Біоінформатика.

БАЗИ ДАНИХ. МОДЕЛЬ NCBI

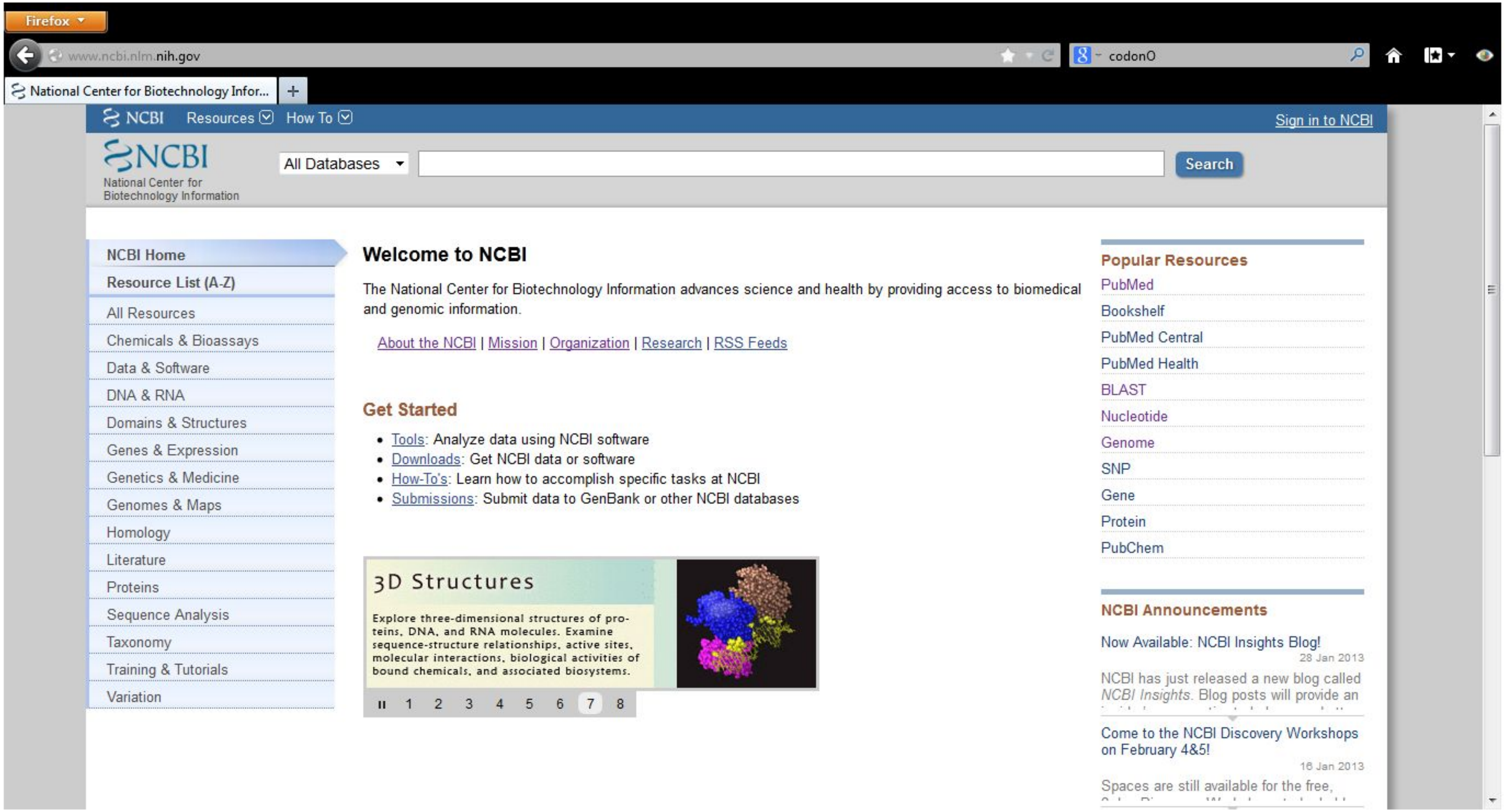
Lab.1

Центральні питання

- ☐ Де знайти літературу про ген/білок, що цікавить
- ☐ Де можна знайти послідовність гена/генома, що цікавить?
- ☐ Що таке депонування НАП?

National Center for Biotechnology Information

- www.ncbi.nlm.nih.gov



Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

About the NCBI | Mission | Organization | Research | RSS Feeds

Get Started

- Tools: Analyze data using NCBI software
- Downloads: Get NCBI data or software
- How-To's: Learn how to accomplish specific tasks at NCBI
- Submissions: Submit data to GenBank or other NCBI databases

3D Structures

Explore three-dimensional structures of proteins, DNA, and RNA molecules. Examine sequence-structure relationships, active sites, molecular interactions, biological activities of bound chemicals, and associated biosystems.

NCBI Announcements

Now Available: NCBI Insights Blog!
28 Jan 2013

NCBI has just released a new blog called *NCBI Insights*. Blog posts will provide an

Come to the NCBI Discovery Workshops on February 4&5!
16 Jan 2013

Spaces are still available for the free,

NCBI

- Фізична локація – м. Бетезда, штат Меріленд (США)
- Подвійна функція – сховище даних і сервер для аналізу

Типи даних

Біомедичні дані

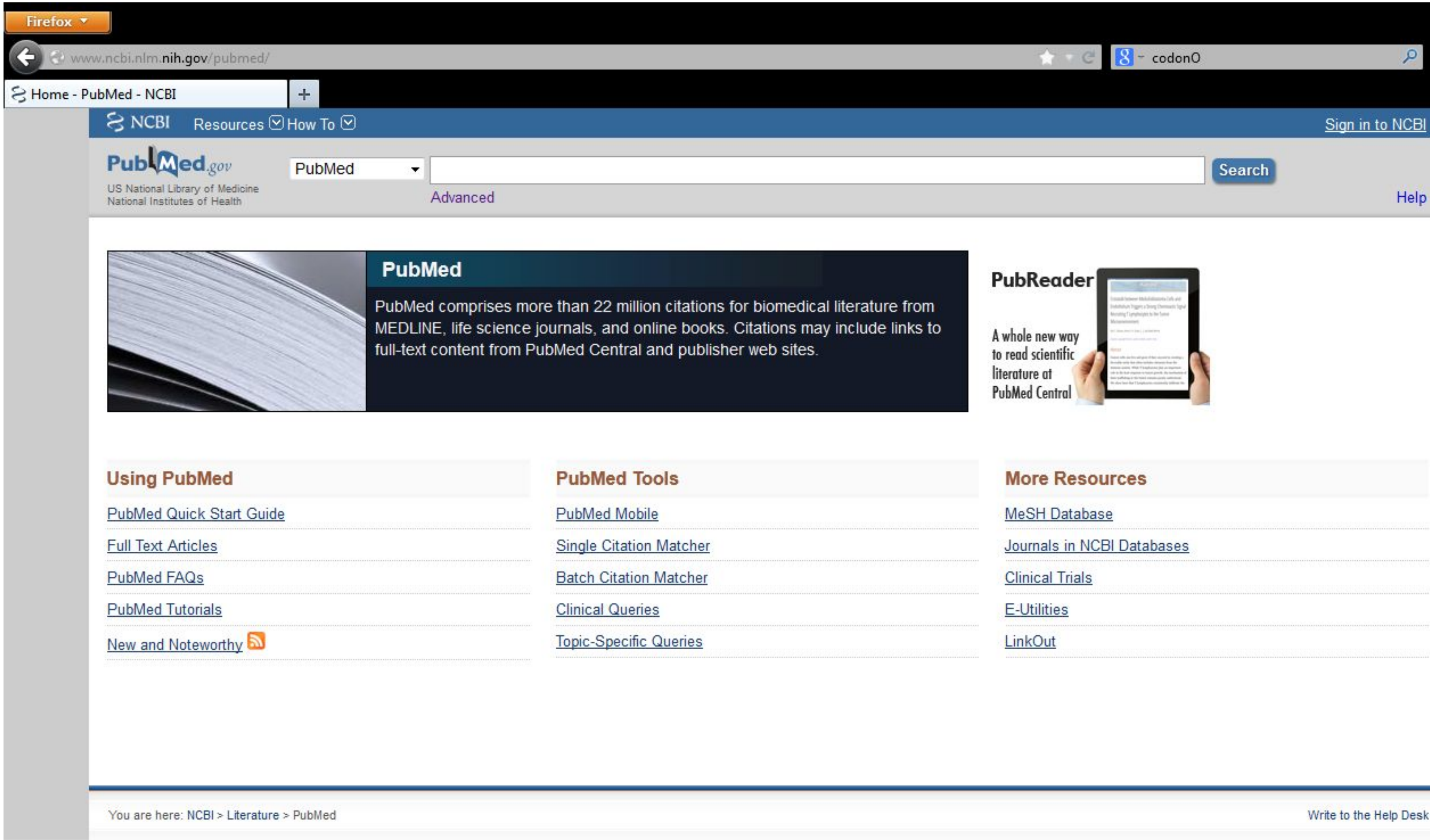
PubMed
PubChem
Taxonomy
Chemicals and assays
Proteins
…..

Геномні дані

Genome
Variations
SNP
OMIM
Genes and expression
…..

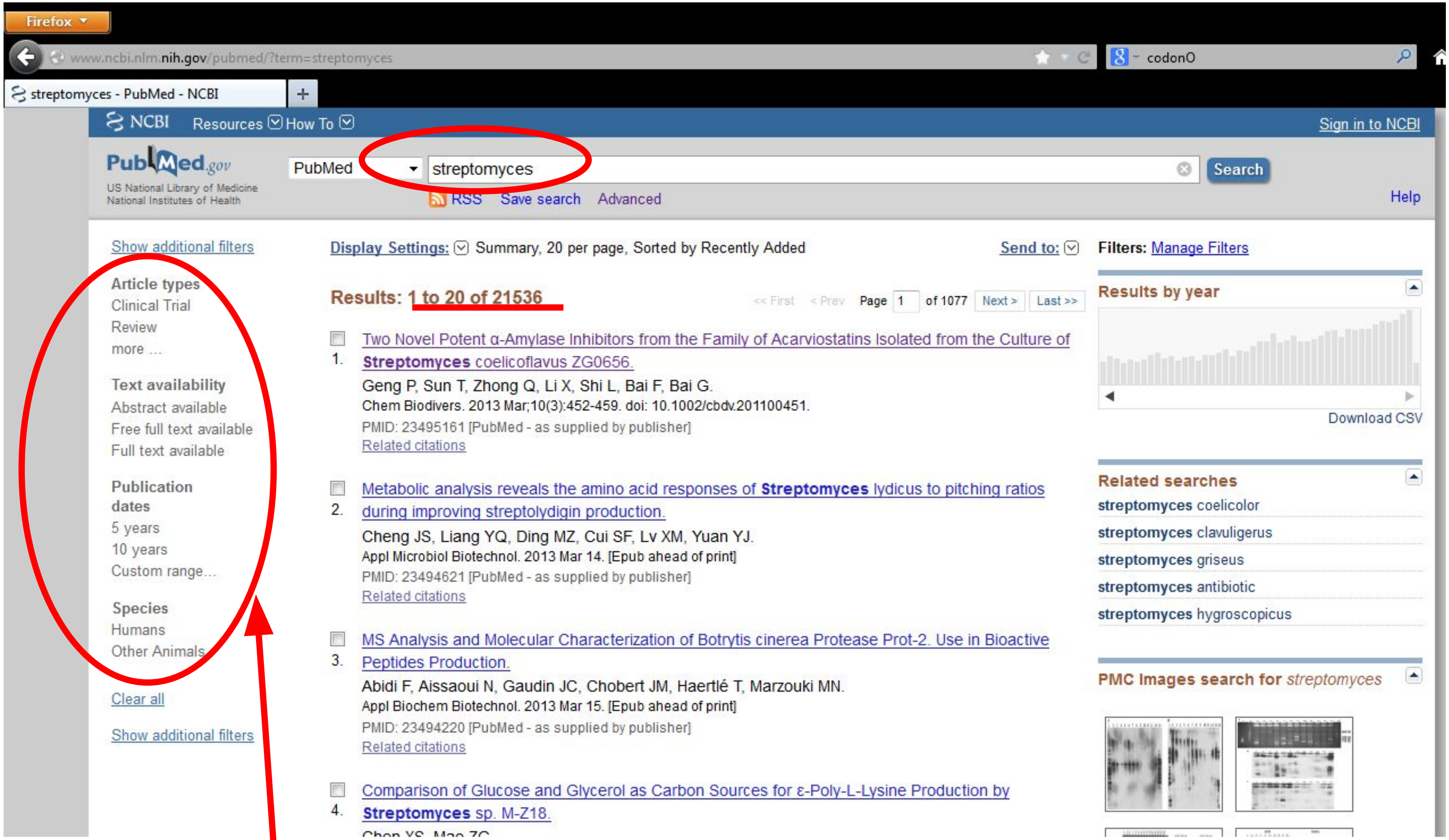
PubMed

- www.ncbi.nlm.nih.gov/pubmed



PubMed comprises more than 22 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites.

PubMed



streptomyces

Display Settings: Summary, 20 per page, Sorted by Recently Added

Results: 1 to 20 of 21536

Article types
Clinical Trial
Review
more ...

Text availability
Abstract available
Free full text available
Full text available

Publication dates
5 years
10 years
Custom range...

Species
Humans
Other Animals

☐ 1. Two Novel Potent α-Amylase Inhibitors from the Family of Acarviostatins Isolated from the Culture of **Streptomyces** coelicoflavus ZG0656.
Geng P, Sun T, Zhong Q, Li X, Shi L, Bai F, Bai G.
Chem Biodivers. 2013 Mar;10(3):452-459. doi: 10.1002/cbdv.201100451.
PMID: 23495161 [PubMed - as supplied by publisher]
Related citations

☐ 2. Metabolic analysis reveals the amino acid responses of **Streptomyces** lydicus to pitching ratios during improving streptolydigin production.
Cheng JS, Liang YQ, Ding MZ, Cui SF, Lv XM, Yuan YJ.
Appl Microbiol Biotechnol. 2013 Mar 14. [Epub ahead of print]
PMID: 23494621 [PubMed - as supplied by publisher]
Related citations

☐ 3. MS Analysis and Molecular Characterization of Botrytis cinerea Protease Prot-2. Use in Bioactive Peptides Production.
Abidi F, Aissaoui N, Gaudin JC, Chobert JM, Haertlé T, Marzouki MN.
Appl Biochem Biotechnol. 2013 Mar 15. [Epub ahead of print]
PMID: 23494220 [PubMed - as supplied by publisher]
Related citations

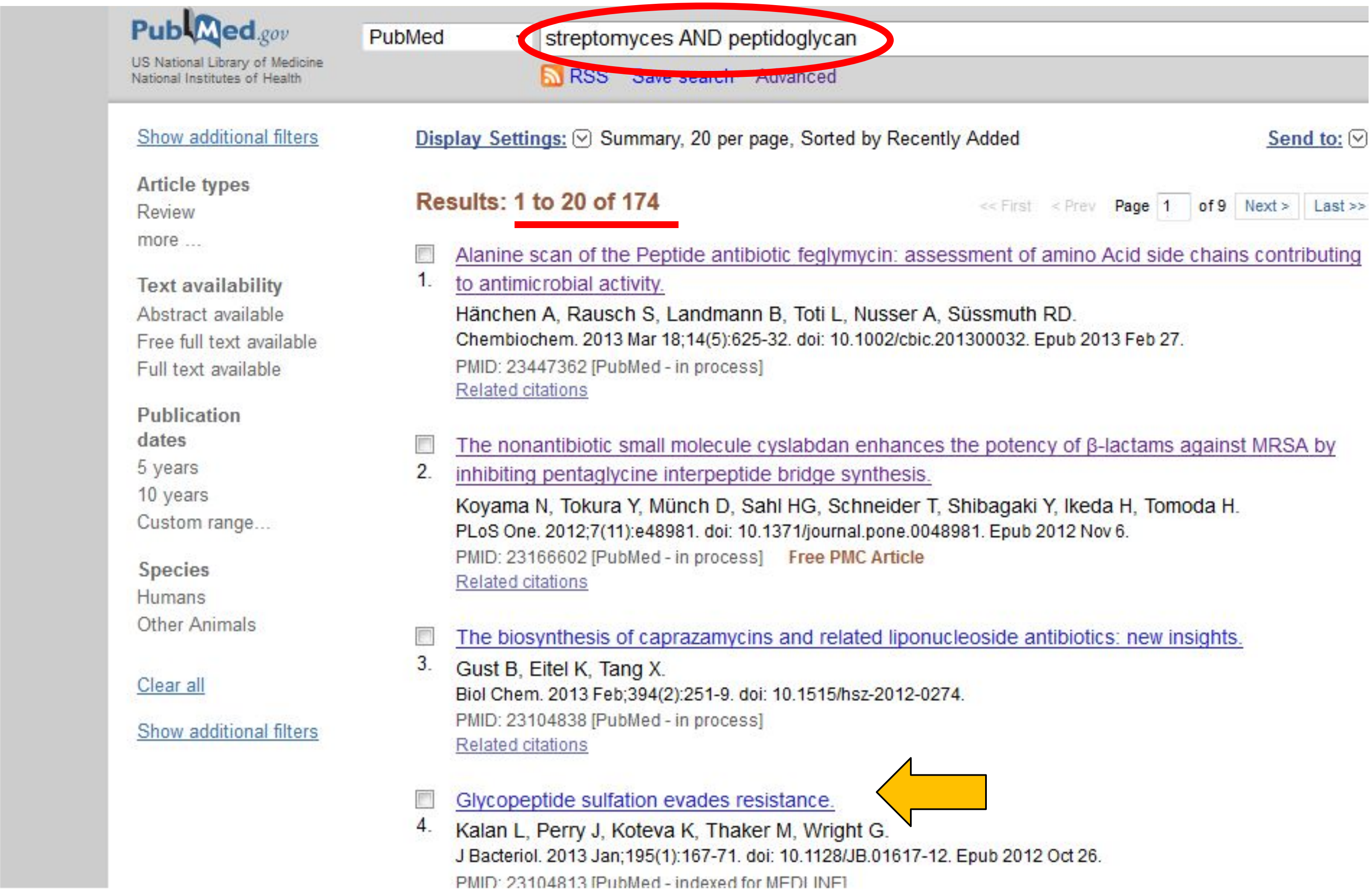
☐ 4. Comparison of Glucose and Glycerol as Carbon Sources for ε-Poly-L-Lysine Production by **Streptomyces** sp. M-Z18.

Related searches
streptomyces coelicolor
streptomyces clavuligerus
streptomyces griseus
streptomyces antibiotic
streptomyces hygroscopicus

Можна звузити коло пошуку до статей, що відповідають певним вимогам (напр., шукати тільки огляди - review)

PubMed

Логічні оператори: AND, OR, *.
Streptomyces; strepto*. * = будь-що



PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed | streptomyces AND peptidoglycan

RSS Save search Advanced

Show additional filters

Article types
Review
more ...

Text availability
Abstract available
Free full text available
Full text available

Publication dates
5 years
10 years
Custom range...

Species
Humans
Other Animals

Clear all

Show additional filters

Display Settings: Summary, 20 per page, Sorted by Recently Added

Send to:

Results: 1 to 20 of 174

<< First < Prev Page 1 of 9 Next > Last >>

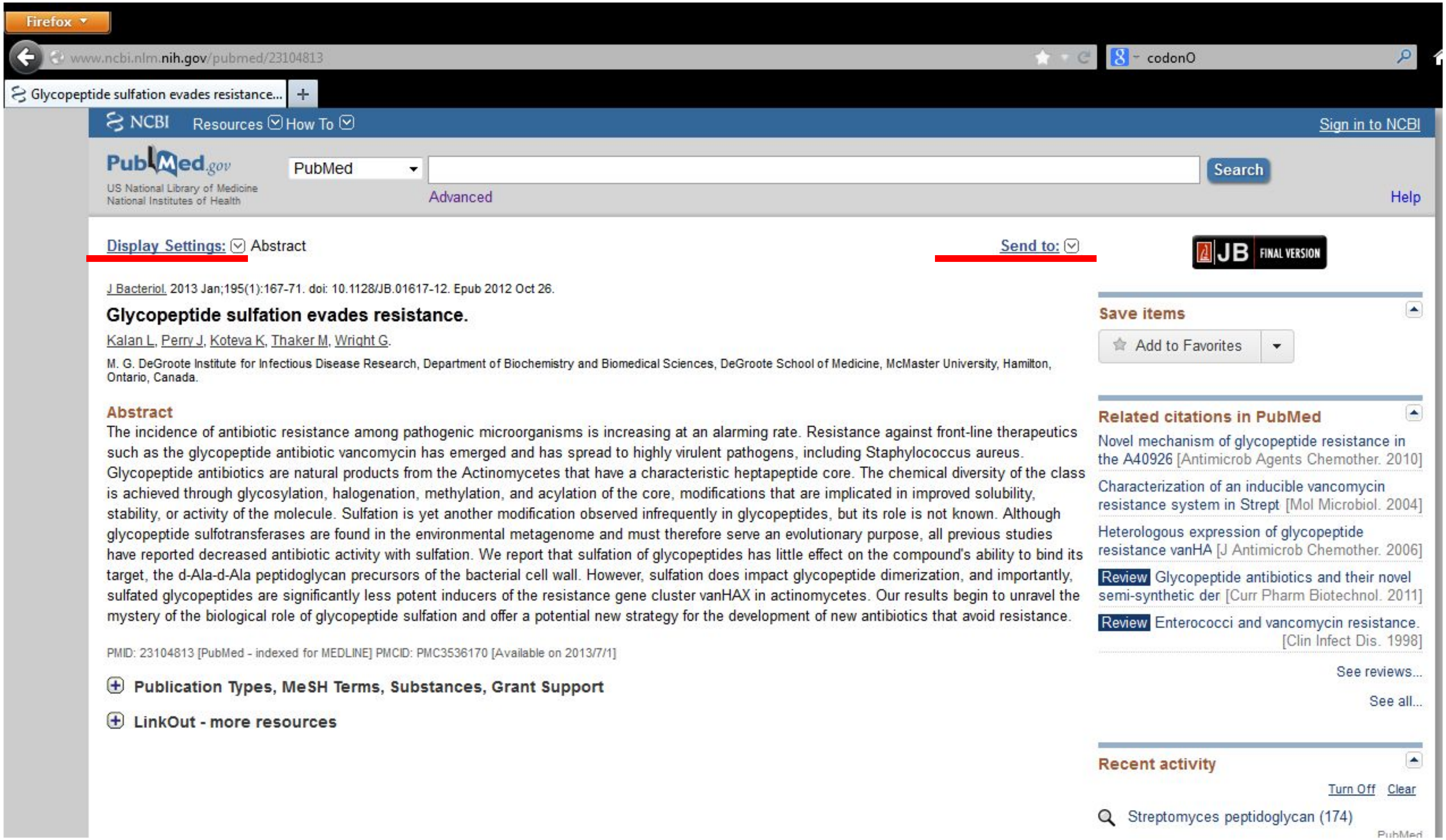
☐ 1. Alanine scan of the Peptide antibiotic feglymycin: assessment of amino Acid side chains contributing to antimicrobial activity.
Hänchen A, Rausch S, Landmann B, Toti L, Nusser A, Süssmuth RD.
Chembiochem. 2013 Mar 18;14(5):625-32. doi: 10.1002/cbic.201300032. Epub 2013 Feb 27.
PMID: 23447362 [PubMed - in process]
Related citations

☐ 2. The nonantibiotic small molecule cyslabdan enhances the potency of β-lactams against MRSA by inhibiting pentaglycine interpeptide bridge synthesis.
Koyama N, Tokura Y, Münch D, Sahl HG, Schneider T, Shibagaki Y, Ikeda H, Tomoda H.
PLoS One. 2012;7(11):e48981. doi: 10.1371/journal.pone.0048981. Epub 2012 Nov 6.
PMID: 23166602 [PubMed - in process] **Free PMC Article**
Related citations

☐ 3. The biosynthesis of caprazamycins and related liponucleoside antibiotics: new insights.
Gust B, Eitel K, Tang X.
Biol Chem. 2013 Feb;394(2):251-9. doi: 10.1515/hsz-2012-0274.
PMID: 23104838 [PubMed - in process]
Related citations

☐ 4. Glycopeptide sulfation evades resistance.
Kalan L, Perry J, Koteva K, Thaker M, Wright G.
J Bacteriol. 2013 Jan;195(1):167-71. doi: 10.1128/JB.01617-12. Epub 2012 Oct 26.
PMID: 23104813 [PubMed - indexed for MEDLINE]

PubMed



Display Settings: Abstract

J Bacteriol. 2013 Jan;195(1):167-71. doi: 10.1128/JB.01617-12. Epub 2012 Oct 26.

Glycopeptide sulfation evades resistance.

Kalan L, Perry J, Koteva K, Thaker M, Wright G.

M. G. DeGroote Institute for Infectious Disease Research, Department of Biochemistry and Biomedical Sciences, DeGroote School of Medicine, McMaster University, Hamilton, Ontario, Canada.

Abstract

The incidence of antibiotic resistance among pathogenic microorganisms is increasing at an alarming rate. Resistance against front-line therapeutics such as the glycopeptide antibiotic vancomycin has emerged and has spread to highly virulent pathogens, including Staphylococcus aureus. Glycopeptide antibiotics are natural products from the Actinomycetes that have a characteristic heptapeptide core. The chemical diversity of the class is achieved through glycosylation, halogenation, methylation, and acylation of the core, modifications that are implicated in improved solubility, stability, or activity of the molecule. Sulfation is yet another modification observed infrequently in glycopeptides, but its role is not known. Although glycopeptide sulfotransferases are found in the environmental metagenome and must therefore serve an evolutionary purpose, all previous studies have reported decreased antibiotic activity with sulfation. We report that sulfation of glycopeptides has little effect on the compound's ability to bind its target, the d-Ala-d-Ala peptidoglycan precursors of the bacterial cell wall. However, sulfation does impact glycopeptide dimerization, and importantly, sulfated glycopeptides are significantly less potent inducers of the resistance gene cluster vanHAX in actinomycetes. Our results begin to unravel the mystery of the biological role of glycopeptide sulfation and offer a potential new strategy for the development of new antibiotics that avoid resistance.

PMID: 23104813 [PubMed - indexed for MEDLINE] PMCID: PMC3536170 [Available on 2013/7/1]

⊕ **Publication Types, MeSH Terms, Substances, Grant Support**

⊕ **LinkOut - more resources**

Related citations in PubMed

- Novel mechanism of glycopeptide resistance in the A40926 [Antimicrob Agents Chemother. 2010]
- Characterization of an inducible vancomycin resistance system in Strept [Mol Microbiol. 2004]
- Heterologous expression of glycopeptide resistance vanHA [J Antimicrob Chemother. 2006]
- Review Glycopeptide antibiotics and their novel semi-synthetic der [Curr Pharm Biotechnol. 2011]
- Review Enterococci and vancomycin resistance. [Clin Infect Dis. 1998]

Recent activity

Streptomyces peptidoglycan (174)

PubMed

Скориставшись лінками, що підкреслено на попередньому слайді, можна отримати бібліографічний опис статті:

1: Kalan L, Perry J, Koteva K, Thaker M, Wright G. Glycopeptide sulfation evades resistance. J Bacteriol. 2013 Jan;195(1):167-71. doi: 10.1128/JB.01617-12. Epub 2012 Oct 26. PubMed PMID: 23104813; PubMed Central PMCID: PMC3536170.

Ідентифікатори статті як цифрового об'єкта (ЦО):

- **DOI (Digital Object Identification) – унікальна комбінація арабських цифр і латинських літер – жоден інший ЦО такої не має. Видавець звертається до www.doi.org, що видає первинний номер (10.1128/JB), який далі продовжується у міру появи нових ЦО. DOI можна присвоїти частинам статті (рисункам)**
- **PMID – присвоюється автоматично кожному посиланню, що потрапляє у базу PubMed**
- **PMCID – номери доступу до повних текстів статей, зібраних із доступних джерел фахівцями NCBI, і поміщених в PubMed Central**

PubMed

Доступ до повних текстів статей

Сайт видавця

J Bacteriol. 2013 Jan;195(1):167-71. doi: 10.1128/JB.01617-12. Epub 2012 Oct 26.

Glycopeptide sulfation evades resistance.

Kalan L, Perry J, Koteva K, Thaker M, Wright G.

M. G. DeGroote Institute for Infectious Disease Research, Department of Biochemistry and Biomedical Sciences, DeGroote School of Medicine, McMaster University, Hamilton, Ontario, Canada.

Abstract

The incidence of antibiotic resistance among pathogenic microorganisms is increasing at an alarming rate. Resistance against front-line therapeutics such as the glycopeptide antibiotic vancomycin has emerged and has spread to highly virulent pathogens, including Staphylococcus aureus. Glycopeptide antibiotics are natural products from the Actinomycetes that have a characteristic heptapeptide core. The chemical diversity of the class is achieved through glycosylation, halogenation, methylation, and acylation of the core, modifications that are implicated in improved solubility, stability, or activity of the molecule. Sulfation is yet another modification observed infrequently in glycopeptides, but its role is not known. Although glycopeptide sulfotransferases are found in the environmental metagenome and must therefore serve an evolutionary purpose, all previous studies have reported decreased antibiotic activity with sulfation. We report that sulfation of glycopeptides has little effect on the compound's ability to bind its target, the d-Ala-d-Ala peptidoglycan precursors of the bacterial cell wall. However, sulfation does impact glycopeptide dimerization, and importantly, sulfated glycopeptides are significantly less potent inducers of the resistance gene cluster vanHAX in actinomycetes. Our results begin to unravel the mystery of the biological role of glycopeptide sulfation and offer a potential new strategy for the development of new antibiotics that avoid resistance.

PMID: 23104813 [PubMed - indexed for MEDLINE] PMCID: PMC3536170 [Available on 2013/7/1]

⊕ **Publication Types, MeSH Terms, Substances, Grant Support**

⊕ **LinkOut - more resources**

Related citations in PubMed

Novel mechanism of glycopeptide resistance in the A40926 [Antimicrob Agents Chemother. 2010]

Characterization of an inducible vancomycin resistance system in Strept [Mol Microbiol. 2004]

Heterologous expression of glycopeptide resistance vanHA [J Antimicrob Chemother. 2006]

Review Glycopeptide antibiotics and their novel semi-synthetic der [Curr Pharm Biotechnol. 2011]

Review Enterococci and vancomycin resistance. [Clin Infect Dis. 1998]

See reviews...

See all...

PubMed

Деякі видання у галузі хімії не реферуються в PubMed

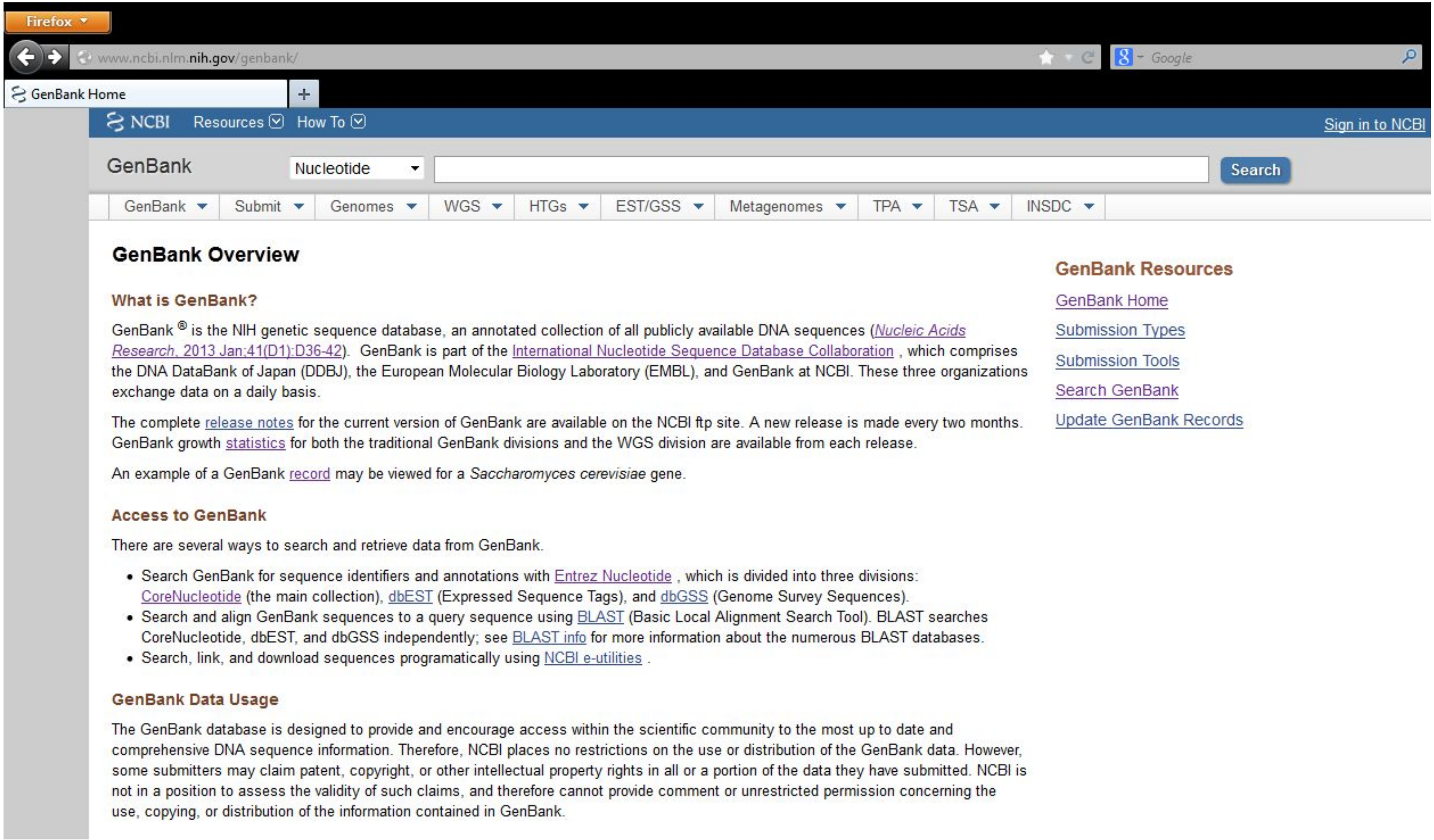
Видавничий дім Elsevier

- ☐ **Журнали Open Access**
- ☐ **NIH вимагає відкривати зміст усіх статей, де описано досліди, що ними фінансувались – через певний період часу (зразу, півроку, рік)**
- ☐ **Запит до автора статті (може переслати персональну копію для особистого вжитку)**
- ☐ **Пошук через Google**
- ☐ **Запит через наукові соцмережі (LinkedIn, ResearchGate)**

GenBank

- ☐ Ця база є частиною міжнародного консорціуму INSDC (International Nucleotide Sequence Databases Collaboration; http://www.insdc.org/), куди входять European Nucleotide Archive (ENA) і DNA Data Base of Japan (DDBJ).
- ☐ GenBank – підрозділ NCBI. www.ncbi.nlm.nih.gov/genbank
- ☐ Т.зв. первинна, або архівна, база даних – метою є збереження нуклеотидних послідовностей (і продуктів їхньої концептуальної трансляції), що подаються (депонуються) дослідниками з усього світу. Зміст поданої інформації загалом не перевіряється – це відповідальність авторів депонування
- ☐ GenBank, ENA і DDBJ щоденно обмінюються інформацією, і зміст їхніх баз ідентичний

GenBank



GenBank Overview

What is GenBank?

GenBank ® is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences (*Nucleic Acids Research*, 2013 Jan;41(D1):D36-42). GenBank is part of the International Nucleotide Sequence Database Collaboration , which comprises the DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI. These three organizations exchange data on a daily basis.

The complete release notes for the current version of GenBank are available on the NCBI ftp site. A new release is made every two months. GenBank growth statistics for both the traditional GenBank divisions and the WGS division are available from each release.

An example of a GenBank record may be viewed for a *Saccharomyces cerevisiae* gene.

Access to GenBank

There are several ways to search and retrieve data from GenBank.

- Search GenBank for sequence identifiers and annotations with Entrez Nucleotide , which is divided into three divisions: CoreNucleotide (the main collection), dbEST (Expressed Sequence Tags), and dbGSS (Genome Survey Sequences).
- Search and align GenBank sequences to a query sequence using BLAST (Basic Local Alignment Search Tool). BLAST searches CoreNucleotide, dbEST, and dbGSS independently; see BLAST info for more information about the numerous BLAST databases.
- Search, link, and download sequences programatically using NCBI e-utilities .

GenBank Data Usage

The GenBank database is designed to provide and encourage access within the scientific community to the most up to date and comprehensive DNA sequence information. Therefore, NCBI places no restrictions on the use or distribution of the GenBank data. However, some submitters may claim patent, copyright, or other intellectual property rights in all or a portion of the data they have submitted. NCBI is not in a position to assess the validity of such claims, and therefore cannot provide comment or unrestricted permission concerning the use, copying, or distribution of the information contained in GenBank.

GenBank Resources

- GenBank Home
- Submission Types
- Submission Tools
- Search GenBank
- Update GenBank Records

Анотовані фрагменти геномів; анотовані геноми; неанотовані фрагменти і геноми; чернетки (драфти геномів), метагеномні дані….

GenBank

Формат FASTA

Шрифт `Courier New`

```
>PMI3114
TTAAGAAGCAATAGGTGTCCATAGTTCACCTAATTTATAGTCCCACT
CTTTTGGTTTAATGGGAATGCGCCCCGCATTGGGGGTAAAGGTCATT
TCACTAAACAGTAGTGTGGTTTCTGACATTAAAAAATCGACCCGACA
ATAGGCAAAACCGTTAGCTAATTTGTTGGCTAATGTGAGCATATTGT
CAAATTGTAGCGGTTTTTCAATATATTCTGGGGTATTGGGATCTTCT
AAAGTAAAAGGCTGTAATTGCCATTGTGTATCATAAACATTGATA
```

Рядок анотації може займати кілька рядків, містити пробіли

```
>WemR
MKKLKKYLTRKKKENYIIFSIYYFIKVTSSIFISDSLYRKYIFKRKY
KLNLKKPTSFNEKIHYRILNDHNPIYTKLADKLLVRDYVREKIGEKY
LIKLINHYNTPSEINFNTLPKSFVLKCNHDVGSVMIINDKSKINEKA
IKKKLKIALKNNIYYQNREWHYKNIKPKIICEELINIFPHNKKNYPE
DYKIHCFNGIPRYIELQFSRFSHDRRINIYDFNWNLQPFLMGYKNTN
ESIEKPKKLQEIYNISKTLSADFDYCRVDFYITP
```

GenBank

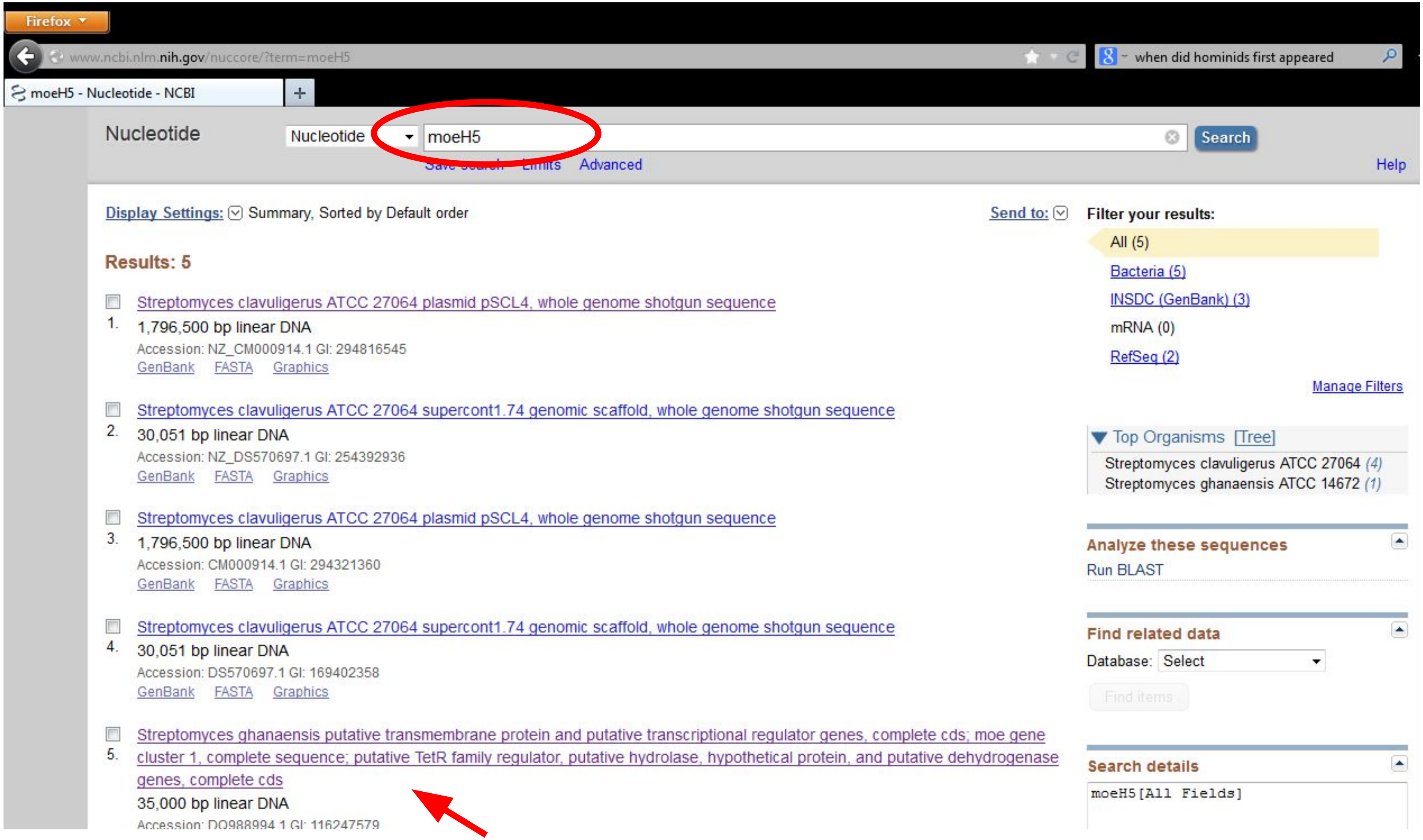
Gene identification (gi) – унікальний номер, що присвоюється кожній послідовності при поданні у GenBank. Після двокрапки – координати у геномі (у п.н.)

Анотація

Locus name

```
>gi|197283915:3424858-3425733|PMI3114| Proteus
mirabilis HI4320 chromosome, complete genome
TTAAGAAGCAATAGGTGTCCATAGTTCACCTAATTTATAGTCCCAC
TCTTTTGGTTTAATGGGAATGCGCCCCGCATTGGGGGTAAAGGTCA
TTTCACTAAACAGTAGTGTGGTTTCTGACATTAAAAAATCGACCCG
ACAATAGGCAAAACCGTTAGCTAATTTGTTGGCTAATGTGAGCATA
TTGTCAAATTGTAGCGGTTTTTCAATATATTCTGGGTATTGGGAT
CTTCTAAAGTAAAAGGCTGTAATTGCCATTGTGTATCATAAACATT
GATA
```

GenBank



Display Settings: Summary, Sorted by Default order

Send to:

Results: 5

1. ☐ Streptomyces clavuligerus ATCC 27064 plasmid pSCL4, whole genome shotgun sequence
   1,796,500 bp linear DNA
   Accession: NZ_CM000914.1 GI: 294816545
   GenBank FASTA Graphics

2. ☐ Streptomyces clavuligerus ATCC 27064 supercont1.74 genomic scaffold, whole genome shotgun sequence
   30,051 bp linear DNA
   Accession: NZ_DS570697.1 GI: 254392936
   GenBank FASTA Graphics

3. ☐ Streptomyces clavuligerus ATCC 27064 plasmid pSCL4, whole genome shotgun sequence
   1,796,500 bp linear DNA
   Accession: CM000914.1 GI: 294321360
   GenBank FASTA Graphics

4. ☐ Streptomyces clavuligerus ATCC 27064 supercont1.74 genomic scaffold, whole genome shotgun sequence
   30,051 bp linear DNA
   Accession: DS570697.1 GI: 169402358
   GenBank FASTA Graphics

5. ☐ Streptomyces ghanaensis putative transmembrane protein and putative transcriptional regulator genes, complete cds; moe gene cluster 1, complete sequence; putative TetR family regulator, putative hydrolase, hypothetical protein, and putative dehydrogenase genes, complete cds
   35,000 bp linear DNA
   Accession: DQ988994.1 GI: 116247579

Filter your results:

All (5)
Bacteria (5)
INSDC (GenBank) (3)
mRNA (0)
RefSeq (2)

Manage Filters

Top Organisms [Tree]

Streptomyces clavuligerus ATCC 27064 (4)
Streptomyces ghanaensis ATCC 14672 (1)

Analyze these sequences

Run BLAST

Find related data

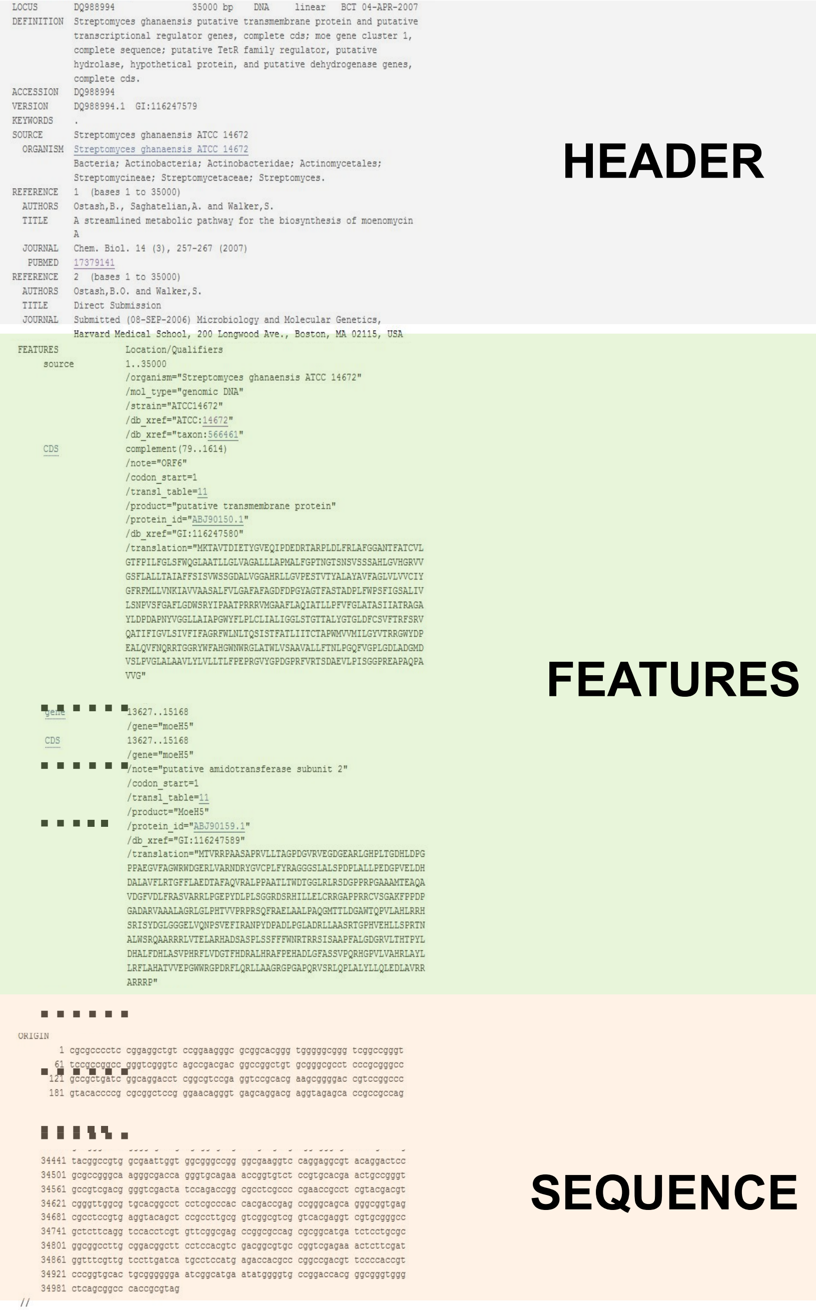
Database: Select

Find items

Search details

moeH5[All Fields]

GenBank



```
LOCUS       DQ988994               35000 bp    DNA     linear   BCT 04-APR-2007
DEFINITION  Streptomyces ghanaensis putative transmembrane protein and putative
            transcriptional regulator genes, complete cds; moe gene cluster 1,
            complete sequence; putative TetR family regulator, putative
            hydrolase, hypothetical protein, and putative dehydrogenase genes,
            complete cds.
ACCESSION   DQ988994
VERSION     DQ988994.1  GI:116247579
KEYWORDS    .
SOURCE      Streptomyces ghanaensis ATCC 14672
  ORGANISM  Streptomyces ghanaensis ATCC 14672
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE   1  (bases 1 to 35000)
  AUTHORS   Ostash,B., Saghatelian,A. and Walker,S.
  TITLE     A streamlined metabolic pathway for the biosynthesis of moenomycin
            A
  JOURNAL   Chem. Biol. 14 (3), 257-267 (2007)
   PUBMED   17379141
REFERENCE   2  (bases 1 to 35000)
  AUTHORS   Ostash,B.O. and Walker,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2006) Microbiology and Molecular Genetics,
            Harvard Medical School, 200 Longwood Ave., Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1..35000
                     /organism="Streptomyces ghanaensis ATCC 14672"
                     /mol_type="genomic DNA"
                     /strain="ATCC14672"
                     /db_xref="ATCC:14672"
                     /db_xref="taxon:566461"
     CDS             complement(79..1614)
                     /note="ORF6"
                     /codon_start=1
                     /transl_table=11
                     /product="putative transmembrane protein"
                     /protein_id="ABJ90150.1"
                     /db_xref="GI:116247580"
                     /translation="MKTAVTDIETYGVEQIPDEDRTARPLDLFRLAFGGANTFATCVL
                     GTFPILFGLSFWQGLAATLLGLVAGALLLAPMALFGPTNGTSNSVSSSAHLGVHGRVV
                     GSFLALLTAIAFFSISVWSSGDALVGGAHRLLGVPESTVTYALAYAVFAGLVLVVCIY
                     GFRFMLLVNKIAVVAASALFVLGAFAFAGDFDPGYAGTFASTADPLFWPSFIGSALIV
                     LSNPVSFGAFLGDWSRYIPAATPRRRVMGAAFLAQIATLLPFVFGLATASIIATRAGA
                     YLDPDAPNYVGGLLAIAPGWYFLPLCLIALIGGLSTGTTALYGTGLDFCSVFTRFSRV
                     QATIFIGVLSIVFIFAGRFWLNLTQSISTFATLIITCTAPWMVVMILGYVTRRGWYDP
                     EALQVFNQRRTGGRYWFAHGWNWRGLATWLVSAAVALLFTNLPGQFVGPLGDLADGMD
                     VSLPVGLALAAVLYLVLLTLFPEPRGVYGPDGPRFVRTSDAEVLPISGGPREAPAQPA
                     VVG"

     ......
     gene            13627..15168
                     /gene="moeH5"
     CDS             13627..15168
                     /gene="moeH5"
     ......          /note="putative amidotransferase subunit 2"
                     /codon_start=1
                     /transl_table=11
                     /product="MoeH5"
     .....           /protein_id="ABJ90159.1"
                     /db_xref="GI:116247589"
                     /translation="MTVRRPAASAPRVLLTAGPDGVRVEGDGEARLGHPLTGDHLDPG
                     PPAEGVFAGWRWDGERLVARNDRYGVCPLFYRAGGGSLALSPDPLALLPEDGPVELDH
                     DALAVFLRTGFFLAEDTAFAQVRALPPAATLTWDTGGLRLRSDGPPRPGAAAMTEAQA
                     VDGFVDLFRASVARRLPGEPYDLPLSGGRDSRHILLELCRRGAPPRRCVSGAKFPPDP
                     GADARVAAALAGRLGLPHTVVPRPRSQFRAELAALPAQGMTTLDGAWTQPVLAHLRRH
                     SRISYDGLGGGELVQNPSVEFIRANPYDPADLPGLADRLLAASRTGPHVEHLLSPRTN
                     ALWSRQAARRRLVTELARHADSASPLSSFFFWNRTRRSISAAPFALGDGRVLTHTPYL
                     DHALFDHLASVPHRFLVDGTFHDRALHRAFPEHADLGFASSVPQRHGPVLVAHRLAYL
                     LRFLAHATVVEPGWWRGPDRFLQRLLAAGRGPGAPQRVSRLQPLALYLLQLEDLAVRR
                     ARRRP"

     ......
ORIGIN
        1 cgcgcccctc cggaggctgt ccggaagggc gcggcacggg tgggggcggg tcggccgggt
       61 tccgccggcc gggtcgggtc agccgacgac ggccggctgt gcgggcgcct cccgcgggcc
     ......
      121 gccgctgatc ggcaggacct cggcgtccga ggtccgcacg aagcggggac cgtccggccc
      181 gtacaccccg cgcggctccg ggaacagggt gagcaggacg aggtagagca ccgccgccag

     ......
    34441 tacggccgtg gcgaattggt ggcgggccgg ggcgaaggtc caggaggcgt acaggactcc
    34501 gcgccgggca agggcgacca gggtgcagaa accggtgtct ccgtgcacga actgccgggt
    34561 gccgtcgacg gggtcgacta tccagaccgg cgcctcgccc cgaaccgcct cgtacgacgt
    34621 cgggttggcg tgcacggcct cctcgcccac cacgaccgag ccgggcagca gggcggtgag
    34681 cgcctccgtg aggtacagct ccgccttgcg gtcggcgtcg gtcacgaggt cgtgcgggcc
    34741 gctcttcagg tccacctcgt gttcggcgag ccggcgccag cgcggcatga tctcctgcgc
    34801 ggcggccttg cggacggctt cctccacgtc gacggcgtgc cggtcgagaa actcttcgat
    34861 ggtttcgttg tccttgatca tgcctccatg agaccacgcc cggccgacgt tccccaccgt
    34921 cccggtgcac tgcgggggga atcggcatga atatggggtg ccggaccacg ggcgggtggg
    34981 ctcagcggcc caccgcgtag
//
```

HEADER

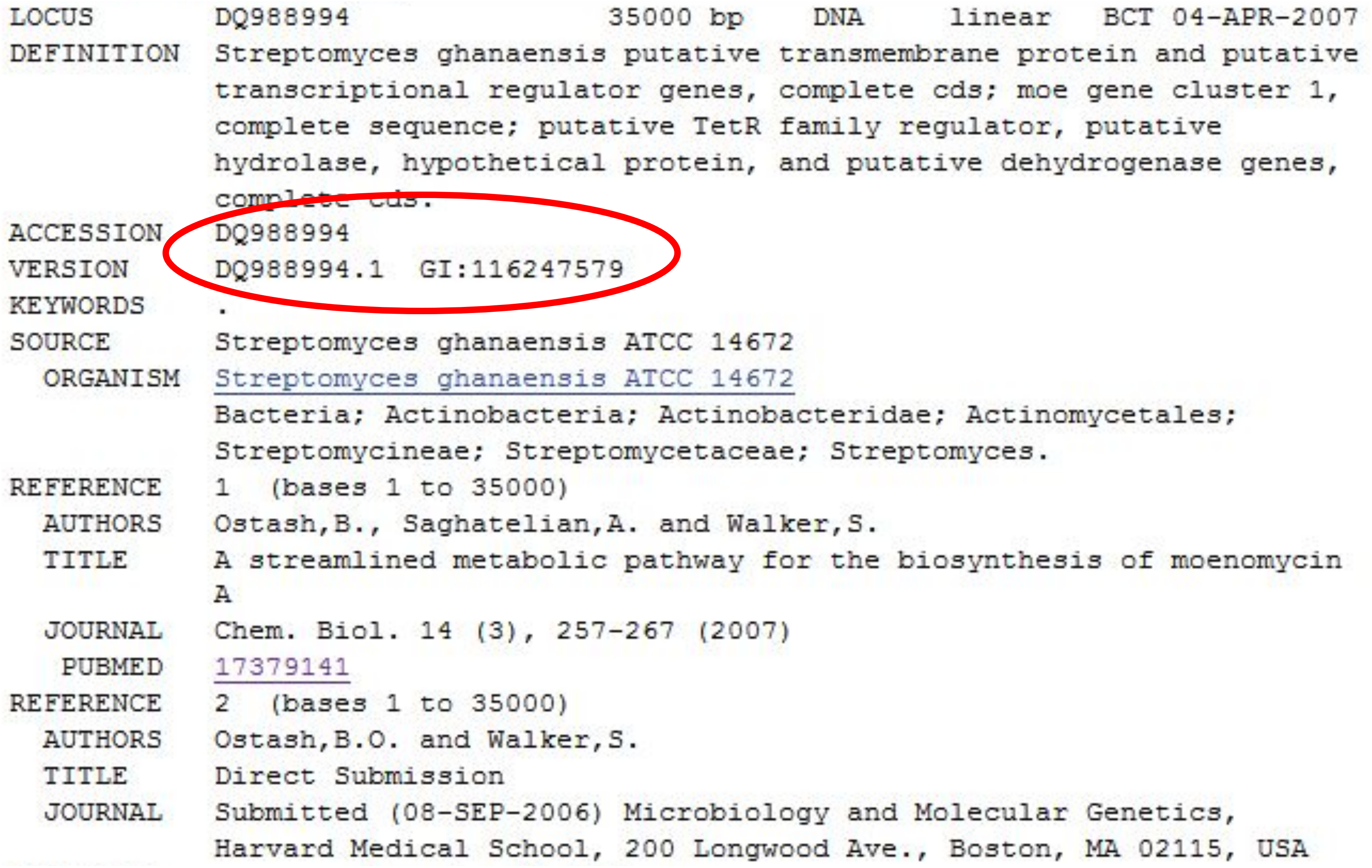
FEATURES

SEQUENCE

Структура GenBank FlatFile (GBFF) - флетфайл

GenBank

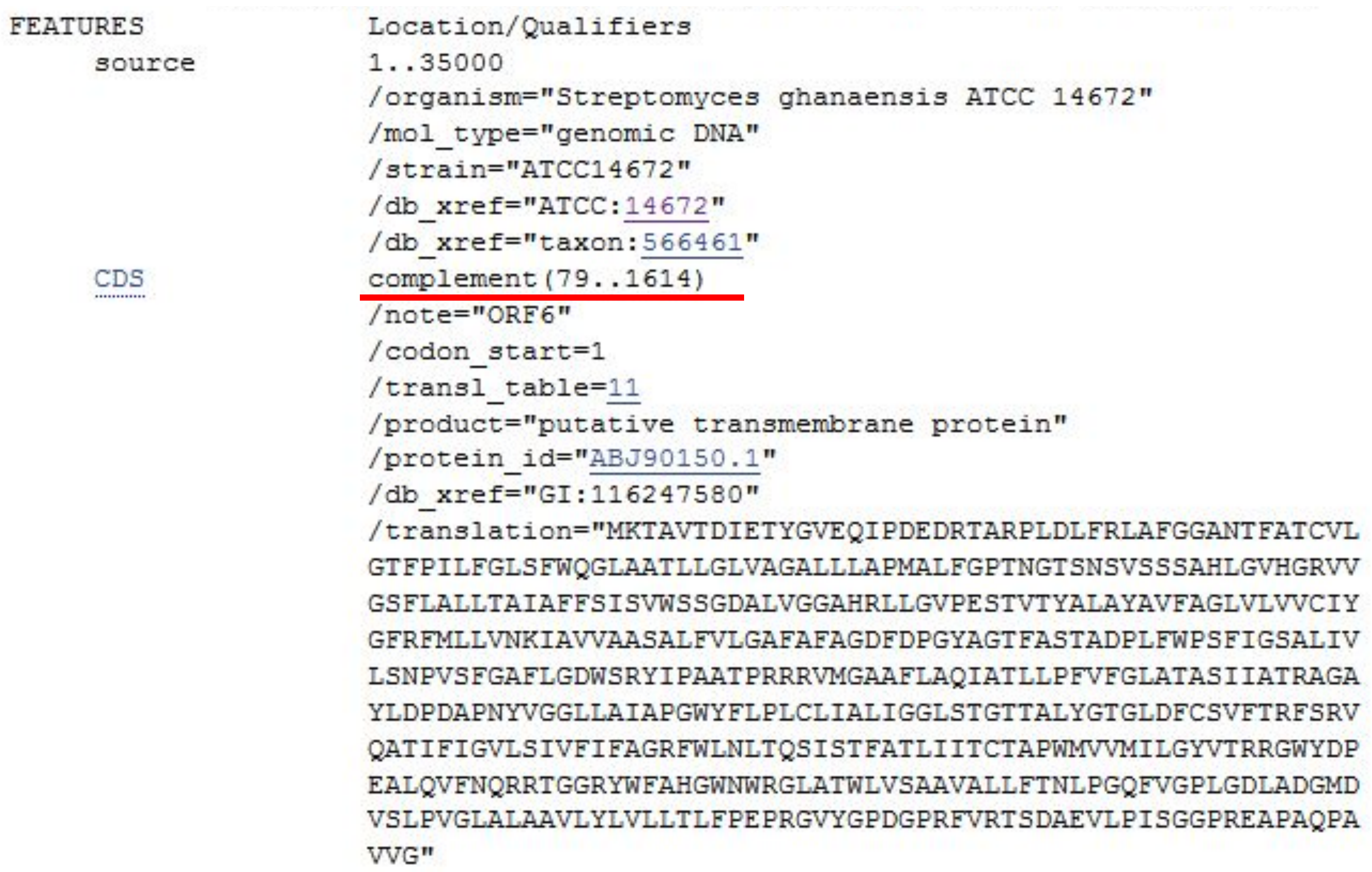
HEADER



```
LOCUS       DQ988994                35000 bp    DNA     linear   BCT 04-APR-2007
DEFINITION  Streptomyces ghanaensis putative transmembrane protein and putative
            transcriptional regulator genes, complete cds; moe gene cluster 1,
            complete sequence; putative TetR family regulator, putative
            hydrolase, hypothetical protein, and putative dehydrogenase genes,
            complete cds.
ACCESSION   DQ988994
VERSION     DQ988994.1  GI:116247579
KEYWORDS    .
SOURCE      Streptomyces ghanaensis ATCC 14672
  ORGANISM  Streptomyces ghanaensis ATCC 14672
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE   1  (bases 1 to 35000)
  AUTHORS   Ostash,B., Saghatelian,A. and Walker,S.
  TITLE     A streamlined metabolic pathway for the biosynthesis of moenomycin
            A
  JOURNAL   Chem. Biol. 14 (3), 257-267 (2007)
   PUBMED   17379141
REFERENCE   2  (bases 1 to 35000)
  AUTHORS   Ostash,B.O. and Walker,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2006) Microbiology and Molecular Genetics,
            Harvard Medical School, 200 Longwood Ave., Boston, MA 02115, USA
```

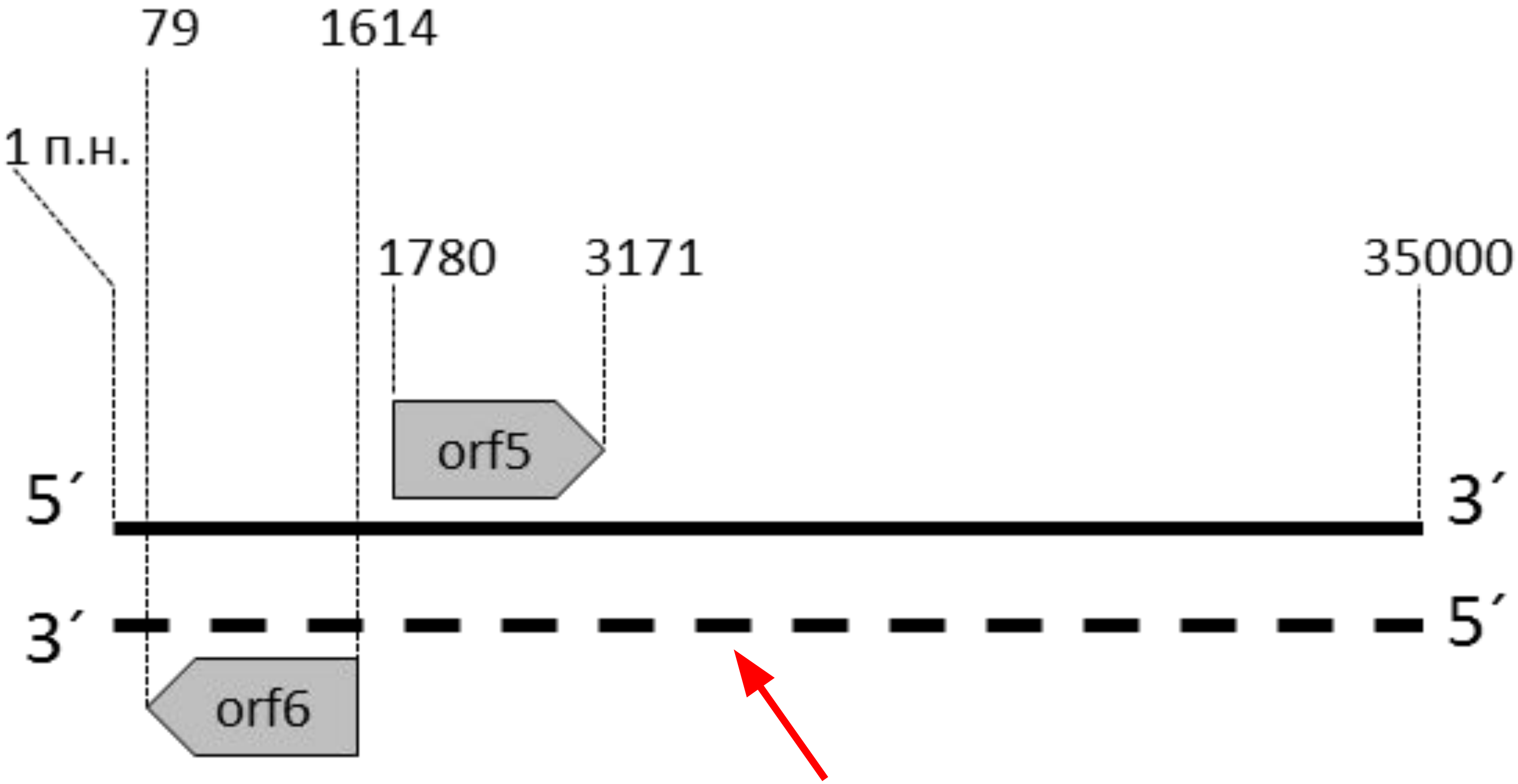
GenBank

FEATURES



```
FEATURES             Location/Qualifiers
     source          1..35000
                     /organism="Streptomyces ghanaensis ATCC 14672"
                     /mol_type="genomic DNA"
                     /strain="ATCC14672"
                     /db_xref="ATCC:14672"
                     /db_xref="taxon:566461"
     CDS             complement(79..1614)
                     /note="ORF6"
                     /codon_start=1
                     /transl_table=11
                     /product="putative transmembrane protein"
                     /protein_id="ABJ90150.1"
                     /db_xref="GI:116247580"
                     /translation="MKTAVTDIETYGVEQIPDEDRTARPLDLFRLAFGGANTFATCVL
                     GTFPILFGLSFWQGLAATLLGLVAGALLLAPMALFGPTNGTSNSVSSSAHLGVHGRVV
                     GSFLALLTAIAFFSISVWSSGDALVGGAHRLLGVPESTVTYALAYAVFAGLVLVVCIY
                     GFRFMLLVNKIAVVAASALFVLGAFAFAGDFDPGYAGTFASTADPLFWPSFIGSALIV
                     LSNPVSFGAFLGDWSRYIPAATPRRRVMGAAFLAQIATLLPFVFGLATASIIATRAGA
                     YLDPDAPNYVGGLLAIAPGWYFLPLCLIALIGGLSTGTTALYGTGLDFCSVFTRFSRV
                     QATIFIGVLSIVFIFAGRFWLNLTQSISTFATLIITCTAPWMVVMILGYVTRRGWYDP
                     EALQVFNQRRTGGRYWFAHGWNWRGLATWLVSAAVALLFTNLPGQFVGPLGDLADGMD
                     VSLPVGLALAAVLYLVLLTLFPEPRGVYGPDGPRFVRTSDAEVLPISGGPREAPAQPA
                     VVG"
```

GenBank



Комплементарна нитка

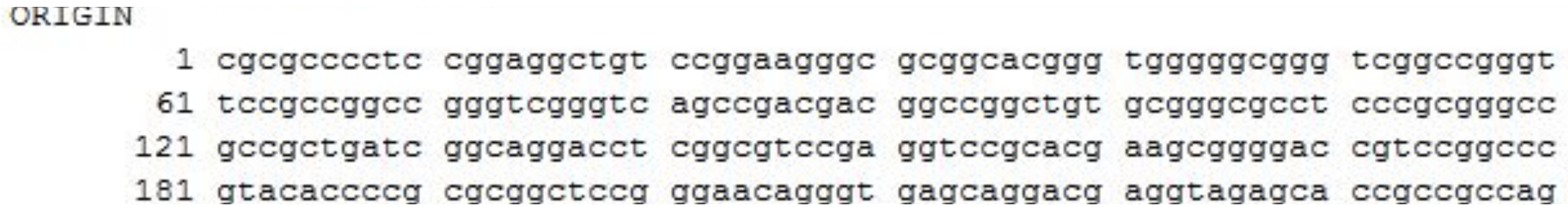
GenBank

FEATURES

```
gene            13627..15168
                /gene="moeH5"
CDS             13627..15168
                /gene="moeH5"
                /note="putative amidotransferase subunit 2"
                /codon_start=1
                /transl_table=11
                /product="MoeH5"
                /protein_id="ABJ90159.1"
                /db_xref="GI:116247589"
                /translation="MTVRRPAASAPRVLLTAGPDGVRVEGDGEARLGHPLTGDHLDPG
                PPAEGVFAGWRWDGERLVARNDRYGVCPLFYRAGGGSLALSPDPLALLPEDGPVELDH
                DALAVFLRTGFFLAEDTAFAQVRALPPAATLTWDTGGLRLRSDGPPRPGAAAMTEAQA
                VDGFVDLFRASVARRLPGEPYDLPLSGGRDSRHILLELCRRGAPPRRCVSGAKFPPDP
                GADARVAAALAGRLGLPHTVVPRPRSQFRAELAALPAQGMTTLDGAWTQPVLAHLRRH
                SRISYDGLGGGELVQNPSVEFIRANPYDPADLPGLADRLLAASRTGPHVEHLLSPRTN
                ALWSRQAARRRLVTELARHADSASPLSSFFFWNRTRRSISAAPFALGDGRVLTHTPYL
                DHALFDHLASVPHRFLVDGTFHDRALHRAFPEHADLGFASSVPQRHGPVLVAHRLAYL
                LRFLAHATVVEPGWWRGPDRFLQRLLAAGRGPGAPQRVSRLQPLALYLLQLEDLAVRR
                ARRRP"
```

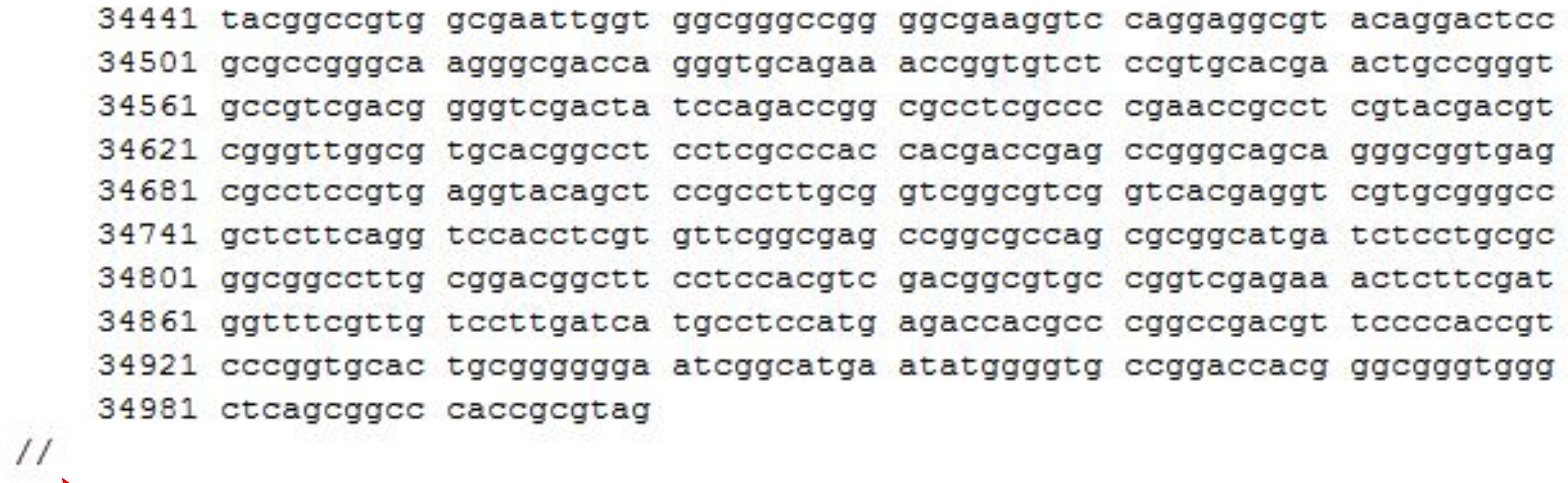
GenBank

ORIGIN



```
ORIGIN
        1 cgcgcccctc cggaggctgt ccggaagggc gcggcacggg tgggggcggg tcggccgggt
       61 tccgccggcc gggtcgggtc agccgacgac ggccggctgt gcgggcgcct cccgcgggcc
      121 gccgctgatc ggcaggacct cggcgtccga ggtccgcacg aagcggggac cgtccggccc
      181 gtacaccccg cgcggctccg ggaacagggt gagcaggacg aggtagagca ccgccgccag
```

………………



```
    34441 tacggccgtg gcgaattggt ggcgggccgg ggcgaaggtc caggaggcgt acaggactcc
    34501 gcgccgggca agggcgacca gggtgcagaa accggtgtct ccgtgcacga actgccgggt
    34561 gccgtcgacg gggtcgacta tccagaccgg cgcctcgccc cgaaccgcct cgtacgacgt
    34621 cgggttggcg tgcacggcct cctcgcccac cacgaccgag ccgggcagca gggcggtgag
    34681 cgcctccgtg aggtacagct ccgccttgcg gtcggcgtcg gtcacgaggt cgtgcgggcc
    34741 gctcttcagg tccacctcgt gttcggcgag ccggcgccag cgcggcatga tctcctgcgc
    34801 ggcggccttg cggacggctt cctccacgtc gacggcgtgc cggtcgagaa actcttcgat
    34861 ggtttcgttg tccttgatca tgcctccatg agaccacgcc cggccgacgt tccccaccgt
    34921 cccggtgcac tgcgggggga atcggcatga atatggggtg ccggaccacg ggcgggtggg
    34981 ctcagcggcc caccgcgtag
//
```

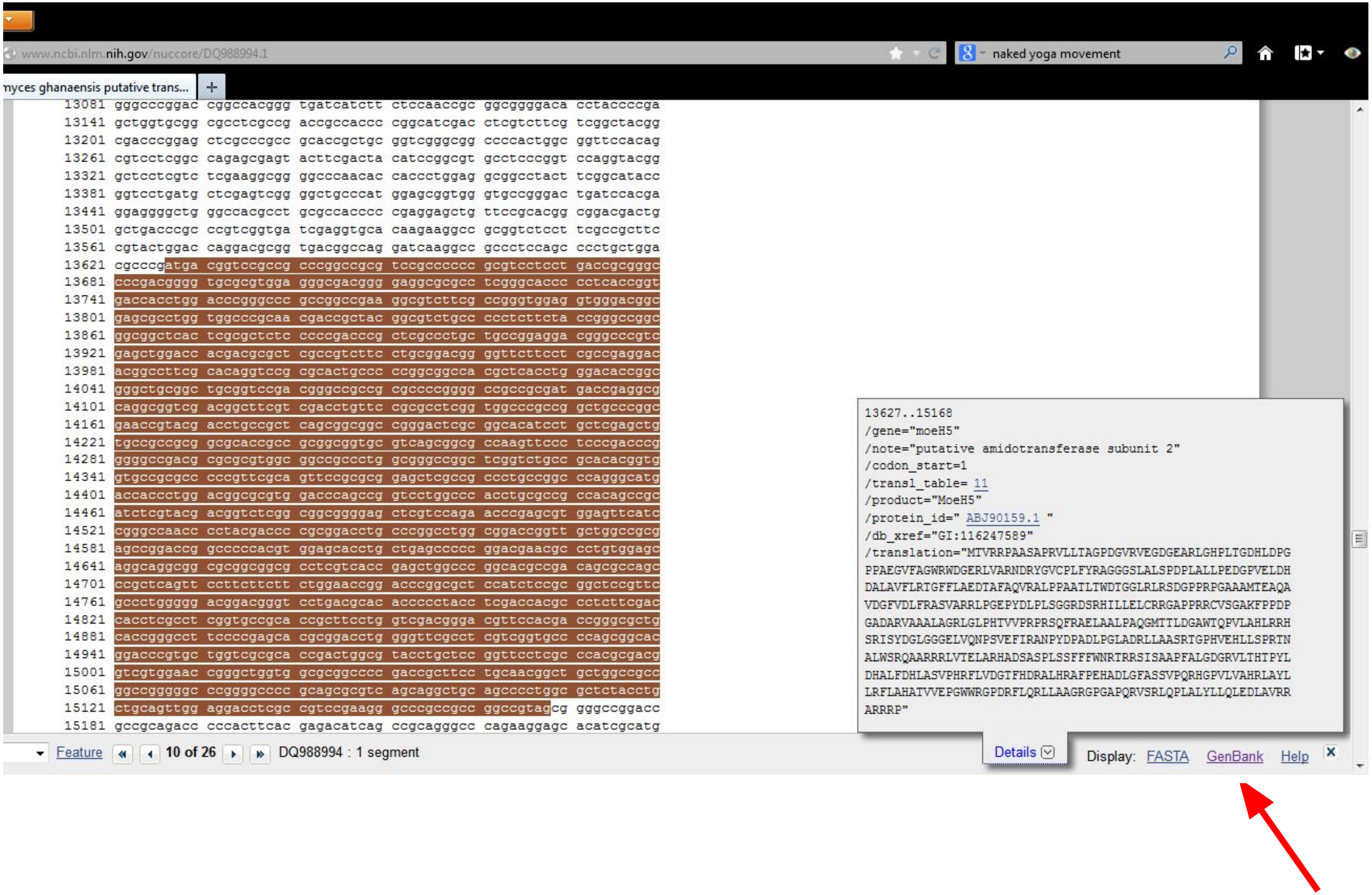
GenBank

FEATURES

```
gene             13627..15168
                 /gene="moeH5"
CDS              13627..15168
                 /gene="moeH5"
                 /note="putative amidotransferase subunit 2"
                 /codon_start=1
                 /transl_table=11
                 /product="MoeH5"
                 /protein_id="ABJ90159.1"
                 /db_xref="GI:116247589"
                 /translation="MTVRRPAASAPRVLLTAGPDGVRVEGDGEARLGHPLTGDHLDPG
                 PPAEGVFAGWRWDGERLVARNDRYGVCPLFYRAGGGSLALSPDPLALLPEDGPVELDH
                 DALAVFLRTGFFLAEDTAFAQVRALPPAATLTWDTGGLRLRSDGPPRPGAAAMTEAQA
                 VDGFVDLFRASVARRLPGEPYDLPLSGGRDSRHILLELCRRGAPPRRCVSGAKFPPDP
                 GADARVAAALAGRLGLPHTVVPRPRSQFRAELAALPAQGMTTLDGAWTQPVLAHLRRH
                 SRISYDGLGGGELVQNPSVEFIRANPYDPADLPGLADRLLAASRTGPHVEHLLSPRTN
                 ALWSRQAARRRLVTELARHADSASPLSSFFFWNRTRRSISAAPFALGDGRVLTHTPYL
                 DHALFDHLASVPHRFLVDGTFHDRALHRAFPEHADLGFASSVPQRHGPVLVAHRLAYL
                 LRFLAHATVVEPGWWRGPDRFLQRLLAAGRGPGAPQRVSRLQPLALYLLQLEDLAVRR
                 ARRRP"
```

GenBank

moeH5



```
13081 gggcccggac cggccacggg tgatcatctt ctccaaccgc ggcggggaca cctaccccga
13141 gctggtgcgg cgcctcgccg accgccaccc cggcatcgac ctcgtcttcg tcggctacgg
13201 cgacccggag ctcgcccgcc gcaccgctgc ggtcgggcgg ccccactggc ggttccacag
13261 cgtcctcggc cagagcgagt acttcgacta catccggcgt gcctcccggt ccaggtacgg
13321 gctcctcgtc tcgaaggcgg ggcccaacac caccctggag gcggcctact tcggcatacc
13381 ggtcctgatg ctcgagtcgg ggctgcccat ggagcggtgg gtgccgggac tgatccacga
13441 ggagggcctg ggccacgcct gcgccacccc cgaggagctg ttccgcacgg cggacgactg
13501 gctgacccgc ccgtcggtga tcgaggtgca caagaaggcc gcggtctcct tcgccgcttc
13561 cgtactggac caggacgcgg tgacggccag gatcaaggcc gccctccagc ccctgctgga
13621 cgcccgatga cggtccgccg cccggccgcg tccgcccccc gcgtcctcct gaccgcgggc
13681 cccgacgggg tgcgcgtgga gggcgacggg gaggcgcgcc tcgggcaccc cctcaccggt
13741 gaccacctgg acccgggccc gccggccgaa ggcgtcttcg ccgggtggag gtgggacggc
13801 gagcgcctgg tggcccgcaa cgaccgctac ggcgtctgcc ccctcttcta ccgggccggc
13861 ggcggctcac tcgcgctctc ccccgacccg ctcgccctgc tgccggagga cgggcccgtc
13921 gagctggacc acgacgcgct cgccgtcttc ctgcggacgg ggttcttcct cgccgaggac
13981 acggccttcg cacaggtccg cgcactgccc ccggcggcca cgctcacctg ggacaccggc
14041 gggctgcggc tgcggtccga cgggccgccg cgccccgggg ccgccgcgat gaccgaggcg
14101 caggcggtcg acggcttcgt cgacctgttc cgcgcctcgg tggcccgccg gctgcccggc
14161 gaaccgtacg acctgccgct cagcggcggc cgggactcgc ggcacatcct gctcgagctg
14221 tgccgccgcg gcgcaccgcc gcggcggtgc gtcagcggcg ccaagttccc tcccgacccg
14281 ggggccgacg cgcgcgtggc ggccgccctg gcgggccggc tcggtctgcc gcacacggtg
14341 gtgccgcgcc cccgttcgca gttccgcgcg gagctcgccg ccctgccggc ccagggcatg
14401 accaccctgg acggcgcgtg gacccagccg gtcctggccc acctgcgccg ccacagccgc
14461 atctcgtacg acggtctcgg cggcggggag ctcgtccaga acccgagcgt ggagttcatc
14521 cgggccaacc cctacgaccc cgcggacctg cccggcctgg cggaccggtt gctggccgcg
14581 agccggaccg gcccccacgt ggagcacctg ctgagccccc ggacgaacgc cctgtggagc
14641 aggcaggcgg cgcggcggcg cctcgtcacc gagctggccc ggcacgccga cagcgccagc
14701 ccgctcagtt ccttcttctt ctggaaccgg acccggcgct ccatctccgc ggctccgttc
14761 gccctggggg acggacgggt cctgacgcac acccccctacc tcgaccacgc cctcttcgac
14821 cacctcgcct cggtgccgca ccgcttcctg gtcgacggga cgttccacga ccgggcgctg
14881 caccgggcct tccccgagca cgcggacctg ggttcgcct cgtcggtgcc ccagcggcac
14941 ggacccgtgc tggtcgcgca ccgactggcg tacctgctcc ggttcctcgc ccacgcgacg
15001 gtcgtggaac cgggctggtg gcgcggcccc gaccgcttcc tgcaacggct gctggccgcc
15061 ggccgggggc ccggggcccc gcagcgcgtc agcaggctgc agcccctggc gctctacctg
15121 ctgcagttgg aggacctcgc cgtccgaagg gcccgccgcc ggccgtagcg gggccggacc
15181 gccgcagacc cccacttcac gagacatcag ccgcagggcc cagaaggagc acatcgcatg
```

```
13627..15168
/gene="moeH5"
/note="putative amidotransferase subunit 2"
/codon_start=1
/transl_table= 11
/product="MoeH5"
/protein_id=" ABJ90159.1 "
/db_xref="GI:116247589"
/translation="MTVRRPAASAPRVLLTAGPDGVRVEGDGEARLGHPLTGDHLDPG
PPAEGVFAGWRWDGERLVARNDRYGVCPLFYRAGGGSLALSPDPLALLPEDGPVELDH
DALAVFLRTGFFLAEDTAFAQVRALPPAATLTWDTGGLRLRSDGPPRPGAAAMTEAQA
VDGFVDLFRASVARRLPGEPYDLPLSGGRDSRHILLELCRRGAPPRRCVSGAKFPPDP
GADARVAAALAGRLGLPHTVVPRPRSQFRAELAALPAQGMTTLDGAWTQPVLAHLRRH
SRISYDGLGGGELVQNPSVEFIRANPYDPADLPGLADRLLAASRTGPHVEHLLSPRTN
ALWSRQAARRRLVTELARHADSASPLSSFFFWNRTRRSISAAPFALGDGRVLTHTPYL
DHALFDHLASVPHRFLVDGTFHDRALHRAFPEHADLGFASSVPQRHGPVLVAHRLAYL
LRFLAHATVVEPGWWRGPDRFLQRLLAAGRGPGAPQRVSRLQPLALYLLQLEDLAVRR
ARRRP"
```

GenBank

GBFF для *moeH5* - гена

Display Settings: GenBank (full)

Showing 1.54kb region from base 13627 to 15168.

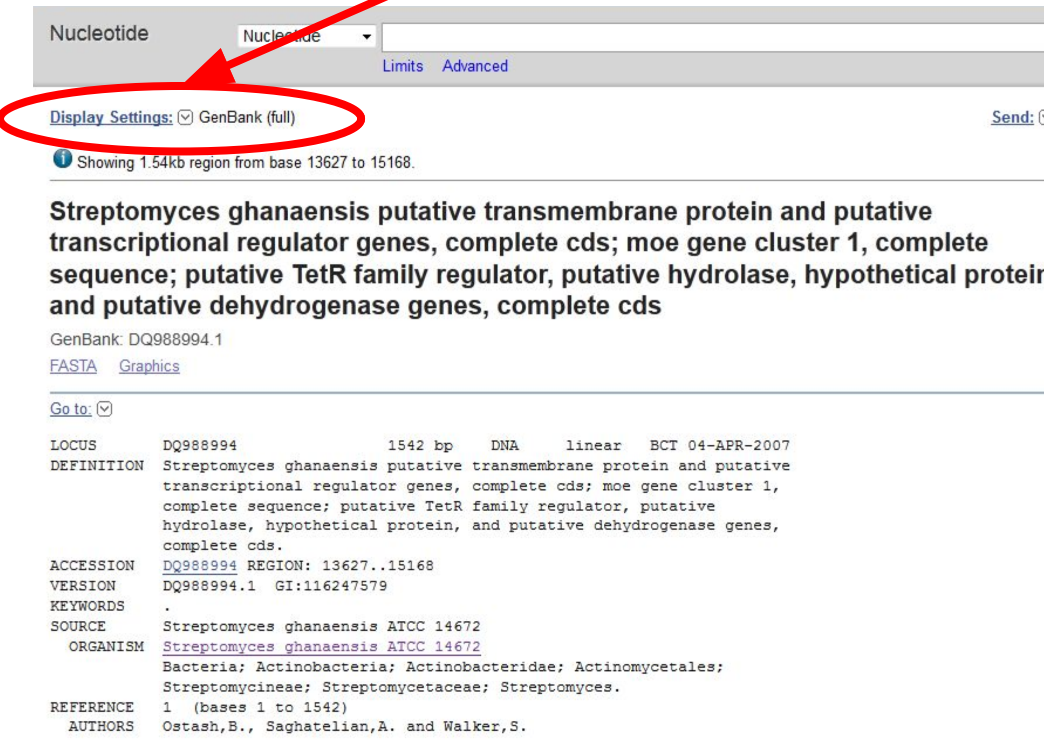
Streptomyces ghanaensis putative transmembrane protein and putative transcriptional regulator genes, complete cds; moe gene cluster 1, complete sequence; putative TetR family regulator, putative hydrolase, hypothetical protein, and putative dehydrogenase genes, complete cds

GenBank: DQ988994.1

FASTA Graphics

Go to:

```
LOCUS       DQ988994                1542 bp    DNA     linear   BCT 04-APR-2007
DEFINITION  Streptomyces ghanaensis putative transmembrane protein and putative
            transcriptional regulator genes, complete cds; moe gene cluster 1,
            complete sequence; putative TetR family regulator, putative
            hydrolase, hypothetical protein, and putative dehydrogenase genes,
            complete cds.
ACCESSION   DQ988994 REGION: 13627..15168
VERSION     DQ988994.1  GI:116247579
KEYWORDS    .
SOURCE      Streptomyces ghanaensis ATCC 14672
  ORGANISM  Streptomyces ghanaensis ATCC 14672
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE   1  (bases 1 to 1542)
  AUTHORS   Ostash,B., Saghatelian,A. and Walker,S.
  TITLE     A streamlined metabolic pathway for the biosynthesis of moenomycin
            A
  JOURNAL   Chem. Biol. 14 (3), 257-267 (2007)
   PUBMED   17379141
REFERENCE   2  (bases 1 to 1542)
  AUTHORS   Ostash,B.O. and Walker,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2006) Microbiology and Molecular Genetics,
            Harvard Medical School, 200 Longwood Ave., Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1..1542
                     /organism="Streptomyces ghanaensis ATCC 14672"
                     /mol_type="genomic DNA"
                     /strain="ATCC14672"
                     /db_xref="ATCC:14672"
                     /db_xref="taxon:566461"
     misc_feature    <1..>1542
                     /note="moe gene cluster 1; involved in moenomycin A
                     biosynthesis"
     gene            <1..4
                     /gene="moeGT1"
     CDS             <1..4
                     /gene="moeGT1"
                     /note="putative glycosyltransferase"
                     /codon_start=1
                     /transl_table=11
                     /product="MoeGT1"
                     /protein_id="ABJ90158.1"
                     /db_xref="GI:116247588"
                     /translation="MAAPDRPLVQVLSPRTWGEFGNYLAATRFSRALRSVIDAEVTLL
                     EAEPILPWIGEAGAQIRTISLESPDAVVRNQRYMALMDRLQARFPEGFEADPTAAQRA
                     DLEPLTRHLRESAPDVVVGTKGFVARLCVAAVRLAGTSTRVVSHVTNPGLLQLPLHRS
                     RYPDLTLVGFPRAKEHLLATAGGDPERVQVVGPLVAQHDLRDFMTSETAVSEAGPWGG
                     DSGPDRPRVIIFSNRGGDTYPELVRRLADRHPGIDLVFVGYGDPELARRTAAVGRPHW
                     RFHSVLGQSEYFDYIRRASRSRYGLLVSKAGPNTTLEAAYFGIPVLMLESGLPMERWV
                     PGLIHEEGLGHACATPEELFRTADDWLTRPSVIEVHKKAAVSFAASVLDQDAVTARIK
                     AALQPLLDAR"
     gene            1..1542
                     /gene="moeH5"
     CDS             1..1542
                     /gene="moeH5"
                     /note="putative amidotransferase subunit 2"
                     /codon_start=1
                     /transl_table=11
                     /product="MoeH5"
                     /protein_id="ABJ90159.1"
                     /db_xref="GI:116247589"
                     /translation="MTVRRPAASAPRVLLTAGPDGVRVEGDGEARLGHPLTGDHLDPG
                     PPAEGVFAGWRWDGERLVARNDRYGVCPLFYRAGGGSLALSPDPLALLPEDGPVELDH
                     DALAVFLRTGFFLAEDTAFAQVRALPPAATLTWDTGGLRLRSDGPPRPGAAAMTEAQA
                     VDGFVDLFRASVARRLPGEPYDLPLSGGRDSRHILLELCRRGAPPRRCVSGAKFPPDP
                     GADARVAAALAGRLGLPHTVVPRPRSQFRAELAALPAQGMTTLDGAWTQPVLAHLRRH
                     SRISYDGLGGGELVQNPSVEFIRANPYDPADLPGLADRLLAASRTGPHVEHLLSPRTN
                     ALWSRQAARRRLVTELARHADSASPLSSFFFWNRTRRSISAAPFALGDGRVLTHTPYL
                     DHALFDHLASVPHRFLVDGTFHDRALHRAFPEHADLGFASSVPQRHGPVLVAHRLAYL
                     LRFLAHATVVEPGWWRGPDRFLQRLLAAGRGPGAPQRVSRLQPLALYLLQLEDLAVRR
                     ARRRP"
ORIGIN
        1 atgacggtcc gccgcccggc cgcgtccgcc ccccgcgtcc tcctgaccgc gggcccccgac
       61 ggggtgcgcg tggagggcga cggggaggcg cgcctcgggc accccctcac cggtgaccac
      121 ctggacccgg gcccgccggc cgaaggcgtc ttcgccgggt ggaggtggga cggcgagcgc
      181 ctggtggccc gcaacgaccg ctacggcgtc tgccccctct tctaccgggc cggcggcggc
      241 tcactcgcgc tctcccccga cccgctcgcc ctgctgccgg aggacgggcc cgtcgagctg
      301 gaccacgacg cgctcgccgt cttcctgcgg acggggttct tcctcgccga ggacacggcc
      361 ttcgcacagg tccgcgcact gcccccggcg gccacgctca cctgggacac cggcgggctg
      421 cggctgcggt ccgacgggcc gccgcgcccc ggggccgccg cgatgaccga ggcgcaggcg
      481 gtcgacggct tcgtcgacct gttccgcgcc tcggtggccc gccggctgcc cggcgaaccg
      541 tacgacctgc cgctcagcgg cggccgggac tcgcggcaca tcctgctcga gctgtgccgc
      601 cgcggcgcac cgccgcggcg gtgcgtcagc ggcgccaagt tccctcccga cccgggggcc
      661 gacgcgcgcg tggcggccgc cctggcgggc cggctcggtc tgccgcacac ggtggtgccg
      721 cgcccccgtt cgcagttccg cgcggagctc gccgccctgc cggcccaggg catgaccacc
      781 ctggacggcg cgtggaccca gccggtcctg gcccacctgc gccgccacag ccgcatctcg
      841 tacgacggtc tcggcggcgg ggagctcgtc cagaacccga gcgtggagtt catccgggcc
      901 aacccctacg accccgcgga cctgcccggc ctggcggacc ggttgctggc cgcgagccgg
      961 accggccccc acgtggagca cctgctgagc cccggacgga acgccctgtg gagcaggcag
      961 accggccccc acgtggagca cctgctgagc cccggacgga acgccctgtg gagcaggcag
     1021 gcggcgcggc ggcgcctcgt caccgagctg gcccggcacg ccgacagcgc cagcccgctc
     1081 agttccttct tcttctggaa ccggacccgg cgctccatct ccgcggctcc gttcgccctg
     1141 ggggacggac gggtcctgac gcacaccccc tacctcgacc acgccctctt cgaccacctc
     1201 gcctcggtgc cgcaccgctt cctggtcgac gggacgttcc acgaccgggc gctgcaccgg
     1261 gccttccccg agcacgcgga cctggggttc gcctcgtcgg tgccccagcg gcacggaccc
     1321 gtgctggtcg cgcaccgact ggcgtacctg ctccggttcc tcgcccacgc gacggtcgtg
     1381 gaaccgggct ggtggcgcgg ccccgaccgc ttcctgcaac ggctgctggc cgccggccgg
     1441 gggcccgggg ccccgcagcg cgtcagcagg ctgcagcccc tggcgctcta cctgctgcag
     1501 ttggaggacc tcgccgtccg aagggcccgc cgccggccgt ag
//
```



Display Settings: GenBank (full)

Showing 1.54kb region from base 13627 to 15168.

Streptomyces ghanaensis putative transmembrane protein and putative transcriptional regulator genes, complete cds; moe gene cluster 1, complete sequence; putative TetR family regulator, putative hydrolase, hypothetical protein and putative dehydrogenase genes, complete cds

GenBank: DQ988994.1

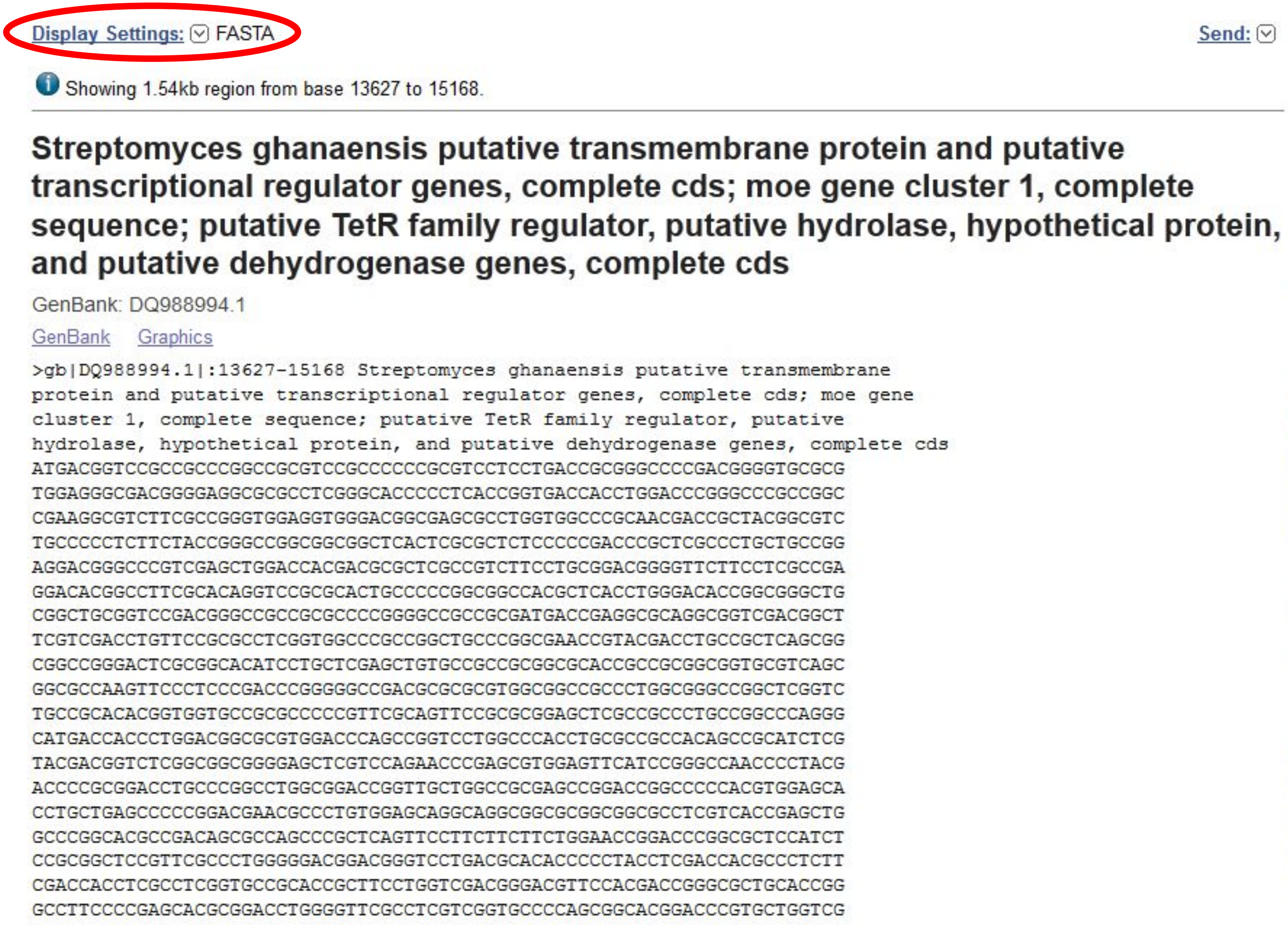
FASTA Graphics

Go to:

```
LOCUS       DQ988994                1542 bp    DNA     linear   BCT 04-APR-2007
DEFINITION  Streptomyces ghanaensis putative transmembrane protein and putative
            transcriptional regulator genes, complete cds; moe gene cluster 1,
            complete sequence; putative TetR family regulator, putative
            hydrolase, hypothetical protein, and putative dehydrogenase genes,
            complete cds.
ACCESSION   DQ988994 REGION: 13627..15168
VERSION     DQ988994.1  GI:116247579
KEYWORDS    .
SOURCE      Streptomyces ghanaensis ATCC 14672
  ORGANISM  Streptomyces ghanaensis ATCC 14672
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE   1  (bases 1 to 1542)
  AUTHORS   Ostash,B., Saghatelian,A. and Walker,S.
```

GenBank

FASTA-файл для *moeH5* - гена



Showing 1.54kb region from base 13627 to 15168.

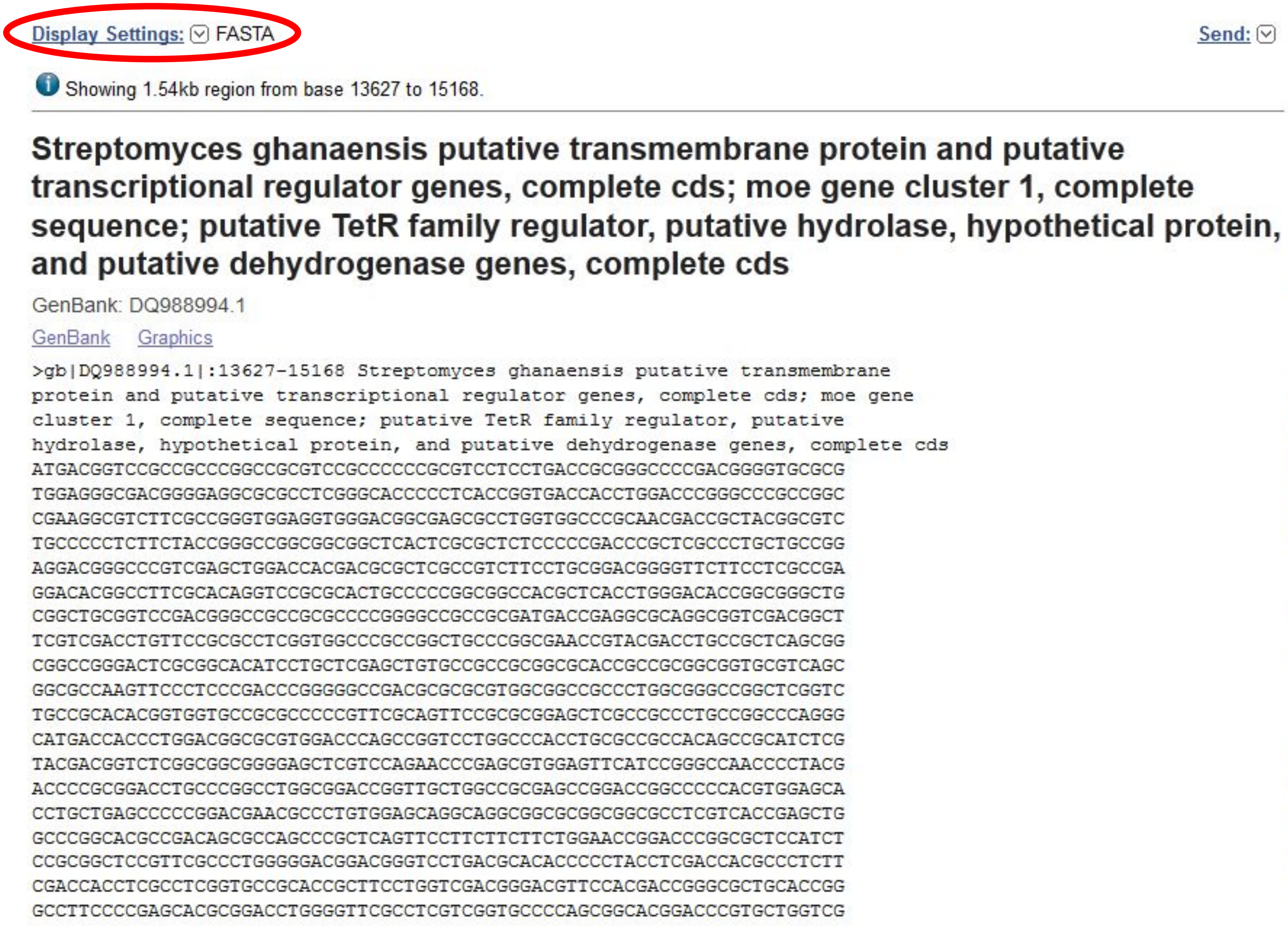
Streptomyces ghanaensis putative transmembrane protein and putative transcriptional regulator genes, complete cds; moe gene cluster 1, complete sequence; putative TetR family regulator, putative hydrolase, hypothetical protein, and putative dehydrogenase genes, complete cds

GenBank: DQ988994.1

```
>gb|DQ988994.1|:13627-15168 Streptomyces ghanaensis putative transmembrane
protein and putative transcriptional regulator genes, complete cds; moe gene
cluster 1, complete sequence; putative TetR family regulator, putative
hydrolase, hypothetical protein, and putative dehydrogenase genes, complete cds
ATGACGGTCCGCCGCCCGGCCGCGTCCGCCCCCCGCGTCCTCCTGACCGCGGGCCCCGACGGGGTGCGCG
TGGAGGGCGACGGGGAGGCGCGCCTCGGGCACCCCCTCACCGGTGACCACCTGGACCCGGGCCCGCCGGC
CGAAGGCGTCTTCGCCGGGTGGAGGTGGGACGGCGAGCGCCTGGTGGCCCGCAACGACCGCTACGGCGTC
TGCCCCCTCTTCTACCGGGCCGGCGGCGGCTCACTCGCGCTCTCCCCCGACCCGCTCGCCCTGCTGCCGG
AGGACGGGCCCGTCGAGCTGGACCACGACGCGCTCGCCGTCTTCCTGCGGACGGGGTTCTTCCTCGCCGA
GGACACGGCCTTCGCACAGGTCCGCGCACTGCCCCCGGCGGCCACGCTCACCTGGGACACCGGCGGGCTG
CGGCTGCGGTCCGACGGGCCGCCGCGCCCCGGGGCCGCCGCGATGACCGAGGCGCAGGCGGTCGACGGCT
TCGTCGACCTGTTCCGCGCCTCGGTGGCCCGCCGGCTGCCCGGCGAACCGTACGACCTGCCGCTCAGCGG
CGGCCGGGACTCGCGGCACATCCTGCTCGAGCTGTGCCGCCGCGGCGCACCGCCGCGGCGGTGCGTCAGC
GGCGCCAAGTTCCCTCCCGACCCGGGGGCCGACGCGCGCGTGGCGGCCGCCCTGGCGGGCCGGCTCGGTC
TGCCGCACACGGTGGTGCCGCGCCCCCGTTCGCAGTTCCGCGCGGAGCTCGCCGCCCTGCCGGCCCAGGG
CATGACCACCCTGGACGGCGCGTGGACCCAGCCGGTCCTGGCCCACCTGCGCCGCCACAGCCGCATCTCG
TACGACGGTCTCGGCGGCGGGGAGCTCGTCCAGAACCCGAGCGTGGAGTTCATCCGGGCCAACCCCTACG
ACCCCGCGGACCTGCCCGGCCTGGCGGACCGGTTGCTGGCCGCGAGCCGGACCGGCCCCCACGTGGAGCA
CCTGCTGAGCCCCCGGACGAACGCCCTGTGGAGCAGGCAGGCGGCGCGGCGGCGCCTCGTCACCGAGCTG
GCCCGGCACGCCGACAGCGCCAGCCCGCTCAGTTCCTTCTTCTTCTGGAACCGGACCCGGCGCTCCATCT
CCGCGGCTCCGTTCGCCCTGGGGGACGGACGGGTCCTGACGCACACCCCCTACCTCGACCACGCCCTCTT
CGACCACCTCGCCTCGGTGCCGCACCGCTTCCTGGTCGACGGGACGTTCCACGACCGGGCGCTGCACCGG
GCCTTCCCCGAGCACGCGGACCTGGGGTTCGCCTCGTCGGTGCCCCAGCGGCACGGACCCGTGCTGGTCG
```

GenBank

FASTA-файл для *moeH5* - гена



Showing 1.54kb region from base 13627 to 15168.

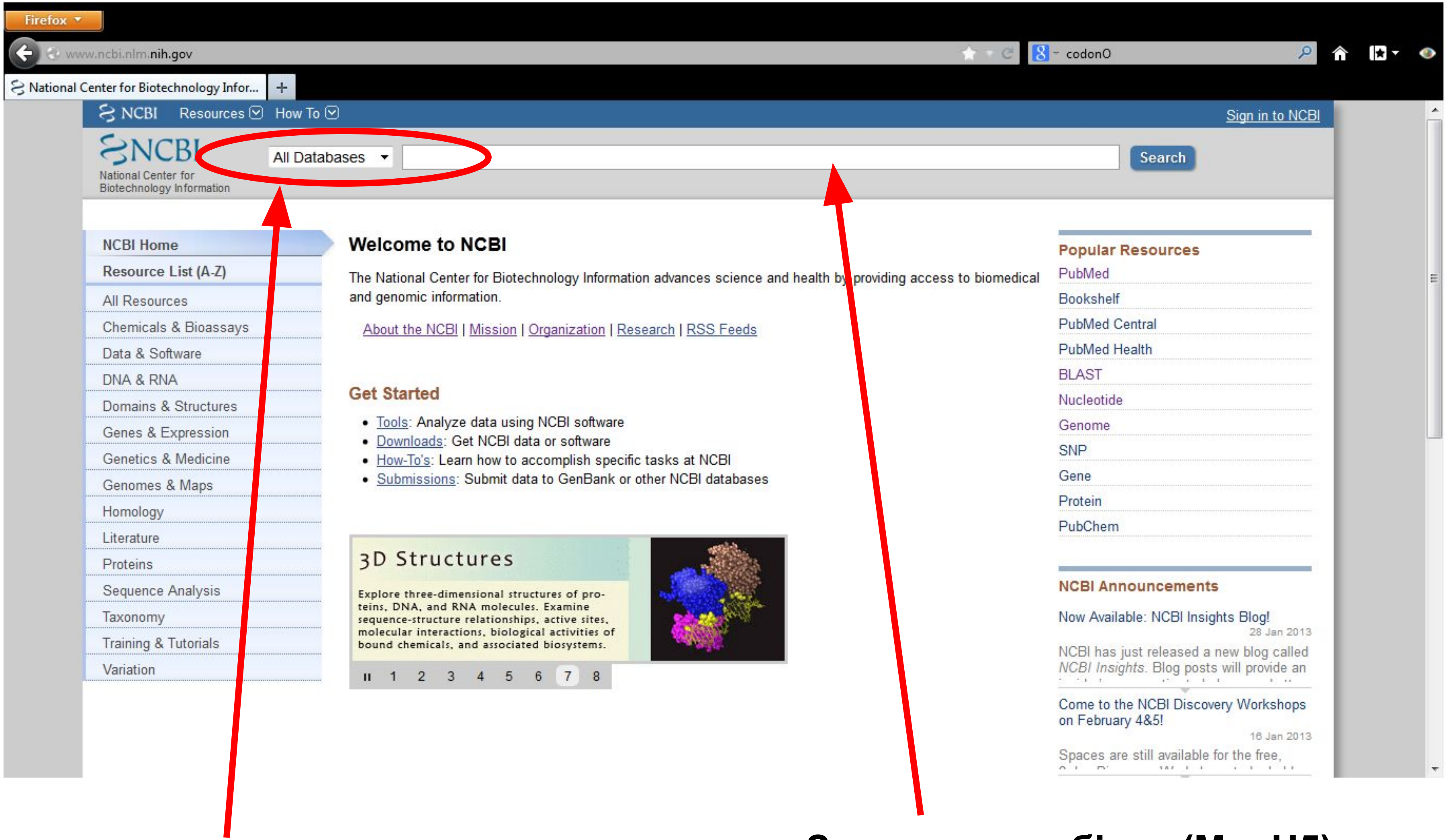
Streptomyces ghanaensis putative transmembrane protein and putative transcriptional regulator genes, complete cds; moe gene cluster 1, complete sequence; putative TetR family regulator, putative hydrolase, hypothetical protein, and putative dehydrogenase genes, complete cds

GenBank: DQ988994.1

```
>gb|DQ988994.1|:13627-15168 Streptomyces ghanaensis putative transmembrane
protein and putative transcriptional regulator genes, complete cds; moe gene
cluster 1, complete sequence; putative TetR family regulator, putative
hydrolase, hypothetical protein, and putative dehydrogenase genes, complete cds
ATGACGGTCCGCCGCCCGGCCGCGTCCGCCCCCCGCGTCCTCCTGACCGCGGGCCCCGACGGGGTGCGCG
TGGAGGGCGACGGGGAGGCGCGCCTCGGGCACCCCCTCACCGGTGACCACCTGGACCCGGGCCCGCCGGC
CGAAGGCGTCTTCGCCGGGTGGAGGTGGGACGGCGAGCGCCTGGTGGCCCGCAACGACCGCTACGGCGTC
TGCCCCCTCTTCTACCGGGCCGGCGGCGGCTCACTCGCGCTCTCCCCCGACCCGCTCGCCCTGCTGCCGG
AGGACGGGCCCGTCGAGCTGGACCACGACGCGCTCGCCGTCTTCCTGCGGACGGGGTTCTTCCTCGCCGA
GGACACGGCCTTCGCACAGGTCCGCGCACTGCCCCCGGCGGCCACGCTCACCTGGGACACCGGCGGGCTG
CGGCTGCGGTCCGACGGGCCGCCGCGCCCCGGGGCCGCCGCGATGACCGAGGCGCAGGCGGTCGACGGCT
TCGTCGACCTGTTCCGCGCCTCGGTGGCCCGCCGGCTGCCCGGCGAACCGTACGACCTGCCGCTCAGCGG
CGGCCGGGACTCGCGGCACATCCTGCTCGAGCTGTGCCGCCGCGGCGCACCGCCGCGGCGGTGCGTCAGC
GGCGCCAAGTTCCCTCCCGACCCGGGGGCCGACGCGCGCGTGGCGGCCGCCCTGGCGGGCCGGCTCGGTC
TGCCGCACACGGTGGTGCCGCGCCCCCGTTCGCAGTTCCGCGCGGAGCTCGCCGCCCTGCCGGCCCAGGG
CATGACCACCCTGGACGGCGCGTGGACCCAGCCGGTCCTGGCCCACCTGCGCCGCCACAGCCGCATCTCG
TACGACGGTCTCGGCGGCGGGGAGCTCGTCCAGAACCCGAGCGTGGAGTTCATCCGGGCCAACCCCTACG
ACCCCGCGGACCTGCCCGGCCTGGCGGACCGGTTGCTGGCCGCGAGCCGGACCGGCCCCCACGTGGAGCA
CCTGCTGAGCCCCCGGACGAACGCCCTGTGGAGCAGGCAGGCGGCGCGGCGGCGCCTCGTCACCGAGCTG
GCCCGGCACGCCGACAGCGCCAGCCCGCTCAGTTCCTTCTTCTTCTGGAACCGGACCCGGCGCTCCATCT
CCGCGGCTCCGTTCGCCCTGGGGGACGGACGGGTCCTGACGCACACCCCCTACCTCGACCACGCCCTCTT
CGACCACCTCGCCTCGGTGCCGCACCGCTTCCTGGTCGACGGGACGTTCCACGACCGGGCGCTGCACCGG
GCCTTCCCCGAGCACGCGGACCTGGGGTTCGCCTCGTCGGTGCCCCAGCGGCACGGACCCGTGCTGGTCG
```

Прямий шлях до MoeH5

Стартова сторінка NCBI



Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

About the NCBI | Mission | Organization | Research | RSS Feeds

Get Started

- Tools: Analyze data using NCBI software
- Downloads: Get NCBI data or software
- How-To's: Learn how to accomplish specific tasks at NCBI
- Submissions: Submit data to GenBank or other NCBI databases

3D Structures

Explore three-dimensional structures of proteins, DNA, and RNA molecules. Examine sequence-structure relationships, active sites, molecular interactions, biological activities of bound chemicals, and associated biosystems.

NCBI Announcements

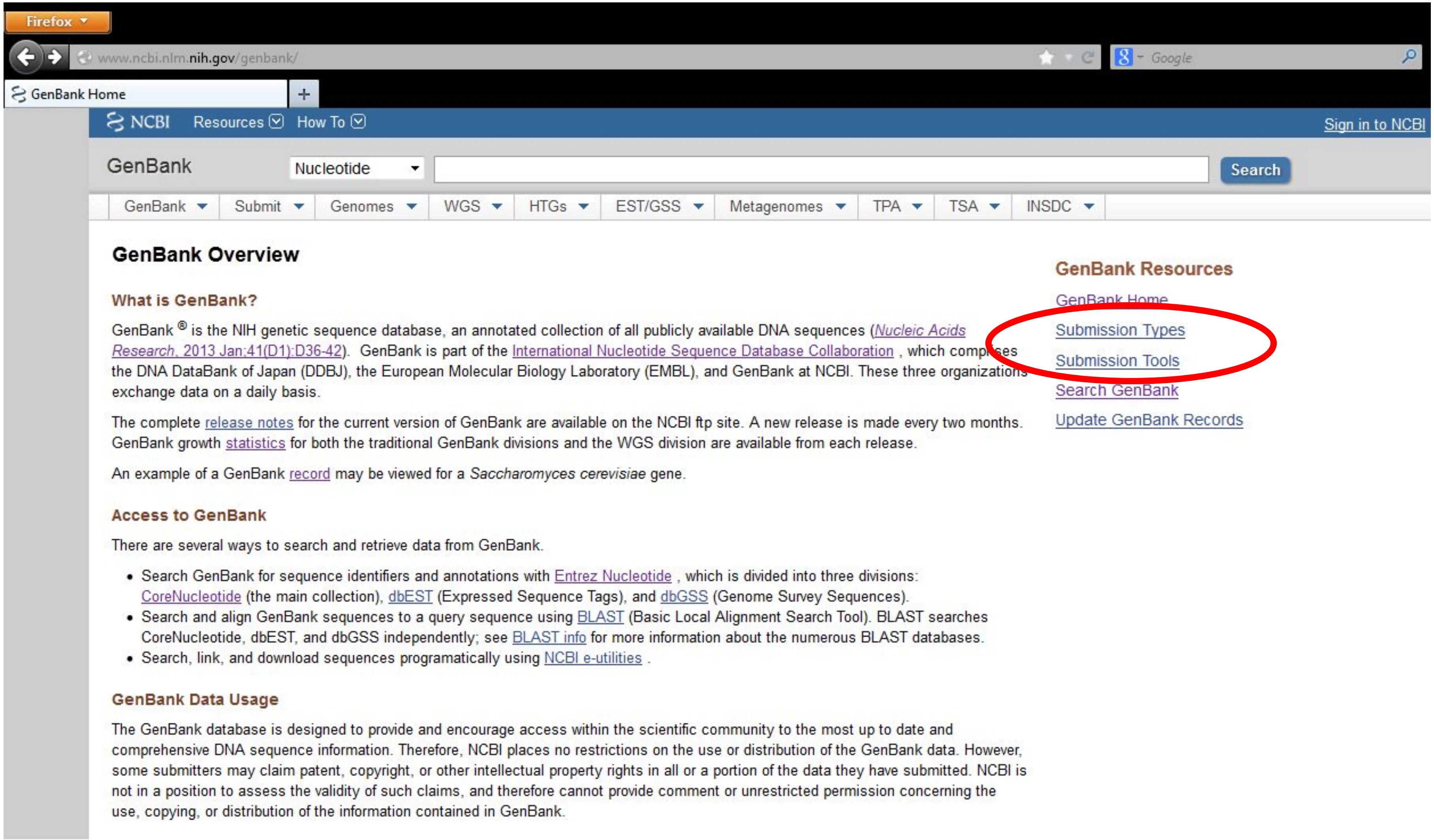
Now Available: NCBI Insights Blog!
28 Jan 2013
NCBI has just released a new blog called *NCBI Insights*. Blog posts will provide an

Come to the NCBI Discovery Workshops on February 4&5!
16 Jan 2013
Spaces are still available for the free,

Вибрати Protein

Задати назву білка (MoeH5)

Депонування НАП



GenBank Overview

What is GenBank?

GenBank ® is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences (*Nucleic Acids Research*, 2013 Jan;41(D1):D36-42). GenBank is part of the International Nucleotide Sequence Database Collaboration , which comprises the DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI. These three organizations exchange data on a daily basis.

The complete release notes for the current version of GenBank are available on the NCBI ftp site. A new release is made every two months. GenBank growth statistics for both the traditional GenBank divisions and the WGS division are available from each release.

An example of a GenBank record may be viewed for a *Saccharomyces cerevisiae* gene.

Access to GenBank

There are several ways to search and retrieve data from GenBank.

- Search GenBank for sequence identifiers and annotations with Entrez Nucleotide , which is divided into three divisions: CoreNucleotide (the main collection), dbEST (Expressed Sequence Tags), and dbGSS (Genome Survey Sequences).
- Search and align GenBank sequences to a query sequence using BLAST (Basic Local Alignment Search Tool). BLAST searches CoreNucleotide, dbEST, and dbGSS independently; see BLAST info for more information about the numerous BLAST databases.
- Search, link, and download sequences programatically using NCBI e-utilities .

GenBank Data Usage

The GenBank database is designed to provide and encourage access within the scientific community to the most up to date and comprehensive DNA sequence information. Therefore, NCBI places no restrictions on the use or distribution of the GenBank data. However, some submitters may claim patent, copyright, or other intellectual property rights in all or a portion of the data they have submitted. NCBI is not in a position to assess the validity of such claims, and therefore cannot provide comment or unrestricted permission concerning the use, copying, or distribution of the information contained in GenBank.

GenBank Resources

- GenBank Home
- Submission Types
- Submission Tools
- Search GenBank
- Update GenBank Records

Депонування НАП

- Нуклеотидні послідовності – у базу, що входить INSDC - GenBank, ENA, DDBJ
- Білкові – SWISSPROT
- GenBank має різні знаряддя для депонування:
  - **BankIt** – для подання відносно коротких і простих (безінтронних) послідовностей
  - **SequIn** – подання складних послідовностей
  - Та інші (тут не розглядаються)
- Подання геномів – окрема історія (див. далі)
- При поданні нуклеотидних послідовностей потрібно знати мінімум інформації про неї, зокрема локалізацію в ній (координати) orf
- У результаті подання послідовності присвоюється номер доступу в GenBank

Депонування НАП

Мінімум інформації про послідовність

- **Розмір (у п.н.)**
- **Локалізація orf (якщо є)**
- **Походження послідовності (що секвенували – хромосома, плазміда тощо; з якого організму)**

1

167

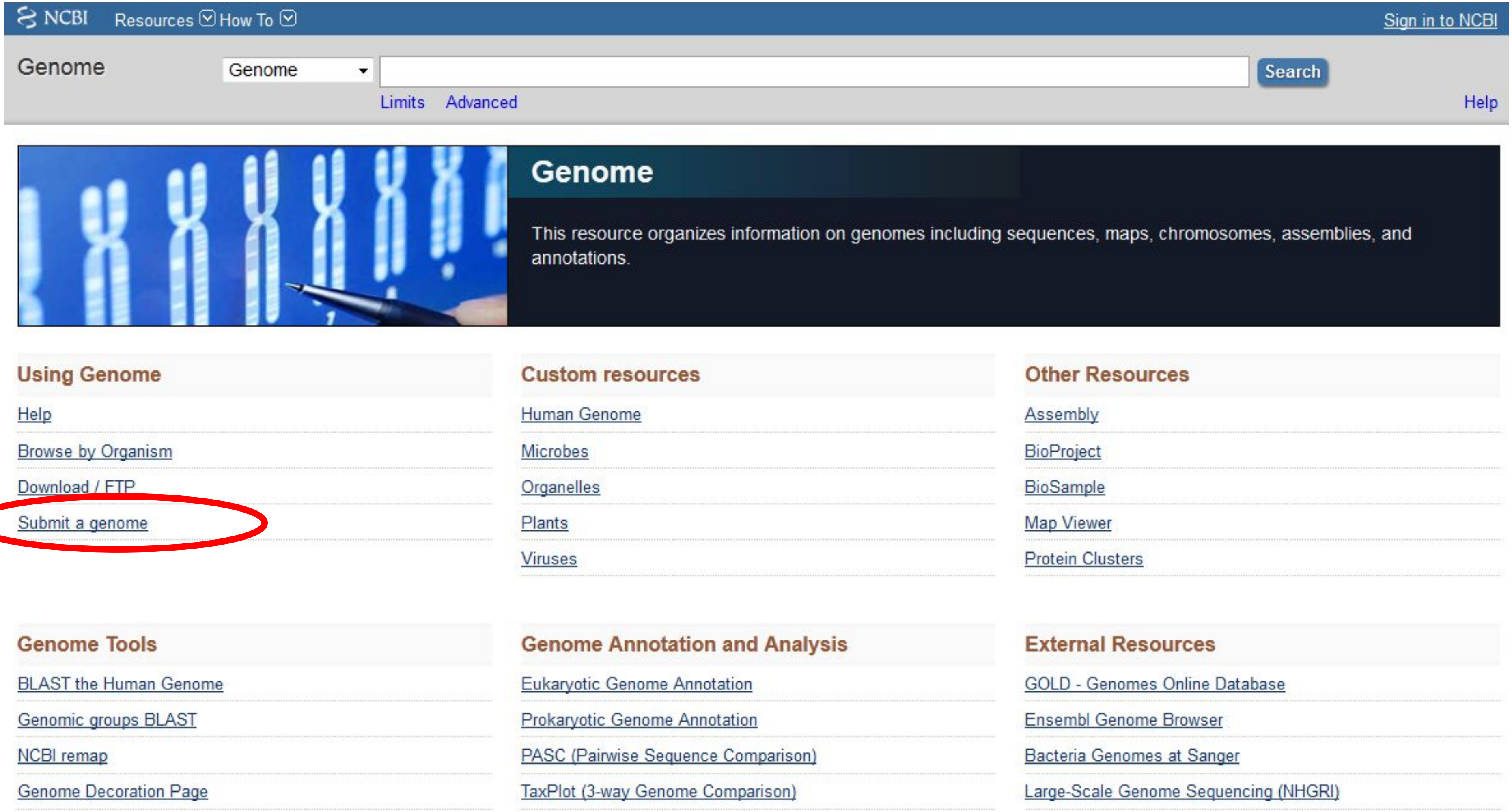
```
CTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCGCAGGAAGCGAGAGCCCCCGGGCCTCCTATCCTGGGCGGGCGCGAA
CGGCTTCGGCCCGACGCGAACGACCGTAGGGGCAAAGGCCATCGGGGGATCGGTTCAGCATGGGCAAGGGCAAGGACAGCCAGGGGTACCCGGCGATGGAGGGCATA
CGGCCGCTCGTGACCGGTGACCCGGCGCGGATCGGGCCCTATCGGCTACTGGGACGGCTCGGTGCGGGCGGGATGGGCCGGGTGTACCTGGCCCGTTCGGCGGGCGG
GCGCACCGTCGCCGTGAAGGTGGTGCACGAGGAGCACATCGCGAACGGGGAGTTCCGGGCCCGCTTCCGTCGGGAGATCGAGGCCGCCCGCCGCGTCGGCGGGCGGT
ACACCGCTCCCGTGCTCGACGCCGACGCCGACGCCGAACGGCCCTGGGTGGCCACCGGCTACGTCCCCGGCCCCTCCCTCGAACAGGCCGTACGGGAGCACGGGCCG
CTGCCCGCCGCCTCGGTCAACGCCCTGGCCGAGGGGCTGTTGAGGGCCCTGCGCGGCATTCACGCCGCCGGGATCGTCCACCGGGACCTCAAGCCGTCCAACGTCCT
GCTCACCGTCGACGGCCCCCGCGTCATCGACTTCGGCATCGCCCGCGCGCTCCAGGTCTCCGTGGAGTCCCTGCTGACCAGTACGGGCATGGTCATCGGCTCTCCCG
GGTTCATGGCGCCCGAGCAGATCCTCGGCGAGGAGACCGGAGCGGGGGCGGACGTCTTCTCCCTCGGCTGTGTCCTGATGTACGCGGCCACGGGGCGGCTGCCGTTC
GGCGCCGGTGCGAGCAACCAGCACGCGGTGATGTTCCGGATCGTCCAGTCCCCGCCGGACCTCGACGCGGTGGAGGACGCCTCCCTGCGCGAGCTGATCGAACGCTG
CCTGACCAAAGCCGCCGCCGAACGCCCGGGAGTGGACGAGCTGGTGGGGTGCCTGGCGCCTGACCGGTCCTCGGACGCCTTGCGTGGTGCCTGGCTGCCGCCGGTGC
TCCTGGCCCGGCTGGCGCAGCAGTCGGCGCTGCTGCTCGACGCGGATGTTCCGGAGACCTCGGACGTACCGGCGCCTGTGCGGGCGGGGGGCGCCGAGGAGCCTGCT
TCTTCGGTTGTTCCTGAGGCTCCCGTTGTTCCTGAGGCTCCCGTGCCCGAGCCGCGTGACCTCGGTACGGTCGATCTGAGGCCGGCACCGGCGAAGGACGACGAGGC
GGGCAGCCTGTCACCGGACCCGGCCCCGGACCCGGACCCGGACTCTTCCCCGGCATCGGCATCGGCATCGGGCCCGGCCCCGGCCTTGTCCCCTGGTCCCGGAGGGG
GCTGGTGCCCGCCGGTGGCGGGCGGACCGGCCAGGGAAGTGGGCGTAGGCGCTCGTGCGTGGGTCGTCGCGGCAGTCGTCGTCGCCGTGCTGGCTGCCGGGGGCACC
ACCGCCTTCCTCAACCGCGGCCCGGGTGGCGCCGACCCCGGGGGCGGAGAGGCTGCCGCGCCGCCCGCCGCCAGTGACGGACCGTCCGGCGCCCCGTCCTCCCCGGC
CGGGGAGGACGACTCCAAGGGCAGCGAGAAGGACAAGGGCAAGGACAAGGACAACGGCAAGGAAGACGGGAAGGACAAGGAAGGTACAGAAGAGGAGCAGGGCGGAG
AGGACGGCGGCGACTCCGGCGGCGGGCAGAGCGACGGGGCCGGTTCCGGCGAGGGCGCGTCCTCGGGCGGAGGCGAGGGGGCCGGAGCTTCCGGCGGAGGCGGCACA
CAGGACGGGTCCGGCTCGTCCGGTTCGGGGGGCTCCGGCGGTGCCGCGAAGCCGCCCGCCCAGGACCCCGCCCCGGACGGCCGGGTGCCGCAGCAGTTCGTCGGGAC
CTGGTCCATCGCGTCCCAGTACGACGCCCTCCAGCCGCACACCGTGGTCATCCGGCGGGTGTCGCCCGGGCAGTCGGCGGTGACCCTCATCGCCGACGTCCAGGGGT
CGGGGCACTGCGAGTACACGGCGAAGCTCAGCTCCGTGGCGGACGGCGGGAACCGGATCAACGTCGGTACCGCCGTGGTGGACAAGGCCCGTTCCGGCGGGGTGTGC
CGCGACACCGACCCCTCGTTCTTCACCGTCGCCGGCTCCGGCATCCTGCACGACGTCGGGCCCGCCCACGGCAGCGGCTATCGCTACAACCGCGCCTGAGCGGGCCC
GTCACTCCGGAATCGACGGTCCGGAACCGGCCGACGTCCGACGCCCG
```

2239

2294

Genome

www.ncbi.nlm.nih.gov/genome



Genome

This resource organizes information on genomes including sequences, maps, chromosomes, assemblies, and annotations.

Using Genome

- Help
- Browse by Organism
- Download / FTP
- Submit a genome

Custom resources

- Human Genome
- Microbes
- Organelles
- Plants
- Viruses

Other Resources

- Assembly
- BioProject
- BioSample
- Map Viewer
- Protein Clusters

Genome Tools

- BLAST the Human Genome
- Genomic groups BLAST
- NCBI remap
- Genome Decoration Page

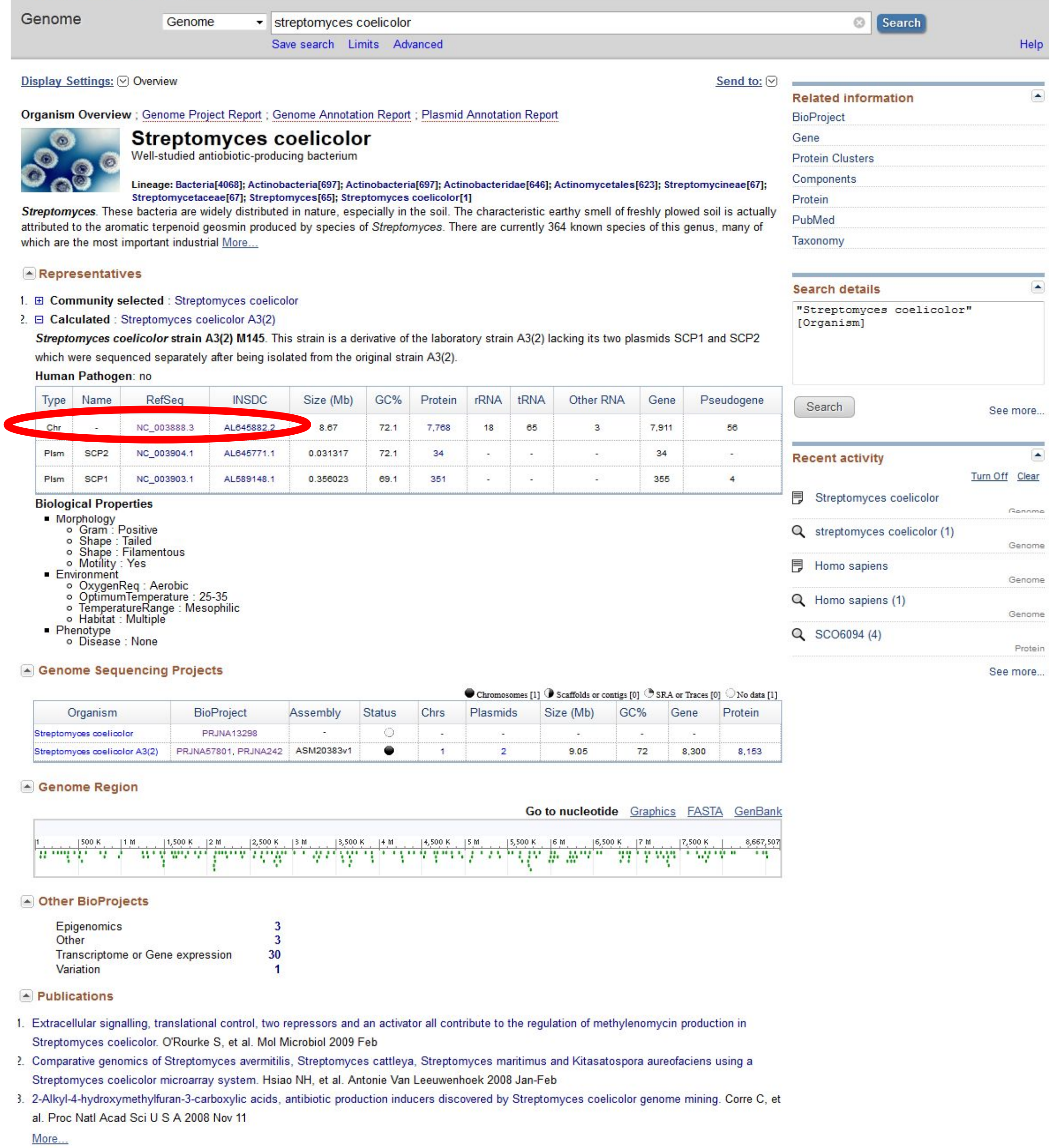
Genome Annotation and Analysis

- Eukaryotic Genome Annotation
- Prokaryotic Genome Annotation
- PASC (Pairwise Sequence Comparison)
- TaxPlot (3-way Genome Comparison)

External Resources

- GOLD - Genomes Online Database
- Ensembl Genome Browser
- Bacteria Genomes at Sanger
- Large-Scale Genome Sequencing (NHGRI)

Genome



streptomyces coelicolor

Display Settings: Overview

Organism Overview ; Genome Project Report ; Genome Annotation Report ; Plasmid Annotation Report

Streptomyces coelicolor

Well-studied antibiotic-producing bacterium

Lineage: Bacteria[4068]; Actinobacteria[697]; Actinobacteria[697]; Actinobacteridae[646]; Actinomycetales[623]; Streptomycineae[67]; Streptomycetaceae[67]; Streptomyces[65]; Streptomyces coelicolor[1]

Streptomyces. These bacteria are widely distributed in nature, especially in the soil. The characteristic earthy smell of freshly plowed soil is actually attributed to the aromatic terpenoid geosmin produced by species of *Streptomyces*. There are currently 364 known species of this genus, many of which are the most important industrial More...

Representatives

1. ⊞ **Community selected** : Streptomyces coelicolor
2. ⊟ **Calculated** : Streptomyces coelicolor A3(2)

***Streptomyces coelicolor* strain A3(2) M145**. This strain is a derivative of the laboratory strain A3(2) lacking its two plasmids SCP1 and SCP2 which were sequenced separately after being isolated from the original strain A3(2).

Human Pathogen: no

Type	Name	RefSeq	INSDC	Size (Mb)	GC%	Protein	rRNA	tRNA	Other RNA	Gene	Pseudogene
Chr	-	NC_003888.3	AL645882.2	8.67	72.1	7,768	18	65	3	7,911	56
Plsm	SCP2	NC_003904.1	AL645771.1	0.031317	72.1	34	-	-	-	34	-
Plsm	SCP1	NC_003903.1	AL589148.1	0.356023	69.1	351	-	-	-	355	4

Biological Properties

- Morphology
  - Gram : Positive
  - Shape : Tailed
  - Shape : Filamentous
  - Motility : Yes
- Environment
  - OxygenReq : Aerobic
  - OptimumTemperature : 25-35
  - TemperatureRange : Mesophilic
  - Habitat : Multiple
- Phenotype
  - Disease : None

Genome Sequencing Projects

● Chromosomes [1] ◑ Scaffolds or contigs [0] ◔ SRA or Traces [0] ○ No data [1]

Organism	BioProject	Assembly	Status	Chrs	Plasmids	Size (Mb)	GC%	Gene	Protein
Streptomyces coelicolor	PRJNA13298	-	○	-	-	-	-	-	
Streptomyces coelicolor A3(2)	PRJNA57801, PRJNA242	ASM20383v1	●	1	2	9.05	72	8,300	8,153

Genome Region

Go to nucleotide Graphics FASTA GenBank

Other BioProjects

Epigenomics	3
Other	3
Transcriptome or Gene expression	30
Variation	1

Publications

1. Extracellular signalling, translational control, two repressors and an activator all contribute to the regulation of methylenomycin production in Streptomyces coelicolor. O'Rourke S, et al. Mol Microbiol 2009 Feb
2. Comparative genomics of Streptomyces avermitilis, Streptomyces cattleya, Streptomyces maritimus and Kitasatospora aureofaciens using a Streptomyces coelicolor microarray system. Hsiao NH, et al. Antonie Van Leeuwenhoek 2008 Jan-Feb
3. 2-Alkyl-4-hydroxymethylfuran-3-carboxylic acids, antibiotic production inducers discovered by Streptomyces coelicolor genome mining. Corre C, et al. Proc Natl Acad Sci U S A 2008 Nov 11

More...

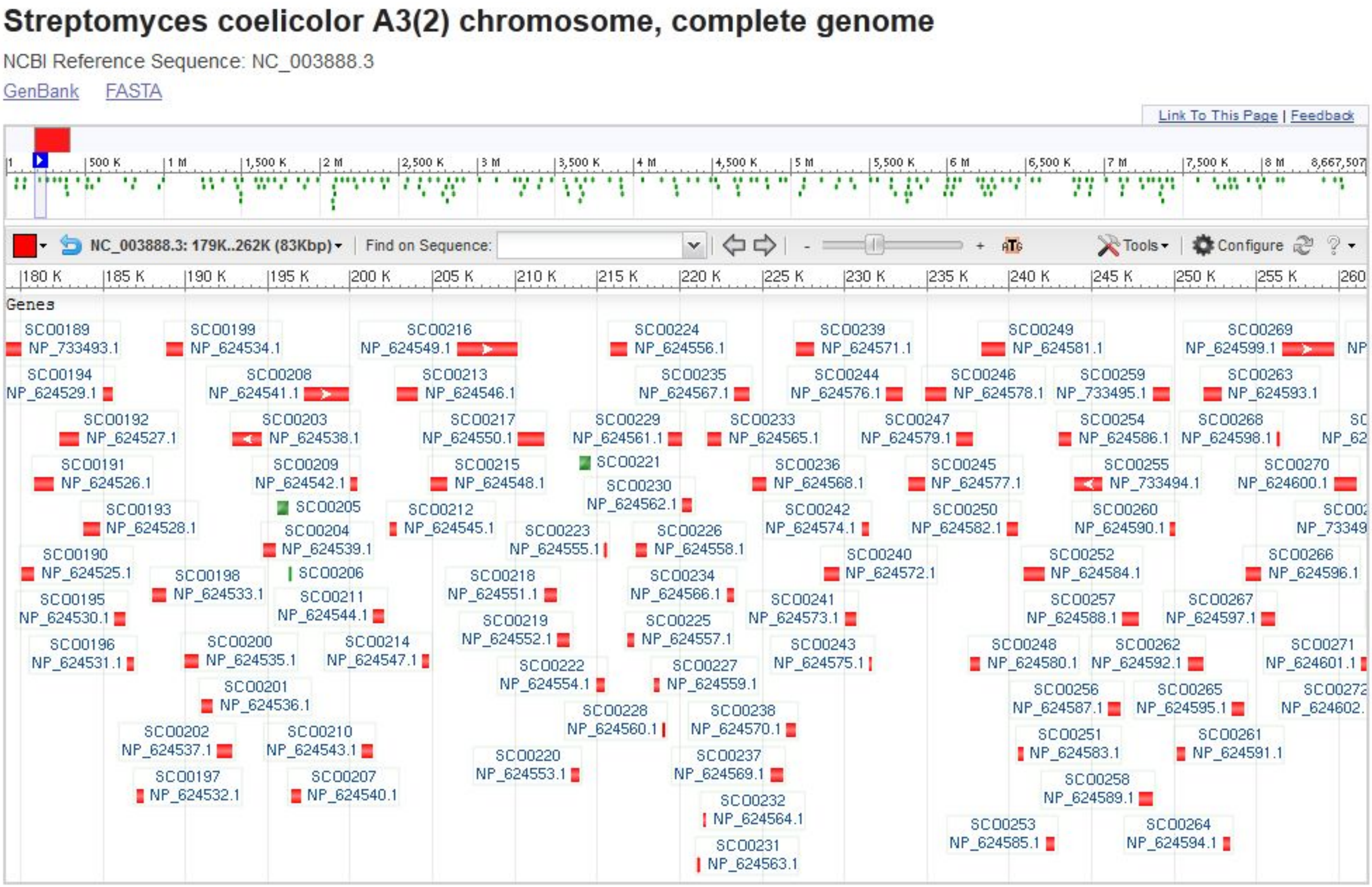
Search details

"Streptomyces coelicolor"[Organism]

Recent activity

- Streptomyces coelicolor — Genome
- streptomyces coelicolor (1) — Genome
- Homo sapiens — Genome
- Homo sapiens (1) — Genome
- SCO6094 (4) — Protein

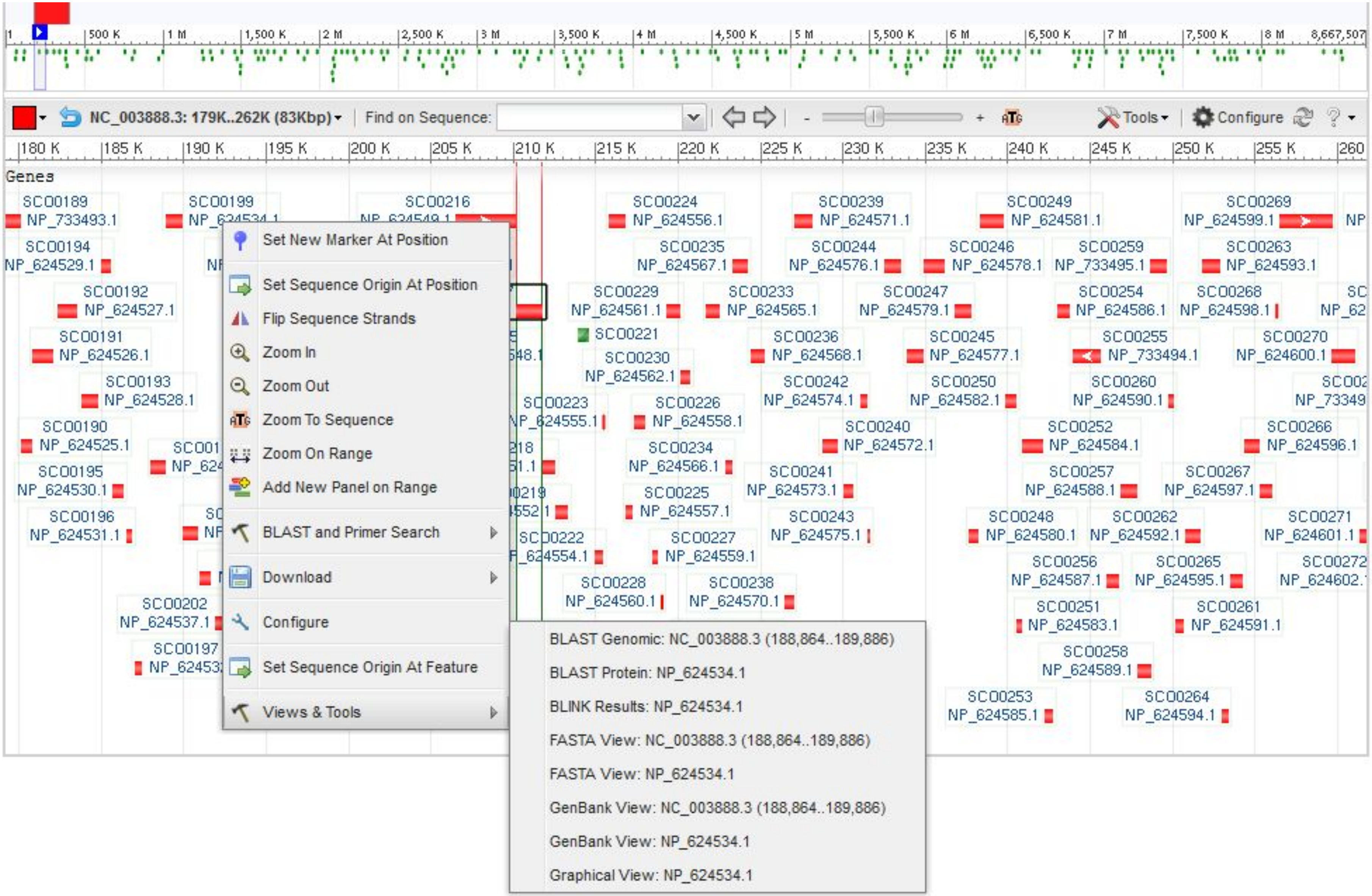
Genome



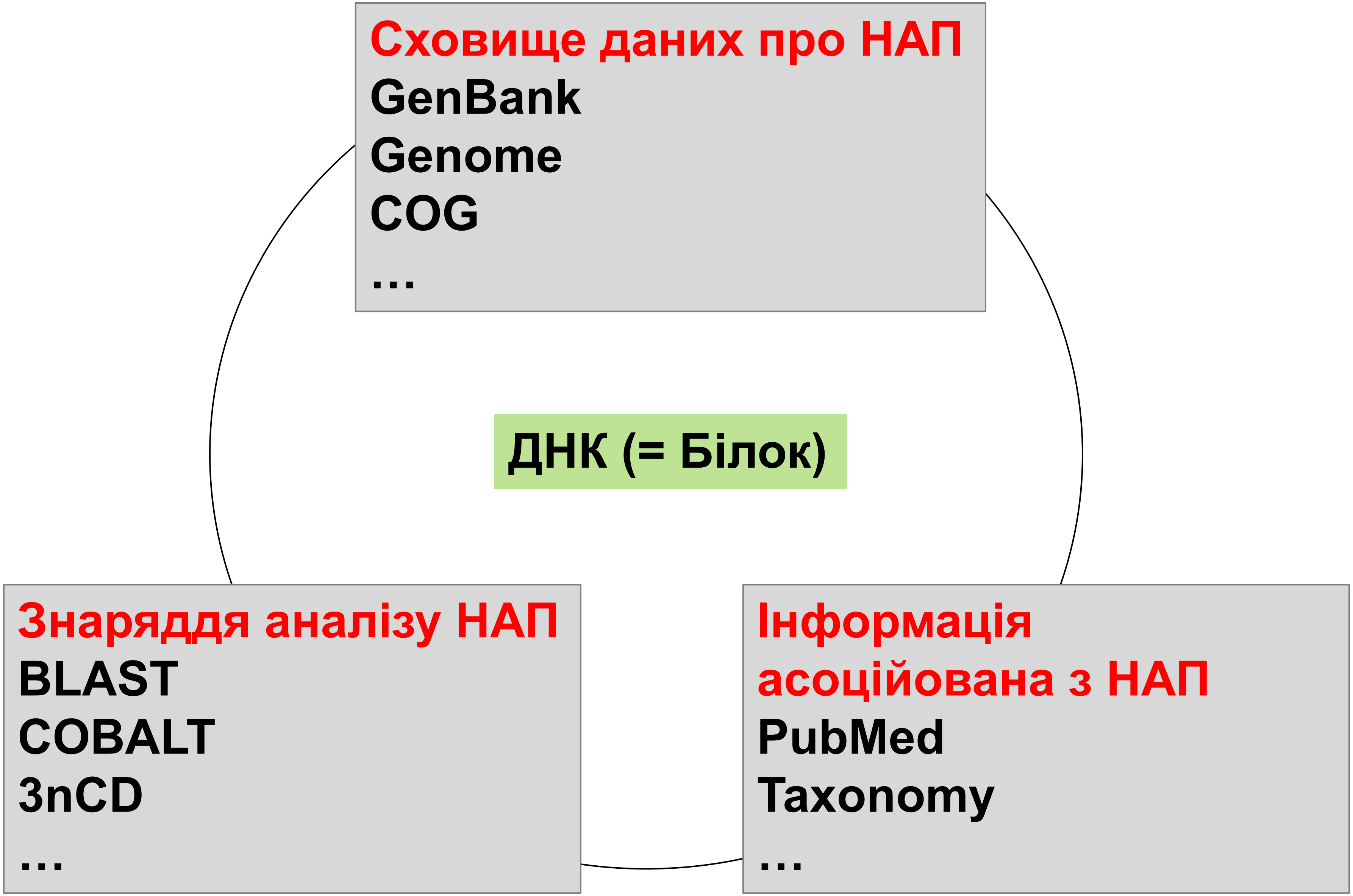
Streptomyces coelicolor A3(2) chromosome, complete genome

NCBI Reference Sequence: NC_003888.3

Genome


1
500 K
1 M
1,500 K
2 M
2,500 K
3 M
3,500 K
4 M
4,500 K
5 M
5,500 K
6 M
6,500 K
7 M
7,500 K
8 M
8,667,507
NC_003888.3: 179K..262K (83Kbp)
Find on Sequence:
Tools
Configure
180 K
185 K
190 K
195 K
200 K
205 K
210 K
215 K
220 K
225 K
230 K
235 K
240 K
245 K
250 K
255 K
260
Genes
SCO0189 NP_733493.1
SCO0199
SCO0216
SCO0224 NP_624556.1
SCO0239 NP_624571.1
SCO0249 NP_624581.1
SCO0269 NP_624599.1
SCO0194 NP_624529.1
SCO0235 NP_624567.1
SCO0244 NP_624576.1
SCO0246 NP_624578.1
SCO0259 NP_733495.1
SCO0263 NP_624593.1
SCO0192 NP_624527.1
SCO0229 NP_624561.1
SCO0233 NP_624565.1
SCO0247 NP_624579.1
SCO0254 NP_624586.1
SCO0268 NP_624598.1
SCO0191 NP_624526.1
SCO0221
SCO0236 NP_624568.1
SCO0245 NP_624577.1
SCO0255 NP_733494.1
SCO0270 NP_624600.1
SCO0230 NP_624562.1
SCO0193 NP_624528.1
SCO0242 NP_624574.1
SCO0250 NP_624582.1
SCO0260 NP_624590.1
SCO0223
SCO0226 NP_624558.1
SCO0190 NP_624525.1
SCO0240 NP_624572.1
SCO0252 NP_624584.1
SCO0266 NP_624596.1
SCO0234 NP_624566.1
SCO0195 NP_624530.1
SCO0241 NP_624573.1
SCO0257 NP_624588.1
SCO0267 NP_624597.1
SCO0225 NP_624557.1
SCO0196 NP_624531.1
SCO0243 NP_624575.1
SCO0248 NP_624580.1
SCO0262 NP_624592.1
SCO0271 NP_624601.1
SCO0222
SCO0227 NP_624559.1
SCO0256 NP_624587.1
SCO0265 NP_624595.1
SCO0228 NP_624560.1
SCO0238 NP_624570.1
SCO0202 NP_624537.1
SCO0251 NP_624583.1
SCO0261 NP_624591.1
SCO0197
SCO0258 NP_624589.1
SCO0253 NP_624585.1
SCO0264 NP_624594.1
Set New Marker At Position
Set Sequence Origin At Position
Flip Sequence Strands
Zoom In
Zoom Out
Zoom To Sequence
Zoom On Range
Add New Panel on Range
BLAST and Primer Search
Download
Configure
Set Sequence Origin At Feature
Views & Tools
BLAST Genomic: NC_003888.3 (188,864..189,886)
BLAST Protein: NP_624534.1
BLINK Results: NP_624534.1
FASTA View: NC_003888.3 (188,864..189,886)
FASTA View: NP_624534.1
GenBank View: NC_003888.3 (188,864..189,886)
GenBank View: NP_624534.1
Graphical View: NP_624534.1

Модель NCBI



Сховище даних про НАП

GenBank
Genome
COG
…

ДНК (= Білок)

Знаряддя аналізу НАП

BLAST
COBALT
3nCD
…

Інформація асоційована з НАП

PubMed
Taxonomy
…