

# Assignment - 2 Bio-Informatics

Name : Neelakanta Sriram

Roll.No : S20170010102

## Task - 1:

### Exact Number of Entries

The screenshot shows the NCBI Gene database search results for the query "dihydrofolate reductase". The search results are displayed in a table with columns: Name/Gene ID, Description, Location, Aliases, and MIM. The table lists various genes and their associated information across different species.

Name/Gene ID	Description	Location	Aliases	MIM
<input type="checkbox"/> DHFR ID: 1719	dihydrofolate reductase [Homo sapiens (human)]	Chromosome 5, NC_000005.10 (80626226..80654983, complement)	DHFRP1, DYR	126060
<input type="checkbox"/> tsdA ID: 944790	dihydrofolate reductase [Escherichia coli str. K-12 substr. MG1655]	NC_000913.3 (49823..50302)	B0048, ECK0049, tmr, tmrA	
<input type="checkbox"/> Dhfr ID: 13361	dihydrofolate reductase [Mus musculus (house mouse)]	Chromosome 13, NC_000079.6 (8254703..9238095)	B430436I03Rik, AA607882, AW655994	
<input type="checkbox"/> Dhfr ID: 24312	dihydrofolate reductase [Rattus norvegicus (Norway rat)]	Chromosome 2, NC_005101.4 (21931887..21958927)	Dhfr1	
<input type="checkbox"/> Dhfr ID: 42003	dihydrofolate reductase [Drosophila melanogaster (fruit fly)]	Chromosome 3R, NT_033777.3 (16481167..16481878)	Dmel_CG14887, CG14887, DHFR, DmelCG14887, dhfr	
<input type="checkbox"/> dhfr ID: 61882	dihydrofolate reductase [Danio rerio (zebrafish)]	Chromosome 5, NC_007116.7 (26138313..26145017)	cb595, zgc:89637	
<input type="checkbox"/> DHFR2 ID: 200895	dihydrofolate reductase 2 [Homo sapiens (human)]	Chromosome 3, NC_000003.12 (94057922..94063389, complement)	DHFR1, DHFRP4	616588
<input type="checkbox"/> DHFR1 ID: 854411	dihydrofolate reductase [Saccharomyces cerevisiae S288C]	Chromosome XV, NC_001147.6 (780906..781541)	YOR236W	
<input type="checkbox"/> DHFRP1 ID: 573971	dihydrofolate reductase pseudogene 1 [Homo sapiens (human)]	Chromosome 18, NC_000018.10 (26167862..26171357, complement)		
<input type="checkbox"/> dhfr ID: 1390554	dihydrofolate reductase [Escherichia coli]	NC_019062.1 (4613..5110)	D616_p72001	
<input type="checkbox"/> DHFR ID: 427317	dihydrofolate reductase [Gallus gallus (chicken)]	Chromosome Z, NC_006127.5 (64257309..64273299)		
<input type="checkbox"/> AT2G21550 ID: 616694	Bifunctional dihydrofolate reductase/thymidylate synthase [Arabidopsis thaliana (thale cress)]	Chromosome 2, NC_003071.7 (8226917..8230205)	AT2G21550, F2G1.18, F2G1.18	

The interface includes a search bar at the top with the query "dihydrofolate reductase" and a "Search" button. Below the search bar, there are filters for "Gene" and "Database". The search results are displayed in a table with columns: Name/Gene ID, Description, Location, Aliases, and MIM. The table lists various genes and their associated information across different species. The results are paginated, showing "Page 1 of 266".

## Task - 2:

### A. Dihydrofolate reductase AND Saccharomyces.

NCBI Resources How To Sign in to NCBI

Gene  Search

Create RSS Save search Advanced

Gene sources: Genomic, Categories: Annotated genes, Protein-coding, Sequence content, RefSeq, Status: Current, Discontinued, Clear all, Show additional filters

Tabular 20 per page Sort by Relevance Send to: Hide sidebar >>

Search results: Items: 1 to 20 of 37 Selected: 1

Name/Gene ID	Description	Location	Aliases
<input checked="" type="checkbox"/> DER1 ID: 854411	dihydrofolate reductase [Saccharomyces cerevisiae S288C]	Chromosome XV, NC_001147.6 (780906..781541)	YOR236W
<input type="checkbox"/> REF1 ID: 850873	Rfx1p [Saccharomyces cerevisiae S288C]	Chromosome XII, NC_001144.5 (507797..510232, complement)	YLR176C, CRT1
<input type="checkbox"/> NEWENTRY ID: 3974689	Record to support submission of GeneRIFs for a gene not in Gene (Plasmodium falciparum; malaria parasite P. falciparum) [Plasmodium falciparum (malaria parasite P. falciparum)]		
<input type="checkbox"/> KLLA0_D07459g ID: 2852927	uncharacterized protein [Kluyveromyces fragilis]	Chromosome D, NC_006040.1 (639883..640581)	KLLA0_D07459g
<input type="checkbox"/> KLTH0010802g ID: 8295418	KLTH0010802p [Lachancea thermotolerans CBS 6345]	Chromosome D, NC_013080.1 (885955..886668)	KLTH0010802g
<input type="checkbox"/> YALJ0_C12771g ID: 2909513	YALJ0C12771p [Yarrowia lipolytica CLB122]	Chromosome C, NC_006069.1 (1773531..1774100, complement)	YALJ0_C12771g, YALJ0C12771g
<input type="checkbox"/> DEHA2C16459g ID: 2900017	DEHA2C16459p [Debaryomyces hansenii CBS757]	Chromosome C, NC_000045.2 (148332..1483952, complement)	DEHA2C16459g
<input type="checkbox"/> LALA0_001e03378g ID: 3463488	uncharacterized protein [Lachancea lancestensis]		LALA0_001e03378g
<input type="checkbox"/> ZYRO0F06930g ID: 3463488	hypothetical protein [Zygosaccharomyces rouxi]	Chromosome F, NC_012995.1	ZYRO0F06930g

Filters: Manage Filters

Results by taxon: Top Organisms [Tree]  
Saccharomyces cerevisiae S288C (2)  
[Candida] glabrata (2)  
Kluyveromyces fragilis (2)  
Enamethium gossypii ATCC 10895 (2)  
Zygosaccharomyces rouxi (2)  
All other taxa (27)  
More...

Find related data: Database: Select Find items

Search details: dihydrofolate reductase[All Fields] AND ("Saccharomyces" [Organism] OR saccharomyces[All Fields]) Search See more...

### B . Dihydrofolate reductase AND Streptococcus pneumoniae

NCBI Resources How To Sign in to NCBI

Gene  Search

Create RSS Save search Advanced

Gene sources: Genomic, Categories: Annotated genes, Protein-coding, Sequence content, RefSeq, Status: Current, Clear all, Show additional filters

Tabular Sort by Relevance Send to: Hide sidebar >>

See CHR1 (REDUCTASE) chalcone reductase CHR1 in the Gene database reductase in Glycine max [Bacotera dorsalis] All 2 Gene records

Search results: Items: 4 Selected: 1 See also 26 discontinued or replaced items.

Name/Gene ID	Description	Location	Aliases
<input type="checkbox"/> NEWENTRY ID: 3979060	Record to support submission of GeneRIFs for a gene not in Gene (Diplococcus pneumoniae; Micrococcus pneumoniae Use when strain, subtype, isolate, etc. is unspecified, or when different from all specified ones in Gene.) [Streptococcus pneumoniae]		
<input checked="" type="checkbox"/> dfr ID: 933911	dihydrofolate reductase [Streptococcus pneumoniae R6]	NC_003098.1 (1412861..1413367, complement)	spr1429
<input type="checkbox"/> dfr ID: 901041	putative dihydrofolate reductase [Streptococcus pyogenes M1 GAS]	NC_002737.2 (731153..731650)	SPY_0883, SPY0883
<input type="checkbox"/> dfrA ID: 1028294	dihydrofolate reductase [Streptococcus mutans UA159]	NC_004350.2 (897334..897846)	SMU_947, SMU947

Filters: Manage Filters

Results by taxon: Top Organisms [Tree]  
Streptococcus pneumoniae (2)  
Streptococcus pyogenes M1 GAS (1)  
Streptococcus mutans UA159 (1)

Find related data: Database: Select Find items

Search details: ("Streptococcus pneumoniae"[Organism] OR Streptococcus pneumoniae[All Fields]) AND dihydrofolate reductase[All Fields] AND alive[prop] Search See more...

## Task - 3

### Information Gathered From Bars :

----> The genomic context of **Streptococcus pneumonia R6**:

**Gene symbol** - DFR

**Scientific name** - Streptococcus pneumonia R6 (strain: R6)

**Gene ID** - 933011

**Gene Type** - Protein coding

**Lineages** - Bacteria, Bacilli, etc

**Genomic Sequence Code** - NC\_003098.1

Red color shows the directionality which is from left to right for DFR.

----> The genomic context of **Saccharomyces cerevisiae S2886** are:

**Gene symbol** - DFR

**Scientific name** - Streptococcus pneumonia R6 (strain: R6)

**Gene ID** - 933011

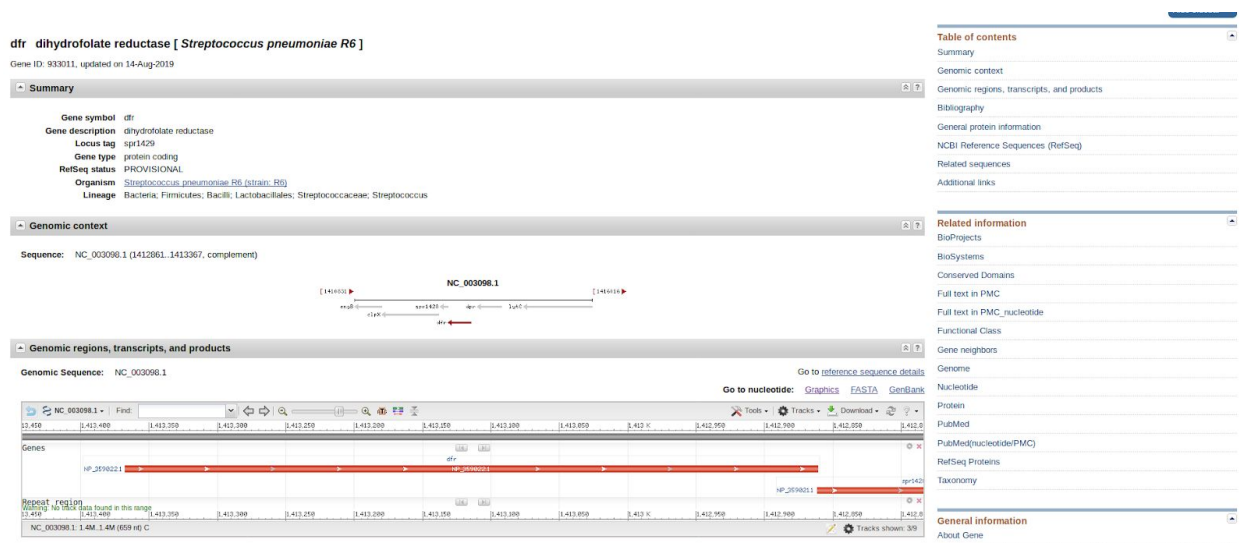
**Gene Type** - Protein coding

**Lineages** - Bacteria, Bacilli, etc

**Genomic Sequence Code** - NC\_003098.1

Red color shows the directionality which is from left to right for DFR

### Streptococcus :



### Saccharomyces :

**DFR1 dihydrofolate reductase [ *Saccharomyces cerevisiae* S288C ]**  
 Gene ID: 854411, updated on 1-Nov-2019

**Summary**

Gene symbol: DFR1  
 Gene description: dihydrofolate reductase  
 Primary source: SGD:SGD0005762  
 Locus tag: YOR239W  
 Gene type: protein coding  
 RNA name: dihydrofolate reductase  
 RefSeq status: REVIEWED  
 Organism: *Saccharomyces cerevisiae* S288C (strain S288C)  
 Lineage: Eukaryota; Fungi; Dikarya; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces

**Genomic context**

Location: chromosome XV  
 Exon count: 1  
 Sequence: Chromosome XV, NC\_001147.6 (780906..781541)

**Genomic regions, transcripts, and products**

Genomic Sequence: NC\_001147.6

Go to reference sequence details: [Graphics](#) [FASTA](#) [GenBank](#)

NC\_001147.6 Find

Genes, RefSeq propagation from SGD, annotation version R64-2.1

DFR1

**Bibliography**

1. nucleotide sequence of the dihydrofolate reductase gene of *Saccharomyces cerevisiae*.

**Related information**

BioAssay by Target (List)  
 BioAssay by Target (Summary)  
 BioProjects  
 BioSystems  
 Conserved Domains  
 Full text in PMC  
 Full text in PMC\_nucleotide  
 Functional Class  
 Gene neighbors  
 Genome  
 GEO Profiles  
 HomoloGene  
 Nucleotide  
 Probe  
 Protein  
 PubMed

## Task - 4

Saccharomyces (Yeast) has **211** amino acids.  
 Streptococcus (Bacteria) has **168** amino acids.

## Task - 5

Features of Saccharomyces (Yeast) are :

FEATURES	Location/Qualifiers
source	1..636 /organism="Saccharomyces cerevisiae S288C" /mol_type="genomic DNA" /strain="S288C" /db_xref="taxon:559292" /chromosome="XV"
gene	<1..>636 /gene="DFR1" /locus_tag="YOR236W" /db_xref="GeneID:854411"
mRNA	<1..>636 /gene="DFR1" /locus_tag="YOR236W" /product="dihydrofolate reductase" /transcript_id="NM_001183655.1" /db_xref="GeneID:854411"
CDS	1..636 /gene="DFR1" /locus_tag="YOR236W" /EC_number="1.5.1.3" /experiment="EXISTENCE:curator inference:G0:0005829 cytosol [PMID:1427091 PMID:2827121 PMID:2838385]" /experiment="EXISTENCE:direct assay:G0:0003729 mRNA binding [PMID:21124907]" /experiment="EXISTENCE:direct assay:G0:0004146 dihydrofolate reductase activity [PMID:2838385]" /experiment="EXISTENCE:direct assay:G0:0005737 cytoplasm [PMID:14562095]" /experiment="EXISTENCE:direct assay:G0:0005739 mitochondrion [PMID:14562095]" /experiment="EXISTENCE:mutant phenotype:G0:0004146 dihydrofolate reductase activity [PMID:1427091 PMID:2827121]" /experiment="EXISTENCE:mutant phenotype:G0:0046452 dihydrofolate metabolic process [PMID:1427091]" /experiment="EXISTENCE:mutant phenotype:G0:0046654 tetrahydrofolate biosynthetic process [PMID:1427091]" /note="Dihydrofolate reductase involved in tetrahydrofolate biosynthesis; required for respiratory metabolism; mutation is functionally complemented by human DHFR" /codon_start=1 /product="dihydrofolate reductase" /protein_id="NP_014879.1" /db_xref="GeneID:854411" /db_xref="SGD:5000005762" /translation="MAGGKIPIVGIVACLPENMGIGFRGGLPWRLPSEMKYFRQVTS LTKDPNKKNALIMGRKTWESIPPKFRPLPNRMNVIISRSFKDDFVHDKERSIVQSN SLANAIMNLESNFKEHLRIYVIGGGEVYSQIFSITDHWLITKINPLDKNATPAMDT FLDAKKLEEVFSEQDPAQLKEFLPPKVELPETDCDQRYSLKEKGYCFEFTLYNRK"
ORIGIN	1 atgctggag gaaagattcc tattgttaga attgtggcat gtttacagcc ggagatgggg

	Gene
Q	Streptococcus pneumoniae AND dihydrofolate reductase AND (alive[p... (Gene
Q	Streptococcus dfri AND (alive[prop]) (56386) Gene
	See more...

Features of Streptococcus (Bacteria) are:

**Streptococcus pneum** x **dhfr dihydrofolate red** x **Streptococcus pneum** x **DHFR1 dihydrofolate r** x +

https://www.ncbi.nlm.nih.gov/genbank/NC\_003098.1/report-genbank&from=1412861&to=1413367

**TITLE** Direct Submission  
**JOURNAL** Submitted (27-JUL-2001) Infectious Diseases Research, Eli Lilly and Company, Lilly Research Labs, Indianapolis, IN 46285-0438, USA  
**COMMENT** REVIEWED REFSEQ: This record has been curated by NCBI staff. The reference sequence was derived from [85987311](#).  
 RefSeq Category: Reference Genome  
 UPK: UniProt Genome  
 COMPLETENESS: full length.  
**FEATURES**  
 source 1..507  
   /organism="Streptococcus pneumoniae R6"  
   /mol\_type="genomic DNA"  
   /strain="R6"  
   /db\_xref="taxon:171101"  
 gene complement(1)  
   /locus\_tag="spr1428"  
   /db\_xref="GeneID:933894"  
 CDS complement(1)  
   /locus\_tag="spr1428"  
   /codon\_start=1  
   /transl\_table=11  
   /product="hypothetical protein"  
   /protein\_id="NP\_359021.1"  
   /db\_xref="GeneID:933894"  
   /translation="MERSLGLFTAFLCFCFLAGADFRKKRYGLSILLWNAFTNLVNSIHAFYMTLF"  
 gene complement(1..507)  
   /gene="dhfr"  
   /locus\_tag="spr1429"  
   /db\_xref="GeneID:933891"  
 CDS complement(1..507)  
   /gene="dhfr"  
   /locus\_tag="spr1429"  
   /EC\_number="1.3.1.3"  
   /codon\_start=1  
   /transl\_table=11  
   /product="dihydrofolate reductase"  
   /protein\_id="NP\_359022.1"  
   /db\_xref="GeneID:933891"  
   /translation="MKKIVAINADDEGLIGKRNLPWHLPAELQHKETTLNHAZLMGRVFDGNGRLLPKRETLILTRNPEEKIDGVATFQDVSVLDWYDQKKNLYIIGKQIFQAFEPYLDVIVTHIHARVEGDTYPEELDLSLFETVSSKFYAKDEKNPYDFTIQYRKNKEV"  
**ORIGIN**  
 1 tttagacttccc ttctctctgc gatattggat ggtataatca taaggattct tcctatcttt  
 61 ggcgtaaact ttgcttgaaa cgtctcaaa aagagcaag tcaagctctt cagggaata  
 121 gatatctctt tccaccgag catgaatg agtgcacatc acttcatcaa ggaagcttc  
 181 aaagccttga aaattttgt tccaccgat aatgtagaga ttctttcttc gatctgata  
 241 ccagtcacga acagactgga cgtctcgaaa agtagaaccc ccactctatc ttcttccgg  
 301 attacgcttc aaatacggg ttctcgttt tgaagcaag cagagcccca tccatcaaa  
 361 ggtcacacgc cccatcaaga tagcatgatt cagattgttt tctttaaagt gctgcaattc  
 421 tgcctgcaaa tgcacagga gacgattttc cttaacatc aaacctcttc catctggggc  
 481 ccaaatagct acgatcttct tagtcat  
 //

**reagents**  
 [ExactAntigen/Labome]  
 reagents [ExactAntigen/Labome]  
 reagents [ExactAntigen/Labome]  
 Order GATC cDNA clone/Protein/Antibody/RNAi [OnGene]  
 Order gta cDNA clone/Protein/Antibody/RNAi [OnGene]  
 Order ADK cDNA clone/Protein/Antibody/RNAi [OnGene]

**Recent activity**  
 Streptococcus pneumoniae R6 chromosome, complete genome Nucleotide  
 dhfr (Streptococcus pneumoniae R6) Gene  
 Streptococcus pneumoniae AND dihydrofolate reductase AND (alive[...]) (Gene)  
 Streptococcus dhfr AND (alive[prep]) (56386) Gene  
 Streptococcus AND (alive[prep]) (56386) Gene  
 See more...

## Task - 6

### Saccharomyces : S288C

This protein is made up of 211 amino acids. Some references related to proteins in the database include 'The nucleotide sequence for of *Saccharomyces cerevisiae* chromosome XV' from the journal Nature and 'Life with 6000 genes' from the Science Journal.

Although, the protein as a whole is a dhfr, it is actually amino acids from 8 to 208 which is actually called as the 'DHFR' region. S288C protein starts with 'm' amino acid and ends with 'k' amino acid. DHFR region starts with 'i' amino acid and ends with 'y' amino acid



## Saccharomyces cerevisiae S288C chromosome XV, complete sequence

NCBI Reference Sequence: NC\_001147.6

[FASTA](#) [Graphics](#)

Go to: ☐

LOCUS NC\_001147 636 bp DNA linear CON 01-NOV-2019  
DEFINITION Saccharomyces cerevisiae S288C chromosome XV, complete sequence.  
ACCESSION [NC\\_001147](#) REGION: 780906..781541  
VERSION NC\_001147.6  
DBLINK BioProject: [PRJNA128](#)  
Assembly: [GCF\\_000146045.2](#)  
KEYWORDS RefSeq.  
SOURCE Saccharomyces cerevisiae S288C  
ORGANISM [Saccharomyces cerevisiae S288C](#)  
Eukaryota; Fungi; Dikarya; Ascomycota; Saccharomycotina;  
Saccharomycetes; Saccharomycetales; Saccharomycetaceae;  
Saccharomyces.  
REFERENCE 1 (bases 1 to 636)  
AUTHORS Dujon,B., Albermann,K., Aldea,M., Alexandraki,D., Ansorge,W.,  
Arino,J., Benes,V., Bohn,C., Bolotin-Fukuhara,M., Bordonne,R.,  
Boyer,J., Camasses,A., Casamayor,A., Casas,C., Cheret,G.,  
Cziepluch,C., Daignan-Fornier,B., Dang,D.V., de Haan,M., Delius,H.,  
Durand,P., Fairhead,C., Feldmann,H., Gaillon,L., Kleine,K. et al.  
TITLE The nucleotide sequence of Saccharomyces cerevisiae chromosome XV  
JOURNAL Nature 387 (6632 SUPPL), 98-102 (1997)  
PUBMED [9169874](#)  
REFERENCE 2 (bases 1 to 636)  
AUTHORS Goffeau,A., Barrell,B.G., Bussey,H., Davis,R.W., Dujon,B.,  
Feldmann,H., Galibert,F., Hoheisel,J.D., Jacq,C., Johnston,M.,  
Louis,E.J., Mewes,H.W., Murakami,Y., Philippsen,P., Tettelin,H. and  
Oliver,S.G.  
TITLE Life with 6000 genes  
JOURNAL Science 274 (5287), 546 (1996)  
PUBMED [8849441](#)  
REFERENCE 3 (bases 1 to 636)  
CONSTRM NCBI Genome Project  
TITLE Direct Submission  
JOURNAL Submitted (31-OCT-2019) National Center for Biotechnology  
Information, NIH, Bethesda, MD 20894, USA  
REFERENCE 4 (bases 1 to 636)  
CONSTRM Saccharomyces Genome Database  
TITLE Direct Submission  
JOURNAL Submitted (16-JAN-2015) Department of Genetics, Stanford  
University, Stanford, CA 94305-5120, USA  
REMARK Protein update by submitter  
REFERENCE 5 (bases 1 to 636)  
CONSTRM Saccharomyces Genome Database  
TITLE Direct Submission  
JOURNAL Submitted (04-MAY-2012) Department of Genetics, Stanford

## Streptococcus:

### R6 .

This is made up of 168 amino acids which is lower than S288C.3 references for work related to this protein is provided in the database. References for these are official government research based references according to the R6 protein details. This is under the NCBI Genome Project which aims at collecting genomic information from various organisms and studying them in order to collect data.

 Due to the large size of this record, sequence and annotated features are not shown. Use the "Customize view" panel to change the display.

## Streptococcus pneumoniae R6 chromosome, complete genome

NCBI Reference Sequence: NC\_003098.1

[FASTA](#) [Graphics](#)

[Go to:](#) ☐

LOCUS	NC_003098	507 bp	DNA	linear	CON 11-JAN-2017
DEFINITION	Streptococcus pneumoniae R6 chromosome, complete genome.				
ACCESSION	<a href="#">NC_003098</a> REGION: 1412861..1413367				
VERSION	NC_003098.1				
DBLINK	BioProject: <a href="#">PRJNA57859</a>				
	Assembly: <a href="#">GCF_000007045.1</a>				
KEYWORDS	RefSeq.				
SOURCE	Streptococcus pneumoniae R6				
ORGANISM	<a href="#">Streptococcus pneumoniae R6</a> Bacteria; Firmicutes; Bacilli; Lactobacillales; Streptococcaceae; Streptococcus.				
REFERENCE	1 (bases 1 to 507)				
AUTHORS	Hoskins,J.A., Alborn,W. Jr., Arnold,J., Blaszcak,L., Burgett,S., DeHoff,B.S., Estrem,S., Fritz,L., Fu,D.-J., Fuller,W., Geringer,C., Gilmour,R., Glass,J.S., Khoja,H., Kraft,A., LaGace,R., LeBlanc,D.J., Lee,L.N., Lefkowitz,E.J., Lu,J., Matsushima,P., McAhren,S., McHenney,M., McLeaster,K., Mundy,C., Nicas,T.I., Norris,F.H., O'Gara,M., Peery,R., Robertson,G.T., Rockey,P., Sun,P.-M., Winkler,M.E., Yang,Y., Young-Bellido,M., Zhao,G., Zook,C., Baltz,R.H., Jaskunas,S.Richard., Rosteck,P.R. Jr., Skatrud,P.L. and Glass,J.I.				
TITLE	Genome of the bacterium Streptococcus pneumoniae strain R6				
JOURNAL	J. Bacteriol. 183 (19), 5709-5717 (2001)				
PUBMED	<a href="#">11544234</a>				
REFERENCE	2 (bases 1 to 507)				
CONSRM	NCBI Genome Project				
TITLE	Direct Submission				
JOURNAL	Submitted (03-OCT-2001) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA				
REFERENCE	3 (bases 1 to 507)				
AUTHORS	Hoskins,J.A., Alborn,W. Jr., Arnold,J., Blaszcak,L., Burgett,S., DeHoff,B.S., Estrem,S., Fritz,L., Fu,D.-J., Fuller,W., Geringer,C., Gilmour,R., Glass,J.S., Hann,A., Khoja,H., Kraft,A., LaGace,R., LeBlanc,D.J., Lee,L.N., Lefkowitz,E.J., Lu,J., Matsushima,P., McAhren,S., McHenney,M., McLeaster,K., Mundy,C., Nicas,T.I., Norris,F.H., O'Gara,M., Peery,R., Robertson,G.T., Rockey,P., Sun,P.-M., Winkler,M.E., Yang,Y., Young-Bellido,M., Zhao,G., Zook,C., Baltz,R.H., Jaskunas,S.Richard., Rosteck,P.R. Jr., Skatrud,P.L. and Glass,J.I.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-JUL-2001) Infectious Diseases Research, Eli Lilly and Company, Lilly Research Labs, Indianapolis, IN 46285-0438, USA				
COMMENT	REVIEWED <a href="#">REFSEQ</a> : This record has been curated by NCBI staff. The reference sequence was derived from <a href="#">AF007317</a>				



## **Task 7:**

### **FASTA for Streptococcus Pneumoniae Protein:**

```
>>NP_359022.1 dihydrofolate reductase [Streptococcus pneumoniae R6]
MTKKIVAIWAQDEEGLIGKENRLPWHLPAELQHFKETTLNHAILMGRVTFDGMGRRLLPKRETLI
LTRNPEEKIDGVATFQDVQSVLDWYQDQEKNLIIIGGKQIFQAFEPYLDEVIVTHIHARVEGDTY
FPEELDLSLFETVSSKFYAKDEKNPYDFTIQYRKREKV
```

### **FASTA for Saccharomyces Protein:**

```
>NP_014879.1 dihydrofolate reductase [Saccharomyces cerevisiae S288C]
MAGGKIPIVGIVACLQPEMGIGFRGGLPWRLPSEMKYFRQVTS�TKDPNKKNALIMGRKTWESI
PPKFRPLPNRMNVIISRSFKDDFVHDKERSIVQSNLANAIMNLESNFKEHLERIYVIGGGEVYS
QIFSITDHWLITKINPLDKNATPAMDTFLDAKKLEEVFSEQDPAQLKEFLPPKVELPETDCDQRY
SLEEKGYCFEFTLYNRK
```

### **FASTA for Streptococcus Pneumoniae Nucleotide:**

```
>NC_003098.1:1412861-1413367 Streptococcus pneumoniae R6
chromosome, complete genome
TTAGACTTCCTTTCTCTTGCGGTATTGGATGGTAAAATCATAAGGATTCTTCTCATCTTTGGC
GTAAAATTTGCTTGAAACTGTCTCAAAAAGAGACAAAGTCAAGCTCTTCAGGGAAATAGGTAT
CTCCTTCCACCCGAGCATGAATGTGAGTGACAATCACTTCATCAAGGTAAGGTTCAAAGC
CTGAAAAATTTGCTTCCCACCGATAATGTAGAGATTCTTTTCTTGATCCTGATACCAGTCAA
GAACAGACTGGACGTCCTGAAAAGTAGCAACCCCATCTATCTTTTCTTCCGGATTACGCGT
CAAATCAGGGTTTCCCGTTTTGGAAGCAAGCGACGCCCATCCCATCAAAGGTCACACG
CCCCATCAAGATAGCATGATTCAGAGTTGTTTCTTTAAAGTGCTGCAATTCTGCTGGCAAAT
GCCAAGGCAGACGATTTTCCTTACCAATCAAACCCTCTTCATCCTGGGCCCAAATAGCTAC
GATTTTCTTAGTCAT
```

### **FASTA for Saccharomyces Nucleotide:**

```
>NC_001147.6:780906-781541 Saccharomyces cerevisiae S288C
chromosome XV, complete sequence
ATGGCTGGAGGAAAGATTCTATTGTAGGAATTGTGGCATGTTTACAGCCGGAGATGGGG
ATAGGATTTTCGTGGAGGTCTACCATGGAGGTTGCCAGTGAAATGAAGTATTTACAGACAGG
TCACTTCATTGACGAAAGATCCAAACAAAAAAATGCTTTGATAATGGGAAGGAAGACATGG
GAATCCATACCGCCCAAGTTTCGCCCACTGCCCAATAGAATGAATGTCATTATATCGAGAA
GCTTCAAGGACGATTTTGTCCACGATAAAGAGAGATCAATAGTCCAAAGTAATTCATTGGCA
AACGCAATAATGAACCTAGAAAGCAATTTTAAGGAGCATCTGGAAAGAATCTACGTGATTG
GGGGTGGCGAAGTTTATAGTCAAATCTTCTCCATTACAGATCATTGGCTCATCACGAAAATA
```

AATCCATTAGATAAAAACGCAACTCCTGCAATGGACACTTTCCTTGATGCGAAGAAATTGGA  
AGAAGTATTTAGCGAGCAAGATCCGGCCCAGCTGAAAGAATTTCTTCCCCCTAAAGTAGAG  
TTGCCCGAAACAGACTGTGATCAACGCTACTCGCTGGAAGAAAAAGGTTATTGCTTCGAAT  
TCACTCTATACAATCGTAAATGA