sequence:

Results

Hit	gi48934712 ref WP_003254832.1 C4-dicarboxylate ABC transporter [Pseudomonas putida] >gi161378136 ref NP_743328.2 TRAP dicarboxylate transporter DctQ subunit [Pseudomonas putida KT2440] >gi148510503 gb ABQ77363.1 Tripartite ATP-independent periplasmic transporter DctQ component [Pseudomonas putida F1] >gi313407529 gb ADR58895.1 TRAP dicarboxylate transporter DctQ subunit putative [Pseudomonas putida BIRD-1] >gi388591631 gb AFK70772.1 tripartite ATP-independent periplasmic transporter DctQ [Pseudomonas putida ND6] >gi397332466 gb AFO48825.1 TRAP dicarboxylate transporter DctQ subunit putative [Pseudomonas putida DOT-T1E] >gi471347394 gb EMR45361.1 TRAP dicarboxylate transporter DctQ subunit [Pseudomonas putida LS46] >gi480562469 gb ENY75643.1 TRAP dicarboxylate transporter DctQ subunit [Pseudomonas putida TRO1] >gi511525681 gb AGN83222.1 C4-dicarboxylate ABC transporter [Pseudomonas putida H2234] >gi512382117 dbj BAN52967.1 putative TRAP dicarboxylate transporter subunit DctQ [Pseudomonas putida NBRC 14164] >gi551052606 gb ERT18412.1 C4-dicarboxylate ABC transporter [Pseudomonas putida SJ9] >gi734575185 gb AJA12399.1 C4-dicarboxylate ABC transporter [Pseudomonas putida S12]
Length	175
E-value	3.93123e-12
Score	167.0
Identity	29
Query	45.3125
Coverage	
Query	CMLVAGLCVLVITIIIPWCYTRYVLCYGSWPEPISILUIVFTFSAGAVCYLSLLHSLPTI
Match	C+*AGL++*T+IIPW++RYVLGSWPEP+S+L++VFTFAY+H+++I
Subject	CIWAGLSILAMTLIIPWIGIFARYVLGTGSSWPEPVSILLMVVFTFVGAAASTRAGAHMAVGMI

Als u een sequentie heeft en u wilt zien of hij in onze database staat en wat voor waardes deze heeft ga dan naar de “Search Database” optie. Je komt op onderstaande pagina terecht en voer in de zoekbalk de sequentie in.

BLASTx Search database

Search local database

Enter the sequence:

Results

Id	Sequence

Hit	E-value	Score	Query Coverage	Identity

Warning: this site is currently only optimized for Google Chrome. Using a different browser may result in unexpected display errors.

Bij een Juist resultaat krijgt u dit:

BLASTx Search database

Search local database

Enter the sequence:

Results

Id	@HWI-M02942.21.000000000-ACNW4.1.1101.14823.1531.1
Sequence	CTCTGTCGATGAAAGAGGTGACAGGACCCGGACCGGAACCTGGCGTCTGCGAGGACCGCCCTGGTGGTGGGATGTTCTTCGGATCCCGCGGATGCGGAACCTCCCGAGTTCCTCGCGCGCCGGTTGACGGCGTTCGGAATCGCCGCCGCGGCAC

Hit	gi576408776 gb EUA02578.1 carbamoyl-phosphate synthase L chain, ATP binding domain protein [Mycobacterium kansasii 732]
E-value	3.61775e-27
Score	287.0
Query Coverage	63
Identity	59

En bij geen hit krijgt u dit:

Enter the sequence:

Results

```
MOD_PYTHON ERROR
ProcessId: 3318
Interpreter: 'dyogo.hobbenschotjansen.nl'
ServerName: 'dyogo.hobbenschotjansen.nl'
DocumentRoot: '/var/sentora/hostdata/zadmin/public_html/dyogo_hobbenschotjansen_nl'
URI: '/search.psp'
Location: None
Directory: '/var/sentora/hostdata/zadmin/public_html/'
Filename: '/var/sentora/hostdata/zadmin/public_html/dyogo_hobbenschotjansen_nl/search.psp'
PathInfo: ''
Phase: 'PythonHandler'
Handler: 'mod_python.psp'

Traceback (most recent call last):
  File "/usr/lib/python2.7/dist-packages/mod_python/importer.py", line 1537, in HandlerDispatch
    default=default_handler, arg=req, silent=hlist.silent)
  File "/usr/lib/python2.7/dist-packages/mod_python/importer.py", line 1229, in _process_target
    result = _execute_target(config, req, object, arg)
  File "/usr/lib/python2.7/dist-packages/mod_python/importer.py", line 1128, in _execute_target
    result = object(arg)
  File "/usr/lib/python2.7/dist-packages/mod_python/psp.py", line 337, in handler
    p.run()
  File "/usr/lib/python2.7/dist-packages/mod_python/psp.py", line 243, in run
    exec code in global_scope
  File "/var/sentora/hostdata/zadmin/public_html/dyogo_hobbenschotjansen_nl/search.psp", line 94, in <module>
    <td><%= header %></td>
NameError: name 'header' is not defined
```

Deze error zouden we eigenlijk nog met exception handling weg kunnen werken.

Softwarearchitectuur

SearchRow1.py

Input main

Een input in de vorm van een sequentie, accession-code of een woord in de parameter invoer. Invoer heeft een default a die global is zodat het script werkt zonder client die input invoert.

Process main

Aanroepen van de functie regex met het variabel invoer erin. Exceptionhandling.

Output main

Alleen als er in het programma een fout optreedt.

Input regex

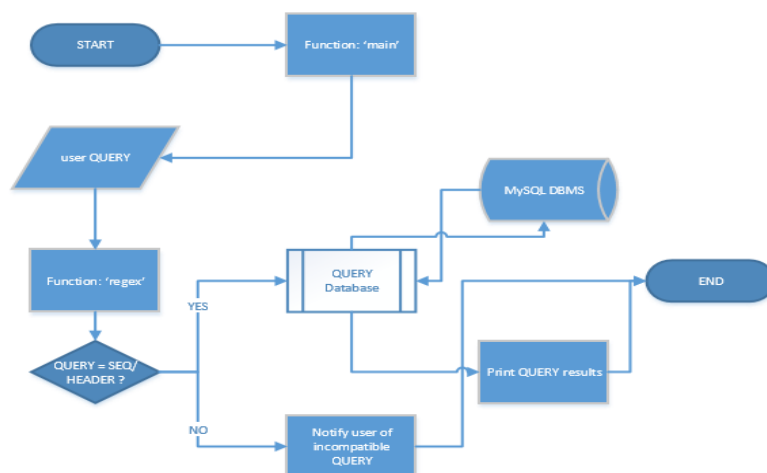
Een input in de vorm van een sequentie, accession-code of een woord in de parameter zoek.

Process regex

Dmv conditionele controle en regular expressions bepalen of zoek een sequentie, accession-code of een woord is. Is dit niet het geval, dan wordt er een melding gegeven. Is zoek een sequentie, wordt het opgezocht in de kolom "nucleotide-sequence" in de database. Is het er iets gevonden, dan worden de eerste resultaten uit deze rij terug gegeven. Is er niks gevonden dan zorgt dit voor een melding. Is zoek een accession-code, wordt dit opgezocht in de kolommen "seq-id" en "result(1-10)-accession-code" in de database. Is het er iets gevonden, dan worden de eerste resultaten uit deze rij terug gegeven. Is er niks gevonden dan zorgt dit voor een melding. Is zoek een woord, wordt dit opgezocht kolommen "result(1-10)-accession-code" in de database. Is het er iets gevonden, dan worden de eerste resultaten uit deze rij terug gegeven. Is er niks gevonden dan zorgt dit voor een melding.

Output regex

Melding wanneer input client niet voldoet aan sequentie, accession-code of woord eisen. Melding wanneer input niet in de database gevonden is.



Singleblast.py

Input main

Eensequentie seq en een action "sb" - Sequence Blast of "csb" - Complement Sequence Blast

Process main

Input action bekijken en bepalen welke functie er vervolgens aangeroepen moet worden, andere soort action wordt opgevangen met een melding en daarnaast staat in deze functie ook de exceptionhandling voor errors die tijdens het runnen van het programma kunnen ontstaan. "sb" roept blastOld aan met seq, "csb" roept blastNew aan met seq.

Output main

Melding bij een andere action dan "sb" of "csb" Meldingen bij errors door de exception handling.

Input blastOld

En sequentie in de parameter old. Main stopt hier het variabel seq in.

Process blastOld

Aanroepen van de functie checkdb met variabel old erin.

Output blastOld

Niets.

Input blastNew

Eensequentie in parameter old. Main stopt hier het variabel seq in.

Process blastNew

Aanroepen van de functie toComplement met variabel old erin. List geretourneerd uit deze functie stoppen in variabel Complement. Vervolgens aanroepen van turnAround met variabel complement erin. String geretourneerd uit deze functiestoppen in variabel new. Aanroepen functie checkdb met argument new erin.

Output blastNew

Weergave van de complementair omgedraaide sequentie.

Input toComplement

Een sequentie in parameter reverse. BlastNew stopt hier het variabel old in.

Process toComplement

lijst seqF maken met de complementaire sequentie waarbij iedere nucleotide een item in de lijst is. A wordt T, T wordt A, G wordt C en C wordt G.

Output toComplement

retourneerd de lijst seqF naar blastNew.

Input turnAround

Eensequentie in parameter list. BlastNew stopt hier het variabel complement in.

Process turnAround

Dmv list wordt per item vanaf de achterste een string f aangevuld. Hierdoor ontstaat in string f de omgekeerde volgorde van de list. Steeds laatste nucleotide uit list in een nieuwe string f zetten om uiteindelijk de omgekeerde list sequentie in f te hebben staan.

Output turnAround

retourneerd de string naar blastNew.

Input checkdb

Een sequentie in parameter zoek. BlastNew stopt hier het variabel new in, blastOld het variabel old.

Process checkdb

De deze functie wordt er ten eerste er bepaald hoeveel rijen er al in de database staan die overeenkomen met een seq id dat de naam seqwebapp bevat, dit komt in variabel count.

Met een regular expression wordt gecontroleerd of het wel echt een DNA sequentie is. Is dit niet zo levert dit een melding op. Is dit wel het geval, wordt de sequentie opgezocht in de database. Staat het in de database, verschijnt er een melding. Staat het niet in de database, dan wordt de functie blast aangeroepen met de variabelen zoek en count erin.

Output checkdb

Sql querys om te controleren. Melding wanneer sequentie geen DNA is. Melding wanneer sequentie al in database staan. Aangeven dat blast begint.

Input blast

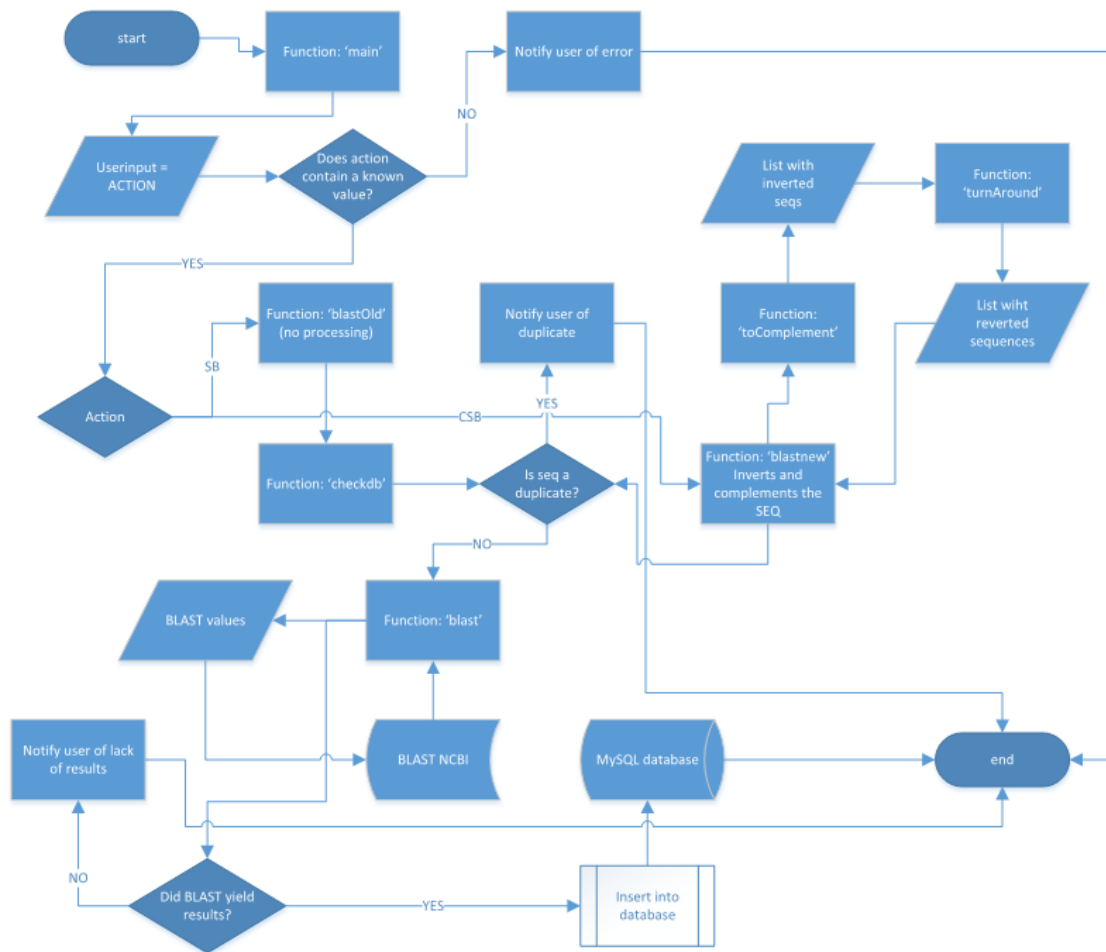
Een sequentie in parameter seq, een aantal in parameter count. Checkdb stopt hier het variabel zoek en count in.

Process blast

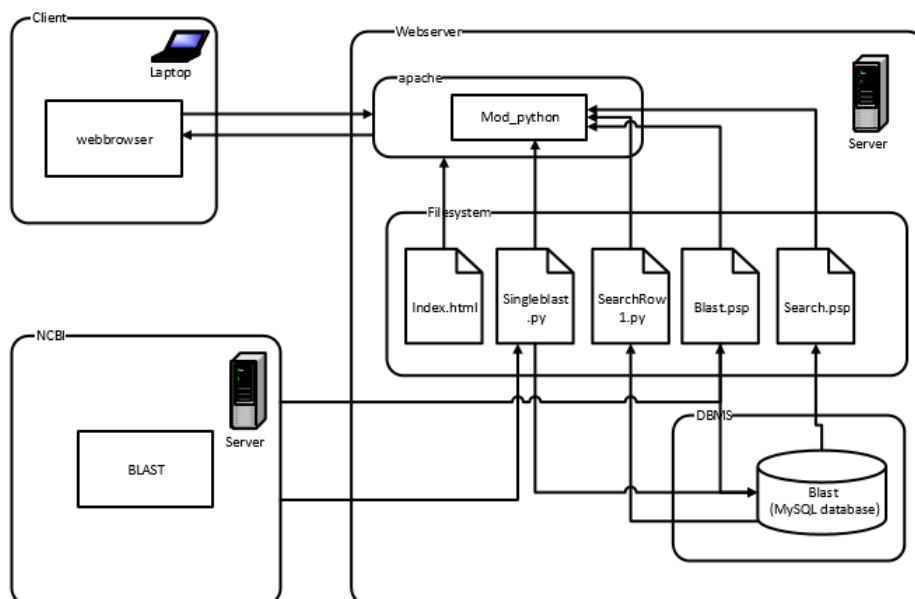
Dmv de modules NCBIWWW en NCBIXML uit Bio.Blast kan een blastx worden uitgevoerd tegen de nr database, e-value threshold van 0.001, score-matrix BLOSUM62. Uit de xml output hiervan worden interessante waarden gehaald en daarmee wordt de database geupdate.

Output blast

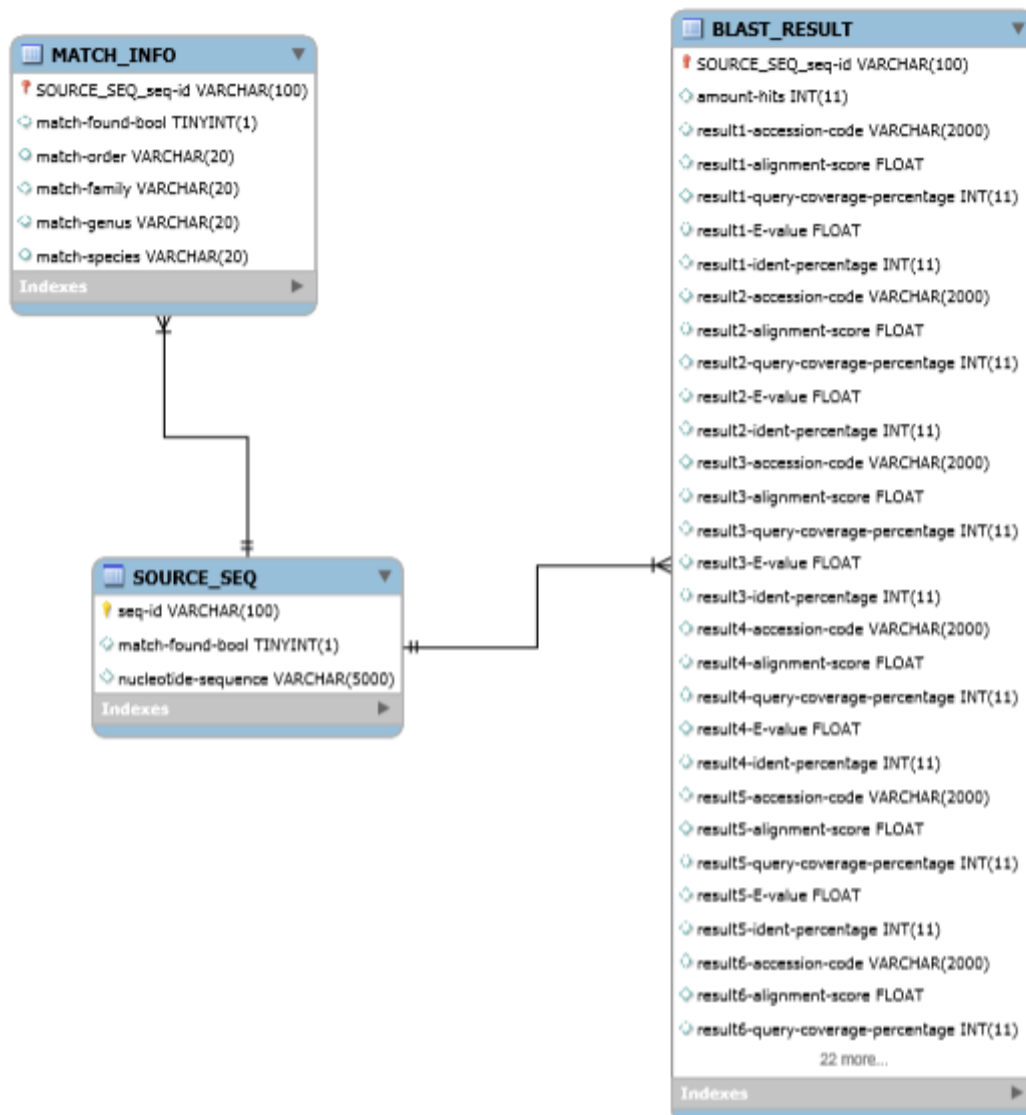
Niets.



Schematische tekening



ERD



SOURCE_SEQ

- Seq-id (header van de reads, om sequentie en resultaten te identificeren)
- Match-found-bool (Boolean om de sequenties met resultaat makkelijk te filteren)
- Nucleotide-sequence (bevat de reads van het sequencen)

BLAST_RESULT

- SOURCE_SEQ_seq-id (foreignkey die verwijst naar de seq-id in de SOURCE_SEQ table)
- amount-hits (integer met het aantal gevonden resultaten per sequence)
- result(1-10) (colommen met daarin de resultaten van het blastscript)

MATCH_INFO

- Heeft nog geen praktisch nut, is een idee voor in de toekomst

Testscripts en resultaten

STAP	Instructie	Verwacht resultaat	Feitelijk resultaat	Opmerking
1	Ga naar dyogo.hobbenschotjansen.nl en ga naar Search database, voer "ATCG" in.	Je krijgt een hit	Klopt	Hij geeft maar 1 sequence aan terwijl het er waarschijnlijk meer kunnen zijn
2	Voer "atcg" in.	Je krijgt een hit, of een error.	Een hit	Bij het zoeken maakt het niet uit of je grote of kleine letters gebruikt
3	Voer "FCTK" in.	Een error of foutmelding	EenMod_Python error	Dit kan nog beter afgevangen worden met exception handling.
4	Ga naar de BLASTx pagina en voer een sequentie van de database in	Je krijgt een uitslag terug	Klopt	Alle resultaten kloppen
5	Voer nu een gedeelte van een sequentie uit de database terug	Je krijgt een uitslag of een error	Je krijgt een uitslag terug	Het heeft het zelfde resultaat als dat ik de sequentie helemaal zou blasten

Referenties

Oracle (MySQL) (z.d.). MySQL 5.5 Reference Manual. Geraadpleegd voor het laatst op 15 Juni 2015, van <http://dev.mysql.com/doc/refman/5.5/en/>

Python (Laatst geupdate, 27 Mei 2015). Python 2.7.10 documentation. Geraadpleegd voor het laatst op 22 Juni 2015, van <https://docs.python.org/2/>

W3schools (z.d.). HTML Element Reference. Geraadpleegd voor het laatst op 21 Juni 2015, van <http://www.w3schools.com/tags/>

BLAST (z.d.). BLAST documentation. Geraadpleegd voor het laatst op 22 Juni 2015, van http://blast.ncbi.nlm.nih.gov/blast/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs