Hello students Unit IV bioinformatics

databases and their sequencing module name,

biological databases,

retrieval system module 21 myself

Ms. Shreeveni Tari outline of the module.

Includes introduction to biological

database retrieval system.

Examples of biological database driver

system learning outcomes describes the

process of biological database retrieval,

system cites.

Biological database retrieval system.

Now we have already seen that the

molecular biology data is stored in

the form of biological databases and

the chief objective of the development

of these databases is to organize

data in a set of structured records

to enable easy retrieval of information.

So we have the information retrieval

that is defined as the activity of

obtaining information system. Resources

that are relevant to an information

need from a collection of those

resources will understand it in a

more better way from a popular example.

That is,

with the use of Google Now.

Here,

whenever we are in need of some information,

what we do is we make use of the

Google search engine OK in the

dialog box you are just typing

a word that is referred to as.

Query and you go for the search option.

Now that will direct the web

pages to the data repository,

so whatever is matching with your

query will be displayed as they

query results and then based on the

displayed results you can very well

choose the entries and you will

be directed to those web pages.

This is the general process that is

utilized during the information retrieval.

Now same is applicable.

For the biological data,

basically retrieval system working

can be considered same as

that of the Google search engine.

OK,

so you have the biological

databases which will be storing the

biological information that will

be in the computer readable form.

So what you have to do is you have to

make use of these biological database

retrieval system for the text based searches.

So those you just have to type your

query and the results will be displayed.

And those results you can use

for your own purpose.

So we have the three biological

database retrieval system,

namely ENTREZ, sequence retrieval

system that is abbreviated as SRS,

and we have DBGET.

OK,

so these systems allows text

searching of multiple molecular

biology databases and also provides

links to relevant information for

entries that match the search.

So once you type your query.

In the search box you will

be directed to the results.

Now we are utilizing the text

based searching.

OK,

that is dependent on certain

concepts we have.

The first concept that is called

as the Boolean search.

Now here you just have to type

a word and you can make use of

the words like and and or OK.

So when you make use of this it will be

referred to as a Boolean search algorithm.

Then you can broaden your search.

So actually if you are typing one of the.

Name of the enzyme.

You can also broaden your search

so all the related information to that

particular enzyme will be displayed.

It will not be restricted

to your particular search.

Then you can narrow down specific word.

You have to type so that will be only

directing you to a particular search.

OK, then you can also go for proximity

searching so whatever available

information is present in their database,

everything related to your

query search will be display.

So these are few of the

basic search concepts.

Now we'll see the first one that is

the ENTREZ that is developed and

maintained by National Center for

Biotechnology Information that is NCBI.

Now these allows text based searches

for a variety of data including the

annotated genetic sequence information,

structural information,

as well as citations, abstracts,

full papers and a taxonomic data.

So all these databases which are offering

this information will be linked together.

That is the key feature of the ENTREZ

OK, and this is possible because

of the cross referencing that is

done between the NCBI databases.

OK,

that is based on pre-existing and local

relationships between the individual entries.

Now the features of these search engine

includes you can make use of the limits

that will be restricting your search.

You can make use of the preview or

indexes so you can search from the

options which are provided to you.

You can make use of the history

like for example for days together.

If you are working on the same topic then

you can have this history and you can

very well browse through your history.

Then also you have the clipboard option.

Now only difference between

the history and clipboard.

Is it the clipboard history or

the clipboard will be only meant

for short duration.

OK it will not be stored for long time.

Now.

This ENTREZ retrival system will provide

access to various databases which includes.

Pub Med that is having the biomedical

literature GenBank that will be for

nucleic acid sequence databases

structure that will be dealing

with a 3 dimensional molecular

structures genome that will be meant

for complete genome assemblies.

OMIM that is the online Mendelian

inheritance in man that is

dealing with the human diseases,

genes,

and genetic disorders and the Taxonomy.

That is the organisms whatever

are present in the gene bank.

Then we have the SRS retrieval system

that is the sequence retreival system.

It is an integrated database RETRIVAL

system maintained by the European

Bioinformatics Institute that is EBI.

Now here the types of databases

which are included our sequence databases,

transcription factors, protein,

3 dimensional structures,

genome mapping mutations and

locus specific mutation.

These offers direct access to certain

sequence analysis applications.

Such as sequence similarity,

searching, example blast and

clustal w sequence alignments.

Now,

queries can be in the form

of accession number that is,

the unique identification number that

is given for each of the entry code,

or it can be based on the name of the author,

or it can be also based on the name of

the Organism,

or it can be also using the Boolean

relationship between the keywords.

That is, with the use of words.

Such as and & or.

Then we have the third retrival system.

That is the DBGET.

This is an integrated database retrieval

system developed and jointly maintained

by the Institute for Chemical Research,

Kyoto University, and the Human Genome

Center of the University of Tokyo.

Now this has the integration

with more than 20 databases.

Now, in this case the search,

commands can be bfind and bget when

you're utilizing bfind commands,

you are basically going for

the text based searching OK,

we just have to type the word, when

you're going for bget command you

have to type in the accession

number that is the unique number

given for each of the entry code.

Now, apart from the sequences,

structural or other related information,

you will be also getting the search

results which will be in the form of

the available and cited information.

Link DB is an associated database that

is present for the DBGET which

provides links between the entries

in the different databases available

and further you will be also directed

to the Organism specific databases.

A unique feature of this division.

Is its association with the Kyoto

and encyclopedia of genes and genomes.

That is the KEGG database that

is dealing with the database of

metabolic and regulatory pathways.

So these are few of the biological

database retrieval systems which are

popularly used and these are mainly

used for the text based searching.

So these are few of the references in

the form of books and web sources.

Thank you.