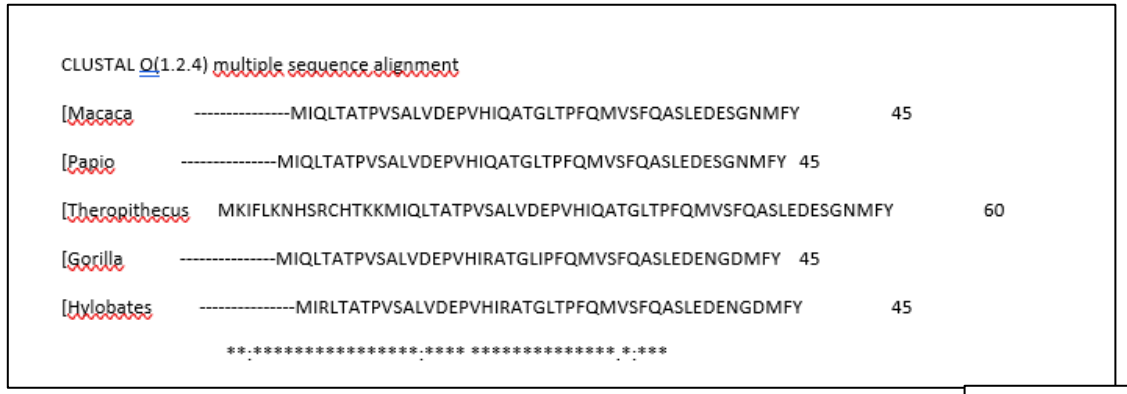
Základy praktické Bioinformatiky

PETRA MATOUŠKOVÁ

2025/2026

font:Courier new



```
CLUSTAL O(1.2.4) multiple sequence alignment

[Macaca          ----------MIQLTATPVSALVDEPVHIQATGLTPFQMVSFQASLEDESGNMFY          45
[Papio           ----------MIQLTATPVSALVDEPVHIQATGLTPFQMVSFQASLEDESGNMFY  45
[Theropithecus   MKIFLKNHSRCHTKKMIQLTATPVSALVDEPVHIQATGLTPFQMVSFQASLEDESGNMFY          60
[Gorilla         ----------MIQLTATPVSALVDEPVHIRATGLIPFQMVSFQASLEDENGDMFY   45
[Hylobates       ----------MIRLTATPVSALVDEPVHIRATGLTPFQMVSFQASLEDENGDMFY          45
                           **.****************.**** **************.*.***
```

Calibri → Courier new

```
CLUSTAL O(1.2.4) multiple sequence alignment

[Macaca          ----------MIQLTATPVSALVDEPVHIQATGLTPFQMVSFQASLEDESGNMFY 45
[Papio           ----------MIQLTATPVSALVDEPVHIQATGLTPFQMVSFQASLEDESGNMFY 45
[Theropithecus   MKIFLKNHSRCHTKKMIQLTATPVSALVDEPVHIQATGLTPFQMVSFQASLEDESGNMFY 60
[Gorilla         ----------MIQLTATPVSALVDEPVHIRATGLIPFQMVSFQASLEDENGDMFY 45
[Hylobates       ----------MIRLTATPVSALVDEPVHIRATGLTPFQMVSFQASLEDENGDMFY 45
                           **:****************:**** **************.*:***
```

Základy praktické bioinformatiky

Téma 5/10

Nukleotidová bioinformatika I

Cíle:

Student bude schopen vyhledat a stáhnout nukleotidové sekvence vybraných genů a získat relevantní informace o těchto genech. Student bude schopen analyzovat základní vlastnosti vybraných nukleotidových sekvencí a porovnat vícero nukleotidových sekvencí.

„Bioinformatika nukleových kyselin“

Vyhledávání NK sekvencí

Analýza vlastností sekvencí-složení, reverse complement, identifikace restrikčních míst (Palindromy)

Práce s kódující DNA=práce s proteiny / překlad DNA sekvence-otvírání čtecího rámce

Návrh primerů pro PCR, rt-PCR

Předpověď sekundárních struktur

Porovnávání sekvencí

Vyhledání SNPs

„čtení“ sekvenačních dat a spojování fragmentů

Vyhledávání hladin expresí jednotlivých genů

mikroRNA

Celé genomy

….

„Bioinformatika nukleových kyselin“

Vyhledávání NK sekvencí

Analýza vlastností sekvencí-složení, reverse complement, identifikace restrikčních míst (Palindromy)

Práce s kódující DNA=práce s proteiny / překlad DNA sekvence-otvírání čtecího rámce

Návrh primerů pro PCR, rt-PCR

Předpověď sekundárních struktur

Porovnávání sekvencí

Vyhledání SNPs

„čtení“ sekvenačních dat a spojování fragmentů

Vyhledávání hladin expresí jednotlivých genů

mikroRNA

Celé genomy

….

DNA/RNA

4 nukleotidy: Adenin, Guanin, Thymidin (Uracil), Cytosin : **5´→ 3´**

(Frederick Sanger/Walter Gilbert 1980 Nobelova cena za určování DNA sekvence)

Genetický kód založený na tripletech:

Zápis sekvence:

A, C, T (U), G

N cokoliv

R A/G purin

Y C/T pyrimidin

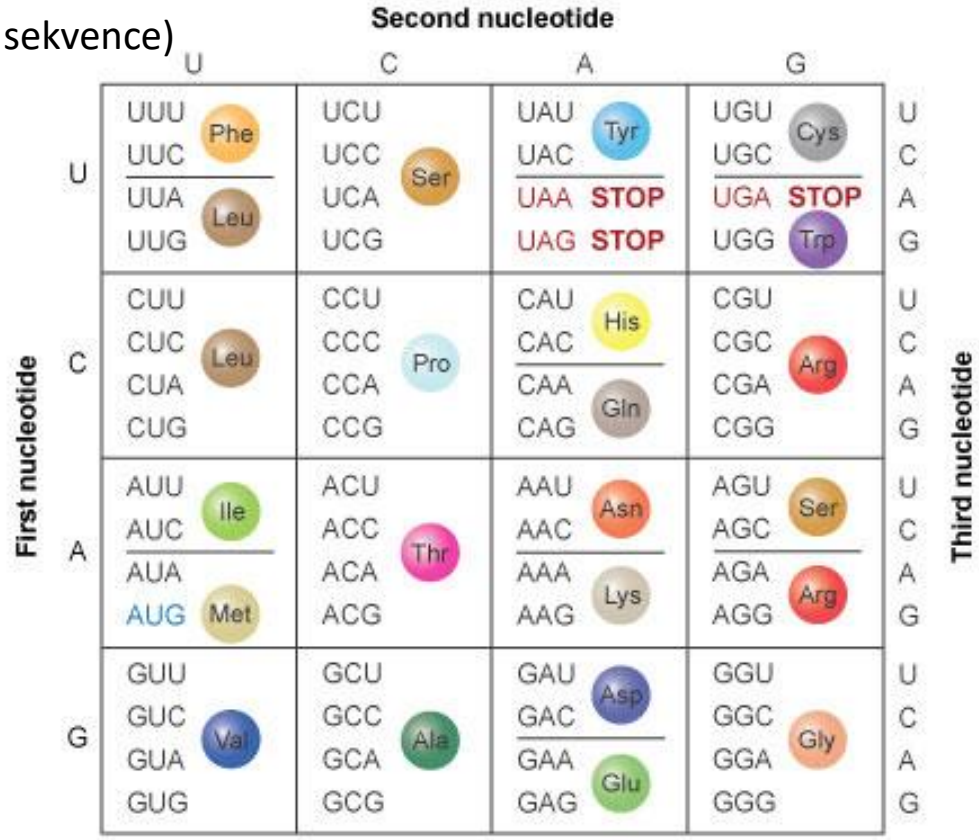
Nucleotide Codes			
A	adenosine	Y	T or C
C	cytidine	M	A or C
G	guanine	W	A or T
T	thymidine	R	G or A
N	A, G, C, or T	B	G, T, or C
U	uridine (matches T)	D	G, A, or T
K	G or T	H	A, C, or T
S	G or C	V	G, C, or A

Komplementarita: A=T, C≡G

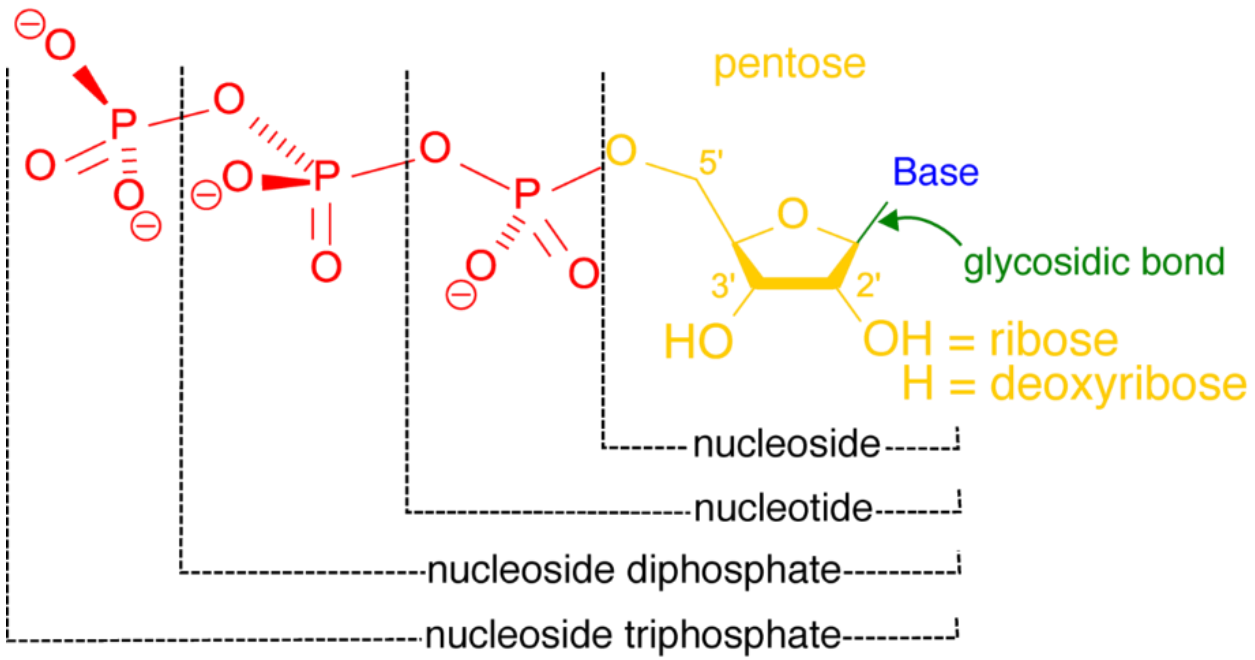
Dvojřetězcová DNA: 5´-ATGCCATAAAA-3´

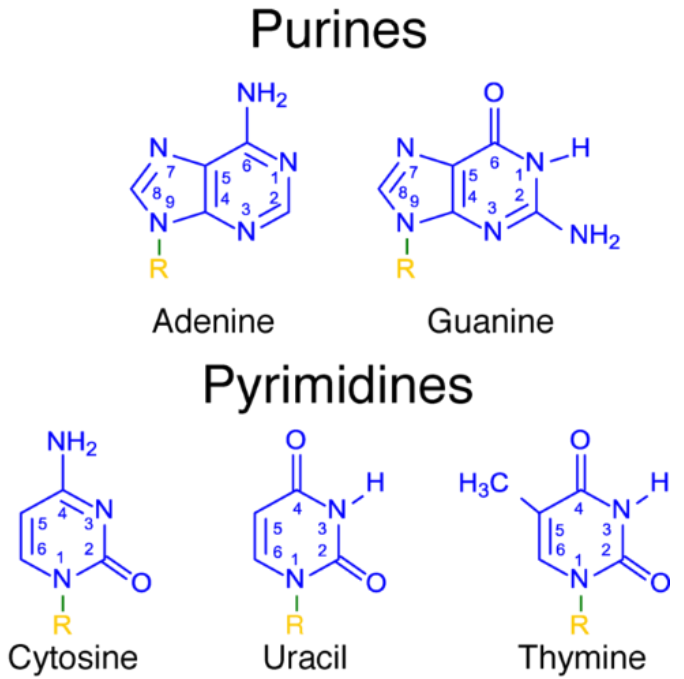
3´-TACGGTATTTT-5´

400-nt (nukleotidů) = 400 bp (base pairs/párů bazí)…kbp, Mbp

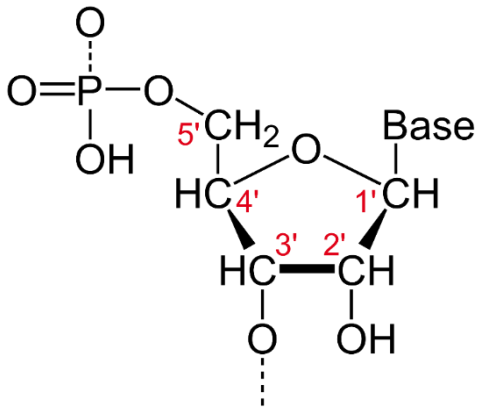


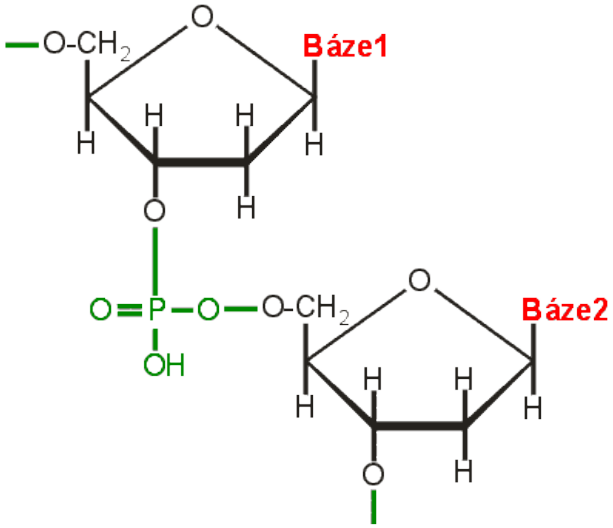
DNA/RNA

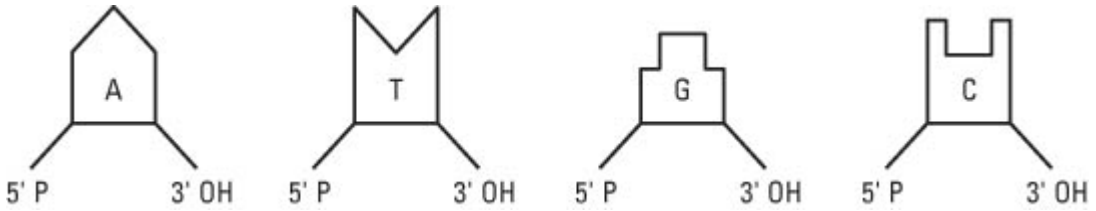

pentose
Base
glycosidic bond
5'
3'
2'
HO
OH = ribose
H = deoxyribose
nucleoside
nucleotide
nucleoside diphosphate
nucleoside triphosphate

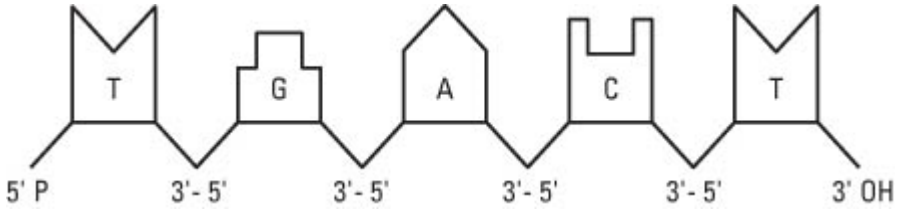

Purines
Adenine
Guanine
Pyrimidines
Cytosine
Uracil
Thymine

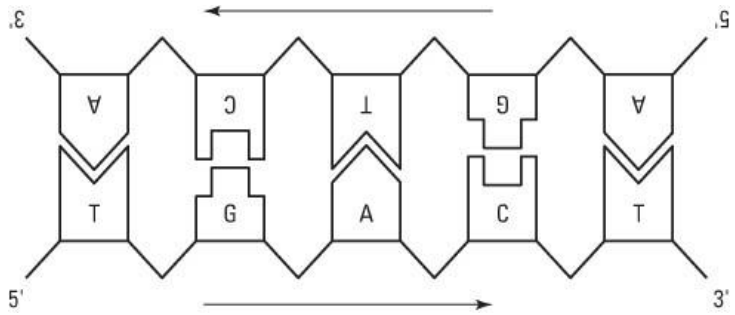
DNA/RNA


O
O=P—O
OH
5'CH2
O
Base
HC4'
1'CH
3'
2'
HC
CH
O
OH


O-CH2
O
Báze1
H
H
H
H
H
O
O=P—O—O-CH2
O
Báze2
OH
H
H
H
H
H
O


A
T
G
C
5' P
3' OH
5' P
3' OH
5' P
3' OH
5' P
3' OH


T
G
A
C
T
5' P
3'- 5'
3'- 5'
3'- 5'
3'- 5'
3' OH


3'
5'
A
C
T
G
A
T
G
A
C
T
5'
3'

DNA/RNA

4 nukleotidy: Adenin, Guanin, Thymidin (Uracil), Cytosin : **5´→ 3´**

(Frederick Sanger/Walter Gilbert 1980 Nobelova cena za určování DNA sekvence)

Genetický kód založený na tripletech:

Zápis sekvence:

A, C, T (U), G

N cokoliv

R A/G purin

Y C/T pyrimidin

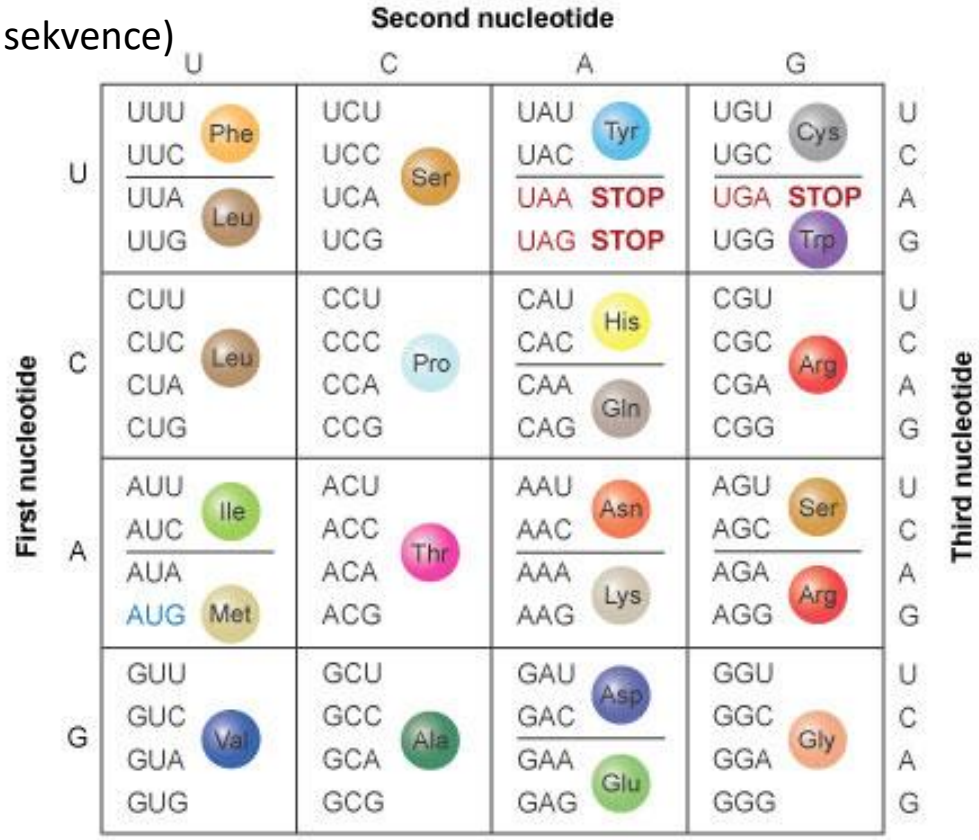
Nucleotide Codes			
A	adenosine	Y	T or C
C	cytidine	M	A or C
G	guanine	W	A or T
T	thymidine	R	G or A
N	A, G, C, or T	B	G, T, or C
U	uridine (matches T)	D	G, A, or T
K	G or T	H	A, C, or T
S	G or C	V	G, C, or A

Komplementarita: A=T, C≡G

Dvojřetězcová DNA: 5´-ATGCCATAAAA-3´

3´-TACGGTATTTT-5´

400-nt (nukleotidů) = 400 bp (base pairs/párů bazí)…kbp, Mbp



Prokaryota

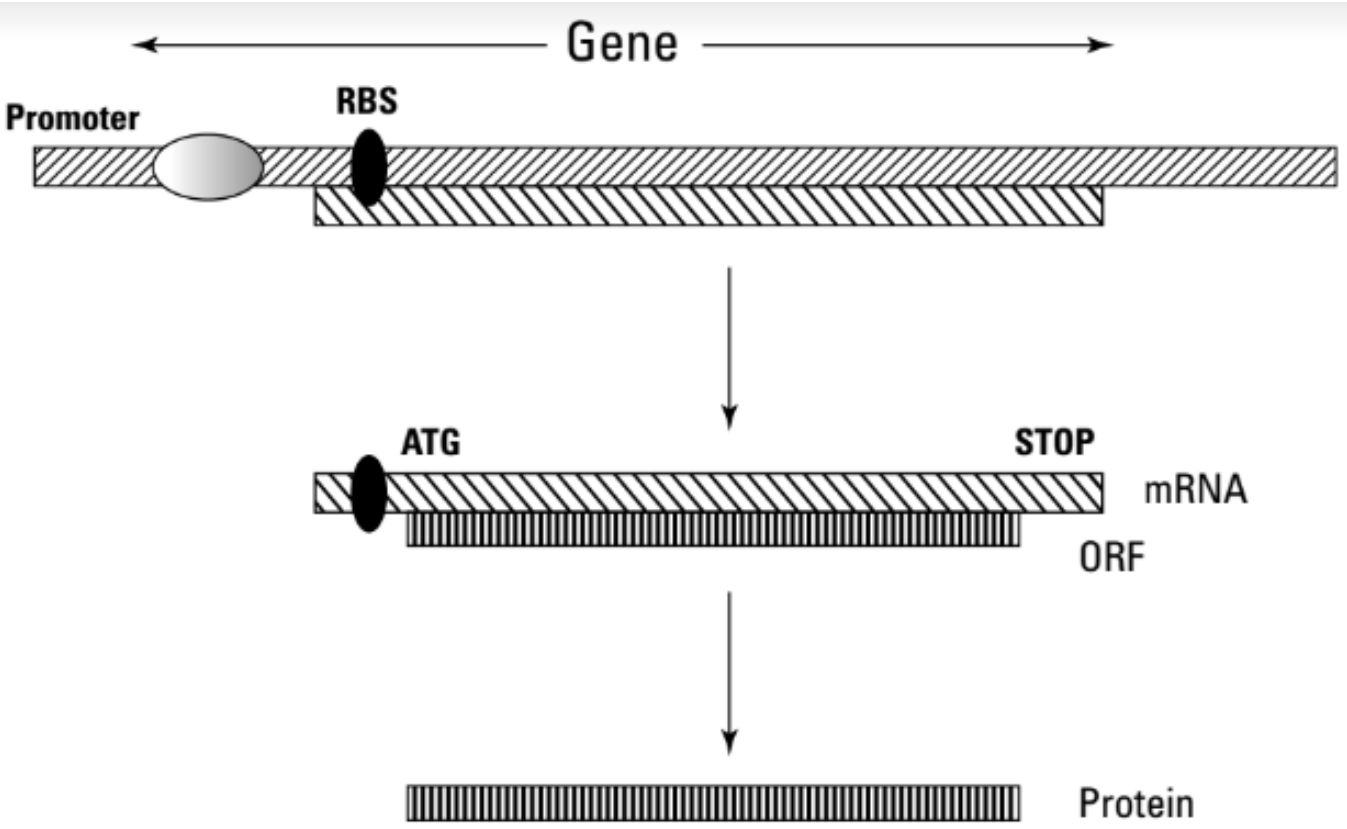
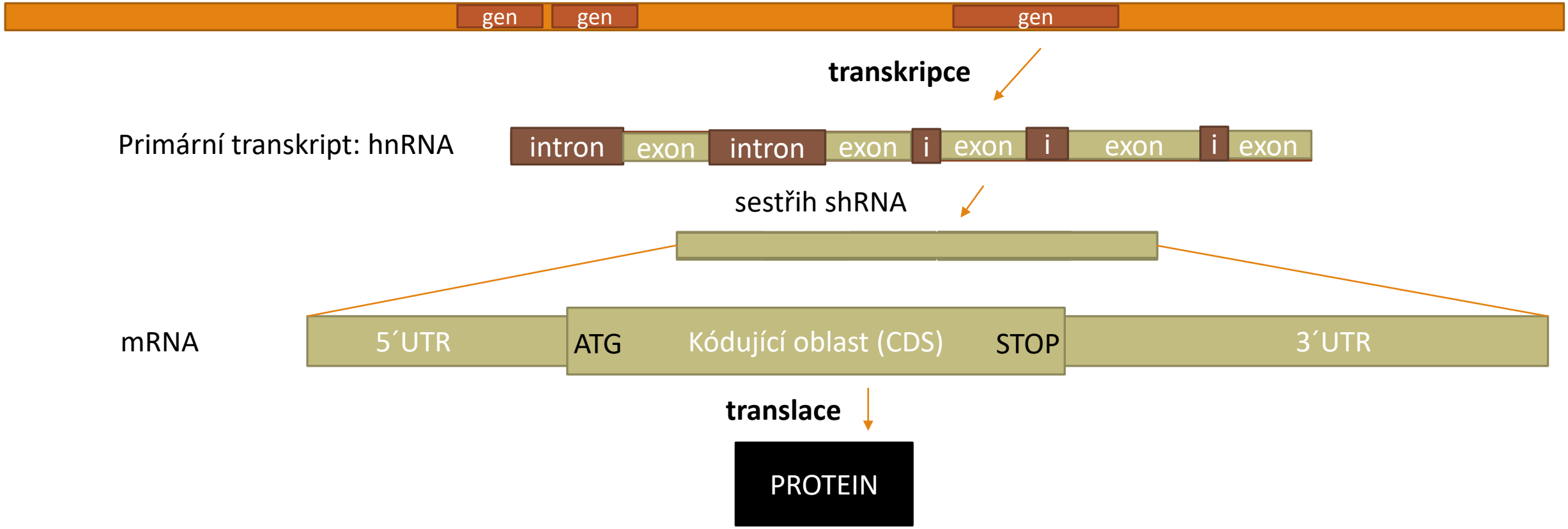


Figure 3-1: Relationship between gene, mRNA, and protein sequence for prokaryotes.

Eukaryota

Genomová DNA / chromozomy



DNA sekvence / databáze (primární)

Hlavní databáze nukleotidových sekvencí (konsorcium):

Evropská **EMBL (EBI)** (1980) / **ENA**

European Molecular Biology Laboratory / European Nucleotide Archive

http://www.ebi.ac.uk/ena/

Americká **GenBank (NCBI)** 1982

U.S. National Center for Biotechnology Information

http://www.ncbi.nlm.nih.gov/nuccore/

		GenBank	
Release	Date	Bases	Sequences
3	Dec 1982	680338	606
74	Dec 1992	120242234	97084
133	Dec 2002	28507990166	22318883
193	Dec 2012	148390863904	161140325
211	Dec 2015	203939111071	189232925
241	Dec 2020	723 003 822 007	221 467 827

Japonská **DDBJ** (1984)

DNA Data Bank of Japan

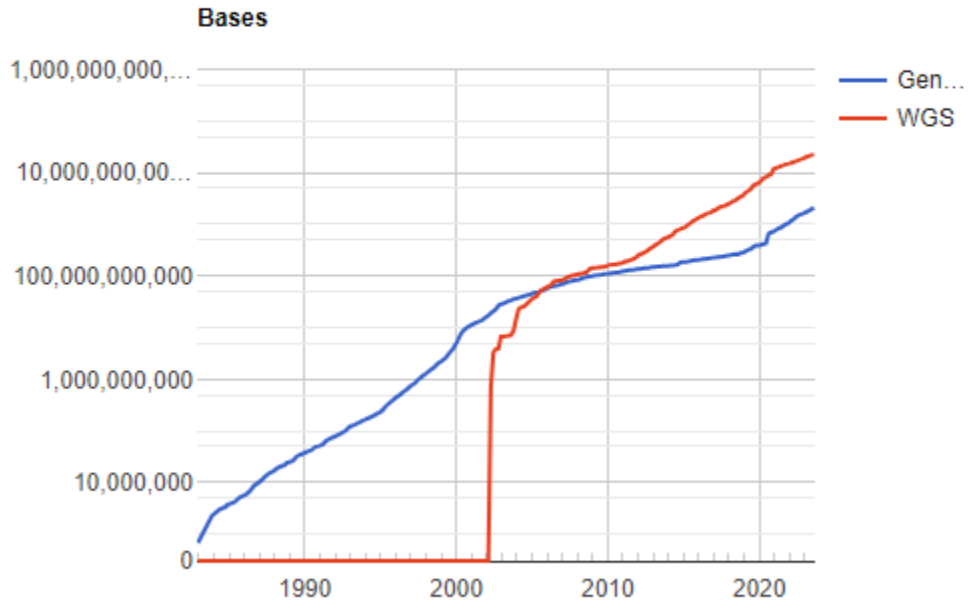
http://www.ddbj.nig.ac.

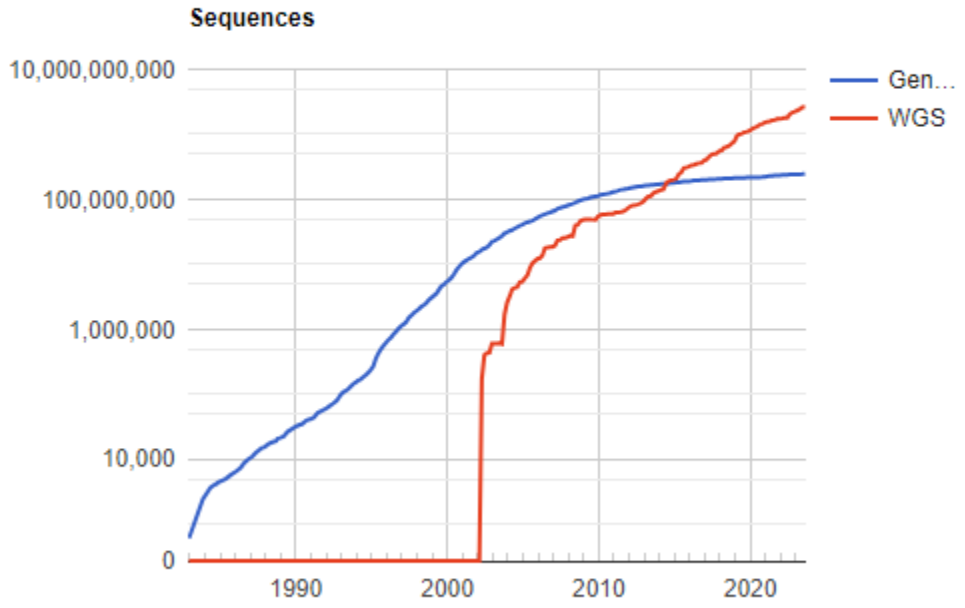
X „nemoderované“ databáze (žádná záruka, že data jsou úplně kvalitní)

Aug 2024: 3 675 462 701 077 bases, 251 998 350 sequences
Feb 2025: 5 415 448 651 743 bases, 255 669 865 sequences

GenBank/NCBI

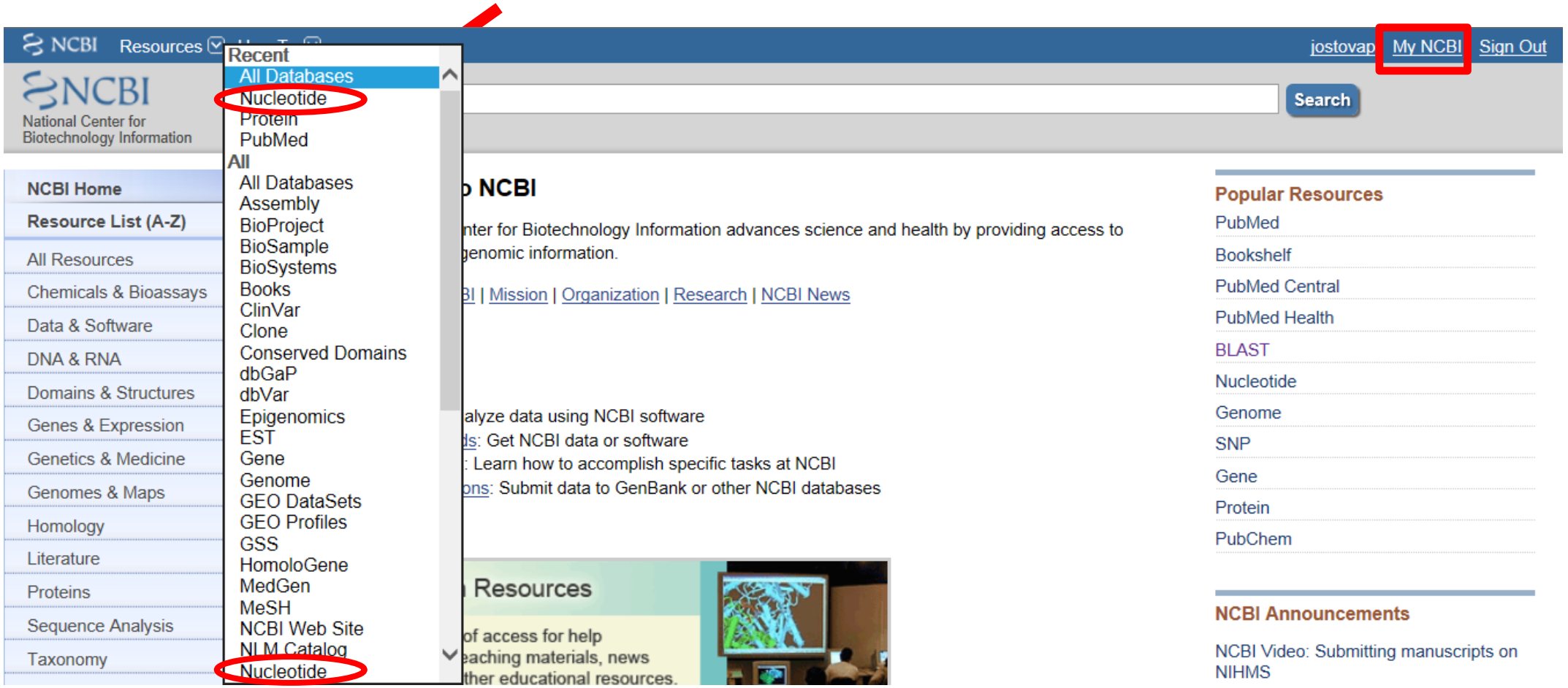
GenBank and WGS Statistics



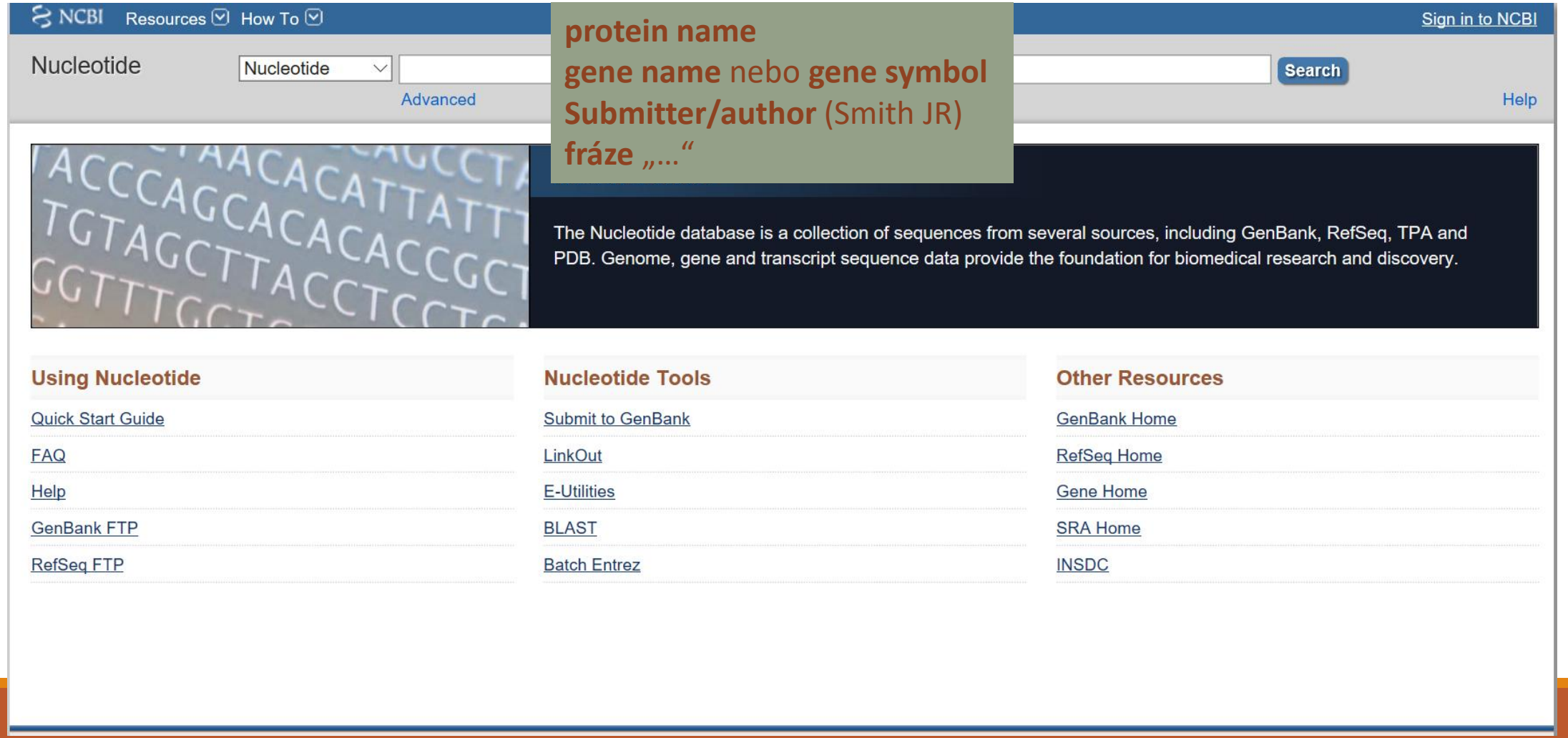


WGS = whole genome sequencing

DNA databáze NCBI



DNA databáze NCBI



DNA databáze NCBI

DNA databáze NCBI

transkript = bez intronů (mRNA)

Showing 4 of 4 (by status, accession number)

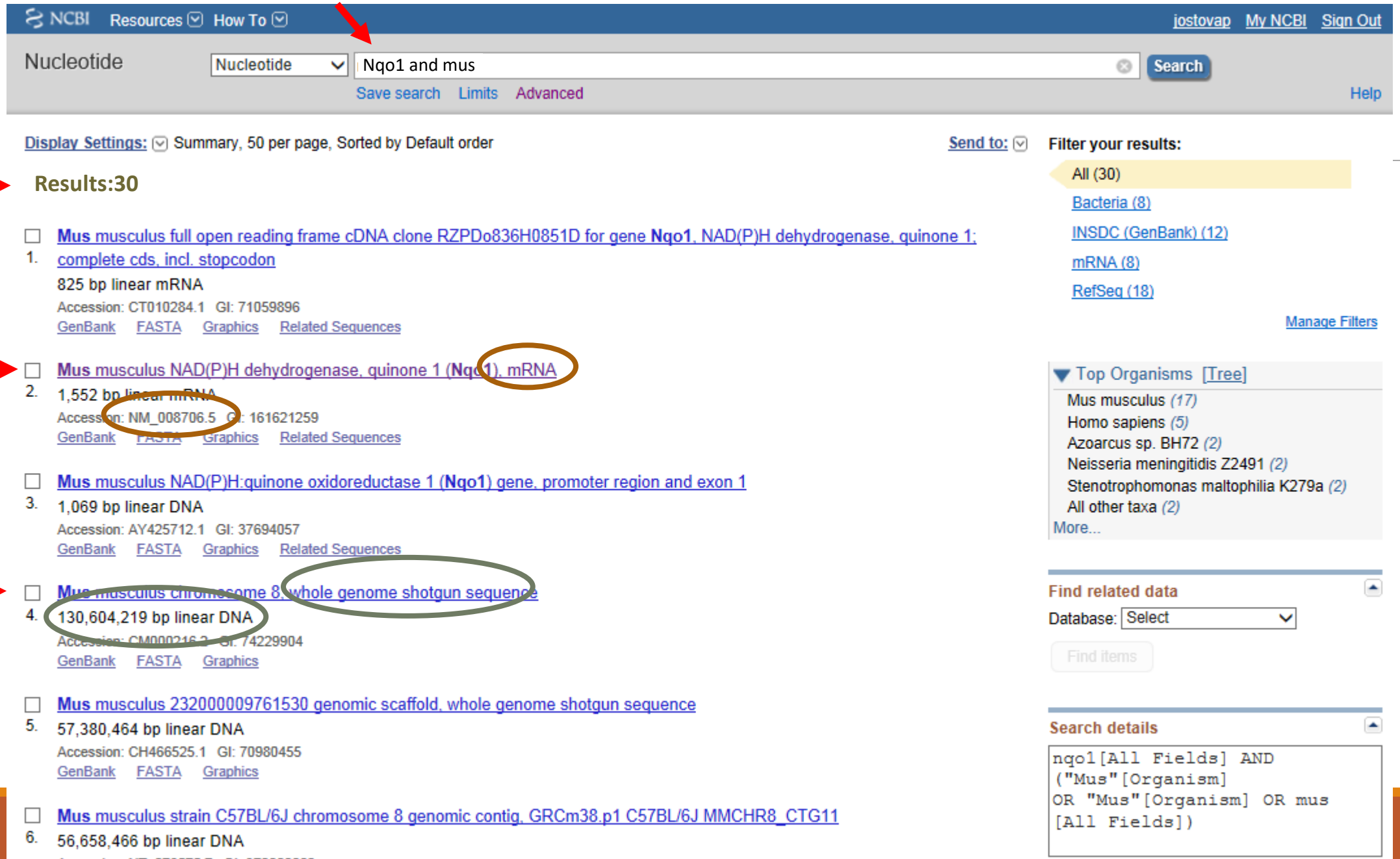
odpovídající protein

Transcript	nt	Protein	aa	Isoform	Status
NM_000903.3	2,521	NP_000894.1	274	a	MANE Select
NM_001025433.2	2,419	NP_001020604.1	240	b	curated
NM_001025434.2	2,407	NP_001020605.1	236	c	curated
NM_001286137.2	2,305	NP_001273066.1	202	d	curated

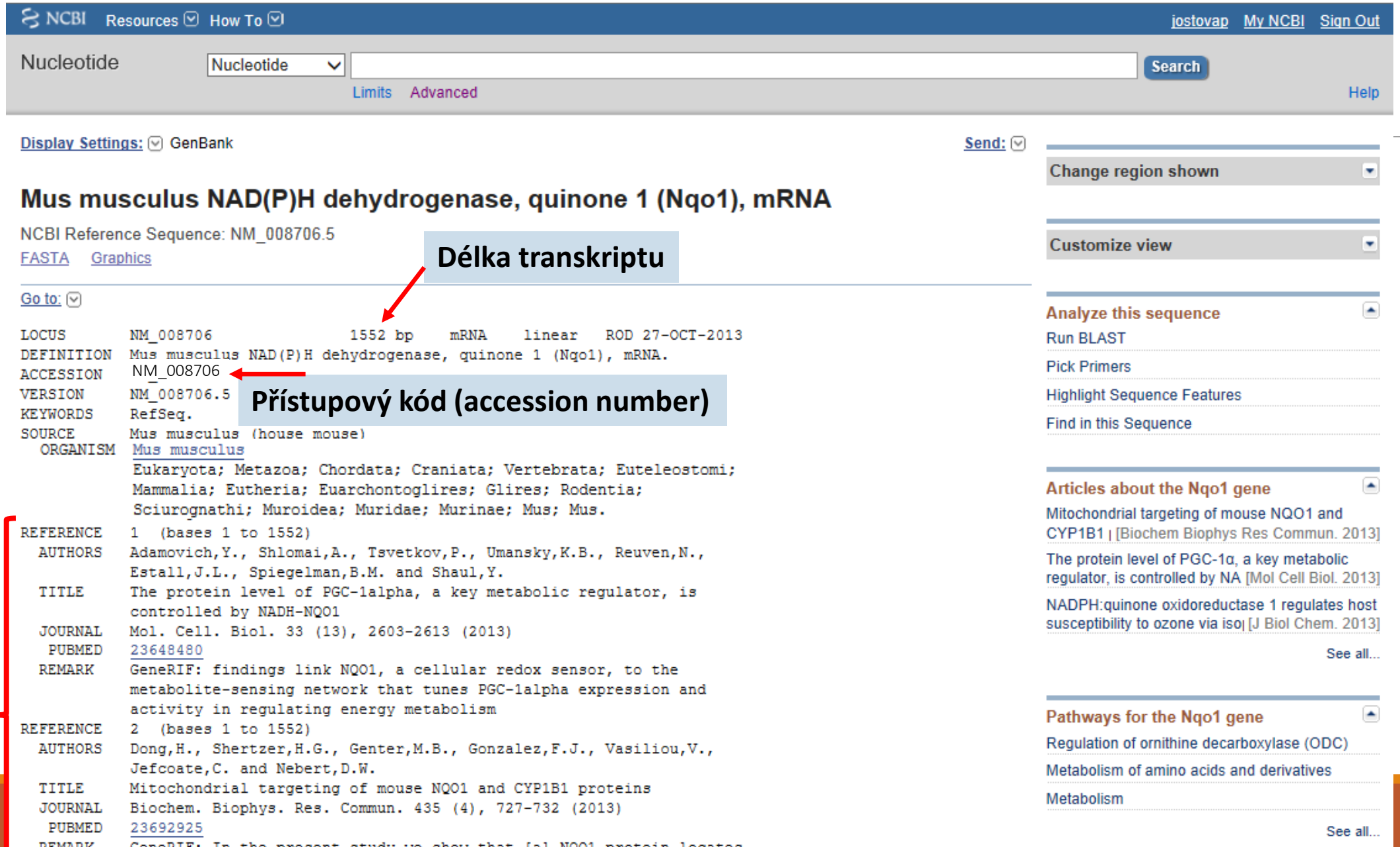
View full table

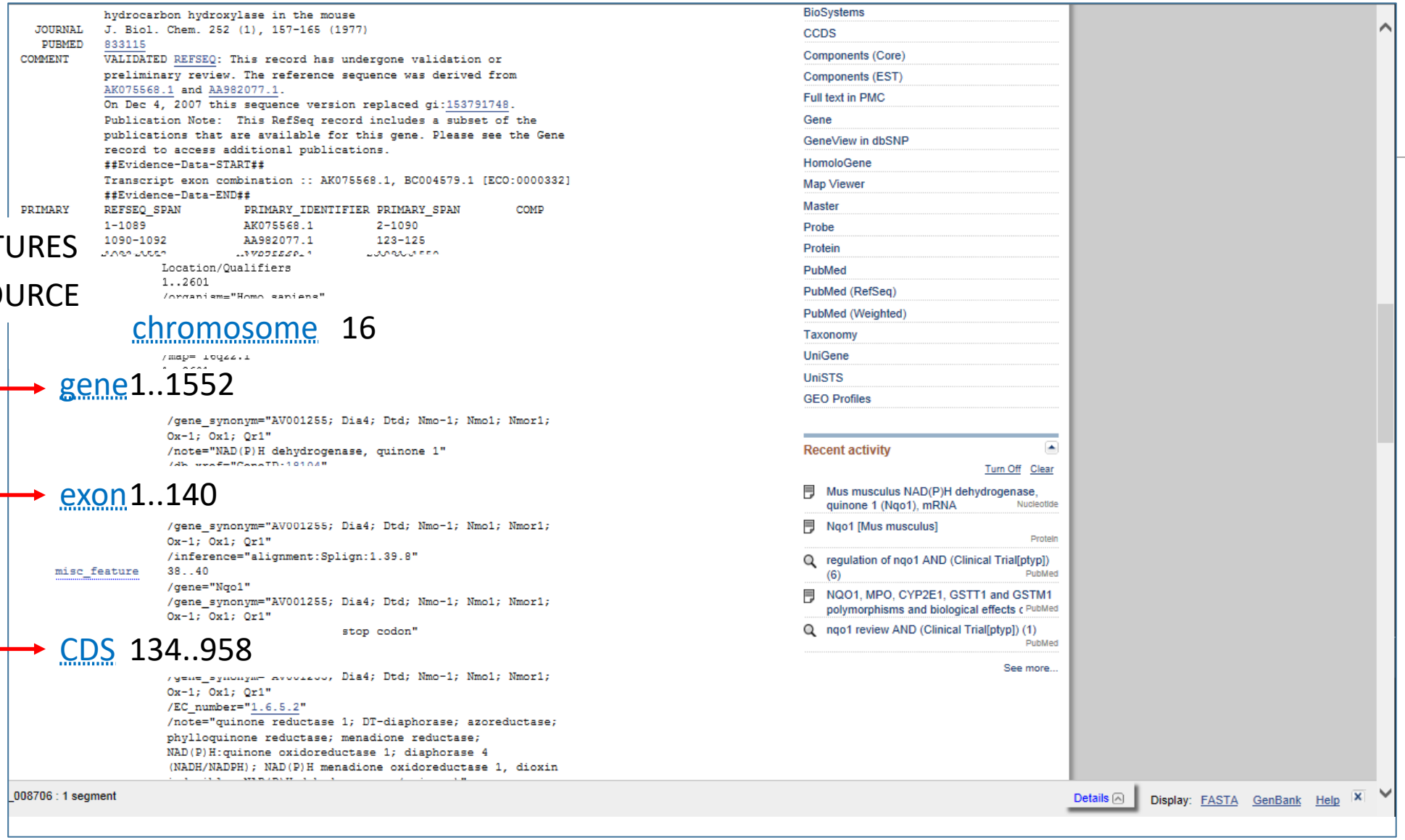
DNA databáze NCBI

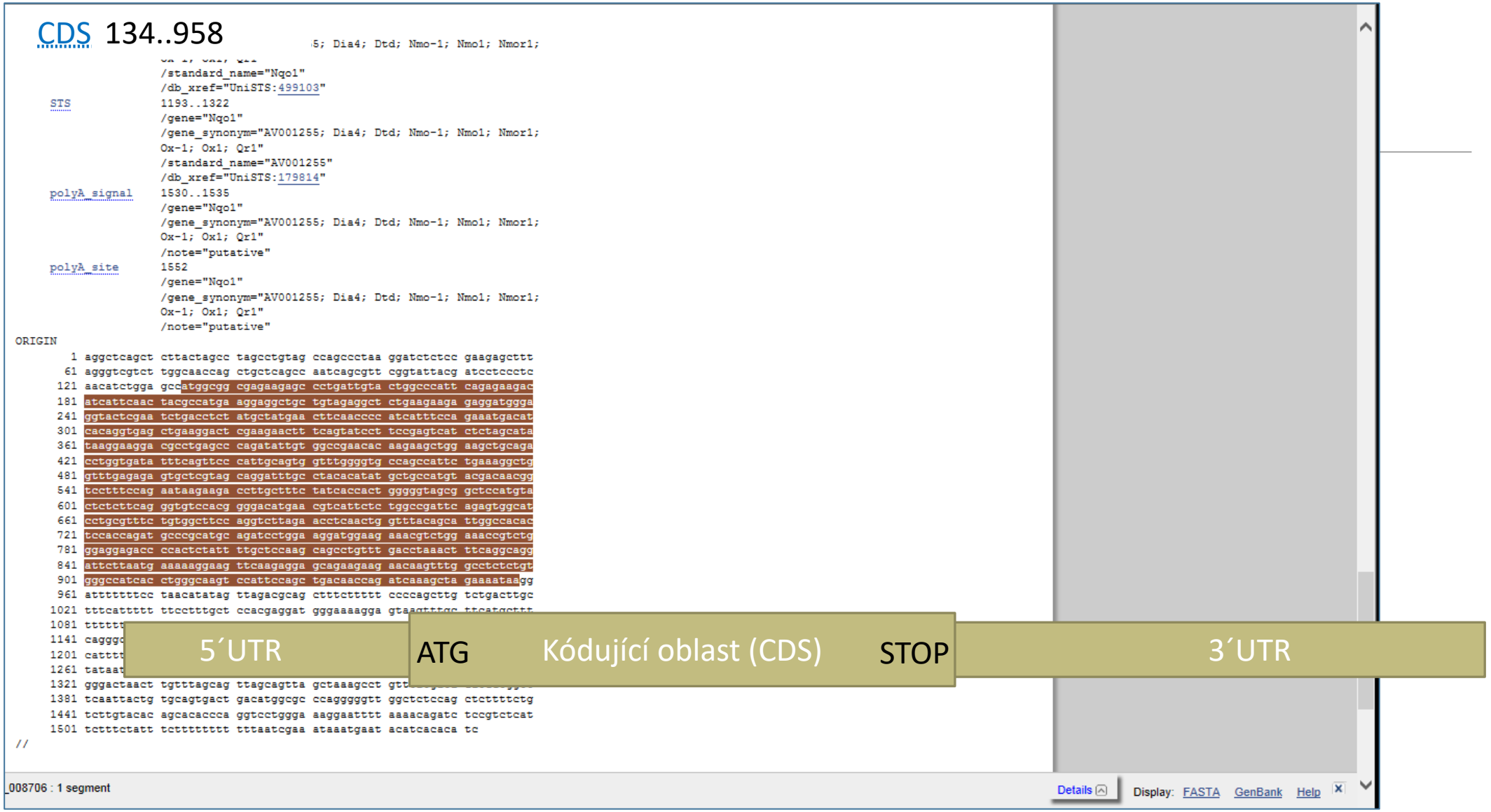
Omezení na:


NCBI Resources How To
jostovap My NCBI Sign Out
Nucleotide
Nucleotide
Nqo1 and mus
Search
Save search Limits Advanced
Help
Display Settings: Summary, 50 per page, Sorted by Default order
Send to:
Results:30
1. Mus musculus full open reading frame cDNA clone RZPDo836H0851D for gene Nqo1, NAD(P)H dehydrogenase, quinone 1; complete cds, incl. stopcodon
825 bp linear mRNA
Accession: CT010284.1 GI: 71059896
GenBank FASTA Graphics Related Sequences
2. Mus musculus NAD(P)H dehydrogenase, quinone 1 (Nqo1), mRNA
1,552 bp linear mRNA
Accession: NM_008706.5 GI: 161621259
GenBank FASTA Graphics Related Sequences
3. Mus musculus NAD(P)H:quinone oxidoreductase 1 (Nqo1) gene, promoter region and exon 1
1,069 bp linear DNA
Accession: AY425712.1 GI: 37694057
GenBank FASTA Graphics Related Sequences
4. Mus musculus chromosome 8, whole genome shotgun sequence
130,604,219 bp linear DNA
Accession: CM000216.2 GI: 74229904
GenBank FASTA Graphics
5. Mus musculus 232000009761530 genomic scaffold, whole genome shotgun sequence
57,380,464 bp linear DNA
Accession: CH466525.1 GI: 70980455
GenBank FASTA Graphics
6. Mus musculus strain C57BL/6J chromosome 8 genomic contig, GRCm38.p1 C57BL/6J MMCHR8_CTG11
56,658,466 bp linear DNA
Filter your results:
All (30)
Bacteria (8)
INSDC (GenBank) (12)
mRNA (8)
RefSeq (18)
Manage Filters
Top Organisms [Tree]
Mus musculus (17)
Homo sapiens (5)
Azoarcus sp. BH72 (2)
Neisseria meningitidis Z2491 (2)
Stenotrophomonas maltophilia K279a (2)
All other taxa (2)
More...
Find related data
Database: Select
Find items
Search details
nqo1[All Fields] AND ("Mus"[Organism] OR "Mus"[Organism] OR mus [All Fields])

„GenBank flatfile“




hydrocarbon hydroxylase in the mouse
JOURNAL J. Biol. Chem. 252 (1), 157-165 (1977)
PUBMED 833115
COMMENT VALIDATED REFSEQ: This record has undergone validation or preliminary review. The reference sequence was derived from AK075568.1 and AA982077.1.
On Dec 4, 2007 this sequence version replaced gi:153791748.
Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications.
##Evidence-Data-START##
Transcript exon combination :: AK075568.1, BC004579.1 [ECO:0000332]
##Evidence-Data-END##
PRIMARY REFSEQ_SPAN PRIMARY_IDENTIFIER PRIMARY_SPAN COMP
1-1089 AK075568.1 2-1090
1090-1092 AA982077.1 123-125
FEATURES
SOURCE
Location/Qualifiers
1..2601
chromosome 16
gene 1..1552
/gene_synonym="AV001255; Dia4; Dtd; Nmo-1; Nmo1; Nmor1; Ox-1; Ox1; Qr1"
/note="NAD(P)H dehydrogenase, quinone 1"
exon 1..140
/gene_synonym="AV001255; Dia4; Dtd; Nmo-1; Nmo1; Nmor1; Ox-1; Ox1; Qr1"
/inference="alignment:Splign:1.39.8"
misc_feature 38..40
/gene="Nqo1"
/gene_synonym="AV001255; Dia4; Dtd; Nmo-1; Nmo1; Nmor1; Ox-1; Ox1; Qr1"
stop codon"
CDS 134..958
Ox-1; Ox1; Qr1"
/EC_number="1.6.5.2"
/note="quinone reductase 1; DT-diaphorase; azoreductase; phylloquinone reductase; menadione reductase; NAD(P)H:quinone oxidoreductase 1; diaphorase 4 (NADH/NADPH); NAD(P)H menadione oxidoreductase 1, dioxin
BioSystems
CCDS
Components (Core)
Components (EST)
Full text in PMC
Gene
GeneView in dbSNP
HomoloGene
Map Viewer
Master
Probe
Protein
PubMed
PubMed (RefSeq)
PubMed (Weighted)
Taxonomy
UniGene
UniSTS
GEO Profiles
Recent activity
Turn Off Clear
Mus musculus NAD(P)H dehydrogenase, quinone 1 (Nqo1), mRNA Nucleotide
Nqo1 [Mus musculus] Protein
regulation of nqo1 AND (Clinical Trial[ptyp]) (6) PubMed
NQO1, MPO, CYP2E1, GSTT1 and GSTM1 polymorphisms and biological effects c PubMed
nqo1 review AND (Clinical Trial[ptyp]) (1) PubMed
See more...
_008706 : 1 segment
Details
Display: FASTA GenBank Help


CDS 134..958
5; Dia4; Dtd; Nmo-1; Nmo1; Nmor1;
/standard_name="Nqo1"
/db_xref="UniSTS:499103"
STS 1193..1322
/gene="Nqo1"
/gene_synonym="AV001255; Dia4; Dtd; Nmo-1; Nmo1; Nmor1;
Ox-1; Ox1; Qr1"
/standard_name="AV001255"
/db_xref="UniSTS:179814"
polyA_signal 1530..1535
/gene="Nqo1"
/gene_synonym="AV001255; Dia4; Dtd; Nmo-1; Nmo1; Nmor1;
Ox-1; Ox1; Qr1"
/note="putative"
polyA_site 1552
/gene="Nqo1"
/gene_synonym="AV001255; Dia4; Dtd; Nmo-1; Nmo1; Nmor1;
Ox-1; Ox1; Qr1"
/note="putative"
ORIGIN
1 aggctcagct cttactagcc tagcctgtag ccagccctaa ggatctctcc gaagagcttt
61 agggtcgtct tggcaaccag ctgctcagcc aatcagcgtt cggtattacg atcctccctc
121 aacatctgga gccatggcgg cgagaagagc cctgattgta ctggcccatt cagagaagac
181 atcattcaac tacgccatga aggaggctgc tgtagaggct ctgaagaaga gaggatggga
241 ggtactcgaa tctgacctct atgctatgaa cttcaacccc atcatttcca gaaatgacat
301 cacaggtgag ctgaaggact cgaagaactt tcagtatcct tccgagtcat ctctagcata
361 taaggaagga cgcctgagcc cagatattgt ggccgaacac aagaagctgg aagctgcaga
421 cctggtgata tttcagttcc cattgcagtg gtttggggtg ccagccattc tgaaaggctg
481 gtttgagaga gtgctcgtag caggatttgc ctacacatat gctgccatgt acgacaacgg
541 tcctttccag aataagaaga ccttgctttc tatcaccact gggggtagcg gctccatgta
601 ctctcttcag ggtgtccacg gggacatgaa cgtcattctc tggccgattc agagtggcat
661 cctgcgtttc tgtggcttcc aggtcttaga acctcaactg gtttacagca ttggccacac
721 tccaccagat gcccgcatgc agatcctgga aggatggaag aaacgtctgg aaaccgtctg
781 ggaggagacc ccactctatt ttgctccaag cagcctgttt gacctaaact ttcaggcagg
841 attcttaatg aaaaaggaag ttcaagagga gcagaagaag aacaagtttg gcctctctgt
901 gggccatcac ctgggcaagt ccattccagc tgacaaccag atcaaagcta gaaaataagg
961 attttttttcc taacatatag ttagacgcag ctttcttttt ccccagcttg tctgacttgc
1021 tttcattttt ttccttgct ccacgaggat gggaaaagga gtaagtttgc ttcatgcttt
1081 ttttt
1141 cagggc
1201 catttt
1261 tataat
1321 gggactaact tgtttagcag ttagcagtta gctaaagcct gtt
1381 tcaattactg tgcagtgact gacatggcgc ccaggggtt ggctctccag ctcttttctg
1441 tcttgtacac agcacaccca ggtcctggga aaggaatttt aaaacagatc tccgtctcat
1501 tctttctatt tctttttttt tttaatcgaa ataaatgaat acatcacaca tc
//
5´UTR
ATG
Kódující oblast (CDS)
STOP
3´UTR
008706 : 1 segment
Details
Display: FASTA GenBank Help

Display Settings: GenBank Send:

Mus musculus NAD(P)H dehydrogenase, quinone 1 (Nqo1), mRNA

e Sequence: NM_008706.5

FASTA

ics

Go to:

```
LOCUS       NM_008706               1552 bp    mRNA    linear   ROD 27-OCT-2013
DEFINITION  Mus musculus NAD(P)H dehydrogenase, quinone 1 (Nqo1), mRNA.
ACCESSION   NM_008706
VERSION     NM_008706.5  GI:161621259
KEYWORDS    RefSeq.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus; Mus.
REFERENCE   1  (bases 1 to 1552)
  AUTHORS   Adamovich,Y., Shlomai,A., Tsvetkov,P., Umansky,K.B., Reuven,N.,
            Estall,J.L., Spiegelman,B.M. and Shaul,Y.
  TITLE     The protein level of PGC-1alpha, a key metabolic regulator, is
            controlled by NADH-NQO1
  JOURNAL   Mol. Cell. Biol. 33 (13), 2603-2613 (2013)
   PUBMED   23648480
  REMARK    GeneRIF: findings link NQO1, a cellular redox sensor, to the
            metabolite-sensing network that tunes PGC-1alpha expression and
            activity in regulating energy metabolism
REFERENCE   2  (bases 1 to 1552)
  AUTHORS   Dong,H., Shertzer,H.G., Genter,M.B., Gonzalez,F.J., Vasiliou,V.,
            Jefcoate,C. and Nebert,D.W.
  TITLE     Mitochondrial targeting of mouse NQO1 and CYP1B1 proteins
  JOURNAL   Biochem. Biophys. Res. Commun. 435 (4), 727-732 (2013)
   PUBMED   23692925
  REMARK    GeneRIF: In the present study we show that [a] NQO1 protein locates
            to cytosol, ER and mitochondria.
REFERENCE   3  (bases 1 to 1552)
  AUTHORS   Kummarapurugu,A.B., Fischer,B.M., Zheng,S., Milne,G.L., Ghio,A.J.,
```

Change region shown

Customize view

Analyze this sequence

Run BLAST

Pick Primers

Highlight Sequence Features

Find in this Sequence

Articles about the Nqo1 gene

Mitochondrial targeting of mouse NQO1 and CYP1B1 | [Biochem Biophys Res Commun. 2013]

The protein level of PGC-1α, a key metabolic regulator, is controlled by NA [Mol Cell Biol. 2013]

NADPH:quinone oxidoreductase 1 regulates host susceptibility to ozone via iso| [J Biol Chem. 2013]

See all...

Pathways for the Nqo1 gene

Regulation of ornithine decarboxylase (ODC)

Metabolism of amino acids and derivatives

Metabolism

See all...

Reference sequence information

FASTA (a RAW) formát

název odvozen:

FASTA = populární program pro porovnávání sekvencí a prohledávání databází

W.R. Pearson a D.J. Lipman 1988

následující formát sekvence:

>NQO1_mus
AGGCTCAGCTCTTACTAGCCTAGCCTGTA
GCCAGCCCTAAGGATCTCTCCGAAGAGCT
AATGTCGG

AGGCTCAGCTCTTACTAGCCTAGCCTGTA
GCCAGCCCTAAGGATCTCTCCGAAGAGCT
AATGTCGG

1. „definiční“ řádek (hlavička) začíná **>** a obsahuje unikátní identifikátor, následovaný krátkým popisem (nemusí být).

2. řádek – vlastní sekvence (DNA/protein- v jedno-písmenném kódování (až po další symbol >, určující začátek další sekvence.

Display Settings: GenBank Send:

Mus musculus NAD(P)H dehydrogenase, quinone 1 (Nqo1), mRNA

e Sequence: NM_008706.5

FASTA

ics

Go to:

```
LOCUS       NM_008706               1552 bp    mRNA    linear   ROD 27-OCT-2013
DEFINITION  Mus musculus NAD(P)H dehydrogenase, quinone 1 (Nqo1), mRNA.
ACCESSION   NM_008706
VERSION     NM_008706.5  GI:161621259
KEYWORDS    RefSeq.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus; Mus.
REFERENCE   1  (bases 1 to 1552)
  AUTHORS   Adamovich,Y., Shlomai,A., Tsvetkov,P., Umansky,K.B., Reuven,N.,
            Estall,J.L., Spiegelman,B.M. and Shaul,Y.
  TITLE     The protein level of PGC-1alpha, a key metabolic regulator, is
            controlled by NADH-NQO1
  JOURNAL   Mol. Cell. Biol. 33 (13), 2603-2613 (2013)
   PUBMED   23648480
  REMARK    GeneRIF: findings link NQO1, a cellular redox sensor, to the
            metabolite-sensing network that tunes PGC-1alpha expression and
            activity in regulating energy metabolism
REFERENCE   2  (bases 1 to 1552)
  AUTHORS   Dong,H., Shertzer,H.G., Genter,M.B., Gonzalez,F.J., Vasiliou,V.,
            Jefcoate,C. and Nebert,D.W.
  TITLE     Mitochondrial targeting of mouse NQO1 and CYP1B1 proteins
  JOURNAL   Biochem. Biophys. Res. Commun. 435 (4), 727-732 (2013)
   PUBMED   23692925
  REMARK    GeneRIF: In the present study we show that [a] NQO1 protein locates
            to cytosol, ER and mitochondria.
REFERENCE   3  (bases 1 to 1552)
  AUTHORS   Kummarapurugu,A.B., Fischer,B.M., Zheng,S., Milne,G.L., Ghio,A.J.,
```

Change region shown

Customize view

Analyze this sequence

Run BLAST

Pick Primers

Highlight Sequence Features

Find in this Sequence

Articles about the Nqo1 gene

Mitochondrial targeting of mouse NQO1 and CYP1B1 | [Biochem Biophys Res Commun. 2013]

The protein level of PGC-1α, a key metabolic regulator, is controlled by NA [Mol Cell Biol. 2013]

NADPH:quinone oxidoreductase 1 regulates host susceptibility to ozone via iso| [J Biol Chem. 2013]

See all...

Pathways for the Nqo1 gene

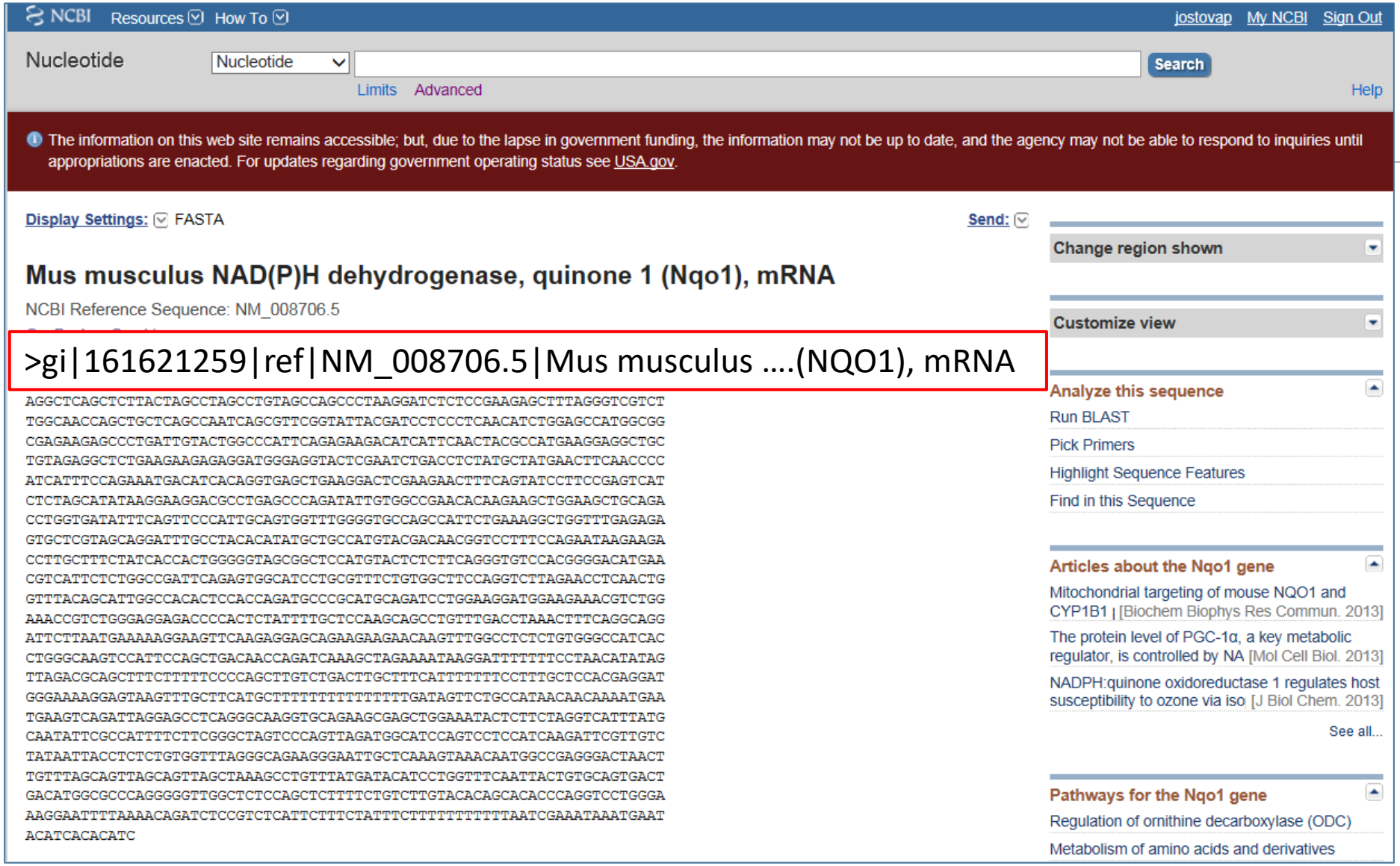
Regulation of ornithine decarboxylase (ODC)

Metabolism of amino acids and derivatives

Metabolism

See all...

Reference sequence information

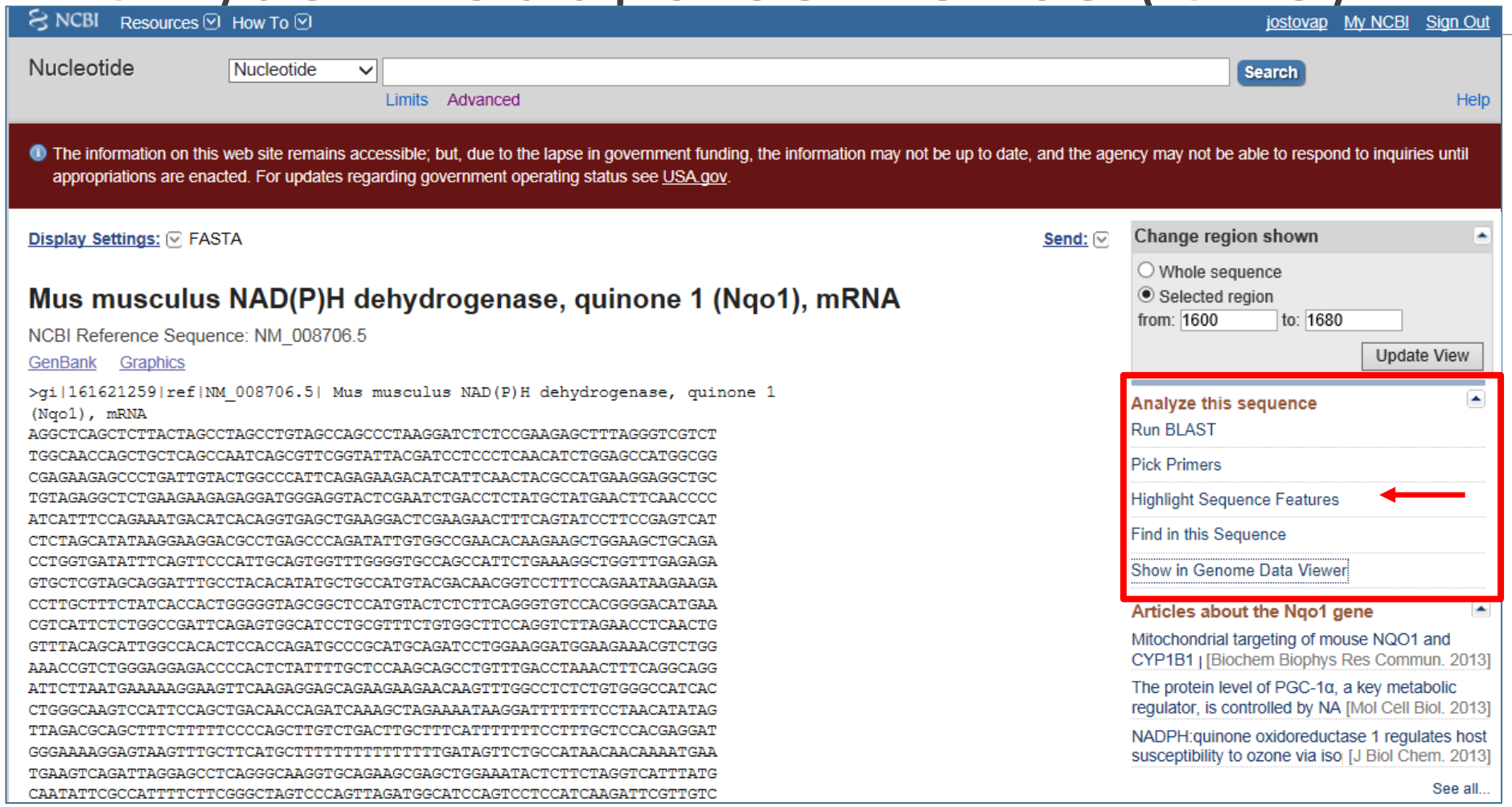

NCBI Resources How To
jostovap My NCBI Sign Out
Nucleotide
Nucleotide
Search
Limits Advanced
Help
The information on this web site remains accessible; but, due to the lapse in government funding, the information may not be up to date, and the agency may not be able to respond to inquiries until appropriations are enacted. For updates regarding government operating status see USA.gov.
Display Settings: FASTA
Send:
Mus musculus NAD(P)H dehydrogenase, quinone 1 (Nqo1), mRNA
NCBI Reference Sequence: NM_008706.5
>gi|161621259|ref|NM_008706.5|Mus musculus ….(NQO1), mRNA
AGGCTCAGCTCTTACTAGCCTAGCCTGTAGCCAGCCCTAAGGATCTCTCCGAAGAGCTTTAGGGTCGTCT
TGGCAACCAGCTGCTCAGCCAATCAGCGTTCGGTATTACGATCCTCCCTCAACATCTGGAGCCATGGCGG
CGAGAAGAGCCCTGATTGTACTGGCCCATTCAGAGAAGACATCATTCAACTACGCCATGAAGGAGGCTGC
TGTAGAGGCTCTGAAGAAGAGAGGATGGGAGGTACTCGAATCTGACCTCTATGCTATGAACTTCAACCCC
ATCATTTCCAGAAATGACATCACAGGTGAGCTGAAGGACTCGAAGAACTTTCAGTATCCTTCCGAGTCAT
CTCTAGCATATAAGGAAGGACGCCTGAGCCCAGATATTGTGGCCGAACACAAGAAGCTGGAAGCTGCAGA
CCTGGTGATATTTCAGTTCCCATTGCAGTGGTTTGGGGTGCCAGCCATTCTGAAAGGCTGGTTTGAGAGA
GTGCTCGTAGCAGGATTTGCCTACACATATGCTGCCATGTACGACAACGGTCCTTTCCAGAATAAGAAGA
CCTTGCTTTCTATCACCACTGGGGGTAGCGGCTCCATGTACTCTCTTCAGGGTGTCCACGGGGACATGAA
CGTCATTCTCTGGCCGATTCAGAGTGGCATCCTGCGTTTCTGTGGCTTCCAGGTCTTAGAACCTCAACTG
GTTTACAGCATTGGCCACACTCCACCAGATGCCCGCATGCAGATCCTGGAAGGATGGAAGAAACGTCTGG
AAACCGTCTGGGAGGAGACCCCACTCTATTTTGCTCCAAGCAGCCTGTTTGACCTAAACTTTCAGGCAGG
ATTCTTAATGAAAAAGGAAGTTCAAGAGGAGCAGAAGAAGAACAAGTTTGGCCTCTCTGTGGGCCATCAC
CTGGGCAAGTCCATTCCAGCTGACAACCAGATCAAAGCTAGAAAATAAGGATTTTTTTCCTAACATATAG
TTAGACGCAGCTTTCTTTTTCCCCAGCTTGTCTGACTTGCTTTCATTTTTTTCCTTTGCTCCACGAGGAT
GGGAAAAGGAGTAAGTTTGCTTCATGCTTTTTTTTTTTTTTTGATAGTTCTGCCATAACAACAAAATGAA
TGAAGTCAGATTAGGAGCCTCAGGGCAAGGTGCAGAAGCGAGCTGGAAATACTCTTCTAGGTCATTTATG
CAATATTCGCCATTTTCTTCGGGCTAGTCCCAGTTAGATGGCATCCAGTCCTCCATCAAGATTCGTTGTC
TATAATTACCTCTCTGTGGTTTAGGGCAGAAGGGAATTGCTCAAAGTAAACAATGGCCGAGGGACTAACT
TGTTTAGCAGTTAGCAGTTAGCTAAAGCCTGTTTATGATACATCCTGGTTTCAATTACTGTGCAGTGACT
GACATGGCGCCCAGGGGGTTGGCTCTCCAGCTCTTTTCTGTCTTGTACACAGCACACCCAGGTCCTGGGA
AAGGAATTTTAAAACAGATCTCCGTCTCATTCTTTCTATTTCTTTTTTTTTTTAATCGAAATAAATGAAT
ACATCACACATC
Change region shown
Customize view
Analyze this sequence
Run BLAST
Pick Primers
Highlight Sequence Features
Find in this Sequence
Articles about the Nqo1 gene
Mitochondrial targeting of mouse NQO1 and CYP1B1 | [Biochem Biophys Res Commun. 2013]
The protein level of PGC-1α, a key metabolic regulator, is controlled by NA [Mol Cell Biol. 2013]
NADPH:quinone oxidoreductase 1 regulates host susceptibility to ozone via iso [J Biol Chem. 2013]
See all...
Pathways for the Nqo1 gene
Regulation of ornithine decarboxylase (ODC)
Metabolism of amino acids and derivatives

fasta

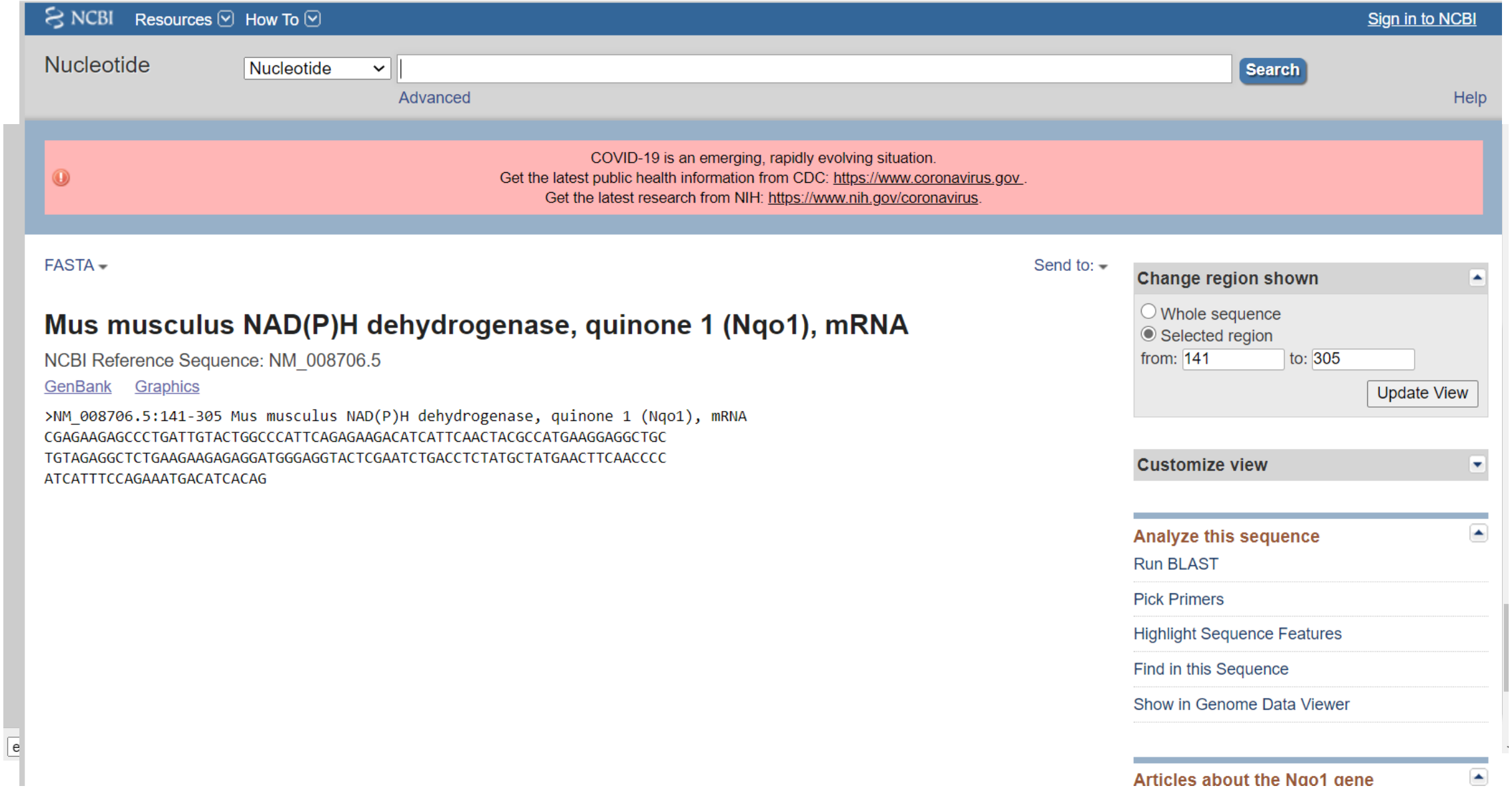
NCBI/Jen kódující sekvence (CDS)

„nakliknout“ oblast CDS → zabarví se příslušná část sekvence hnědě (FASTA odkaz vpravo dole ukáže jen tuto část)

NCBI/Jen kódující sekvence (CDS)



V řádku dole:


NCBI Resources How To
Sign in to NCBI
Nucleotide
Nucleotide
Search
Advanced
Help
COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: https://www.coronavirus.gov .
Get the latest research from NIH: https://www.nih.gov/coronavirus.
FASTA
Send to:
Mus musculus NAD(P)H dehydrogenase, quinone 1 (Nqo1), mRNA
NCBI Reference Sequence: NM_008706.5
GenBank Graphics
>NM_008706.5:141-305 Mus musculus NAD(P)H dehydrogenase, quinone 1 (Nqo1), mRNA
CGAGAAGAGCCCTGATTGTACTGGCCCATTCAGAGAAGACATCATTCAACTACGCCATGAAGGAGGCTGC
TGTAGAGGCTCTGAAGAAGAGAGGATGGGAGGTACTCGAATCTGACCTCTATGCTATGAACTTCAACCCC
ATCATTTCCAGAAATGACATCACAG
Change region shown
Whole sequence
Selected region
from: 141 to: 305
Update View
Customize view
Analyze this sequence
Run BLAST
Pick Primers
Highlight Sequence Features
Find in this Sequence
Show in Genome Data Viewer
Articles about the Nqo1 gene

Display Settings: GenBank Send:

Mus musculus NAD(P)H dehydrogenase, quinone 1 (Nqo1), mRNA

NCBI R… …M_008706.5

FASTA GRAPHIC

Go to:

```
LOCUS       NM_008706               1552 bp    mRNA    linear   ROD 27-OCT-2013
DEFINITION  Mus musculus NAD(P)H dehydrogenase, quinone 1 (Nqo1), mRNA.
ACCESSION   NM_008706
VERSION     NM_008706.5  GI:161621259
KEYWORDS    RefSeq.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus; Mus.
REFERENCE   1  (bases 1 to 1552)
  AUTHORS   Adamovich,Y., Shlomai,A., Tsvetkov,P., Umansky,K.B., Reuven,N.,
            Estall,J.L., Spiegelman,B.M. and Shaul,Y.
  TITLE     The protein level of PGC-1alpha, a key metabolic regulator, is
            controlled by NADH-NQO1
  JOURNAL   Mol. Cell. Biol. 33 (13), 2603-2613 (2013)
   PUBMED   23648480
  REMARK    GeneRIF: findings link NQO1, a cellular redox sensor, to the
            metabolite-sensing network that tunes PGC-1alpha expression and
            activity in regulating energy metabolism
REFERENCE   2  (bases 1 to 1552)
  AUTHORS   Dong,H., Shertzer,H.G., Genter,M.B., Gonzalez,F.J., Vasiliou,V.,
            Jefcoate,C. and Nebert,D.W.
  TITLE     Mitochondrial targeting of mouse NQO1 and CYP1B1 proteins
  JOURNAL   Biochem. Biophys. Res. Commun. 435 (4), 727-732 (2013)
   PUBMED   23692925
  REMARK    GeneRIF: In the present study we show that [a] NQO1 protein locates
            to cytosol, ER and mitochondria.
REFERENCE   3  (bases 1 to 1552)
  AUTHORS   Kummarapurugu,A.B., Fischer,B.M., Zheng,S., Milne,G.L., Ghio,A.J.,
```

Change region shown

Customize view

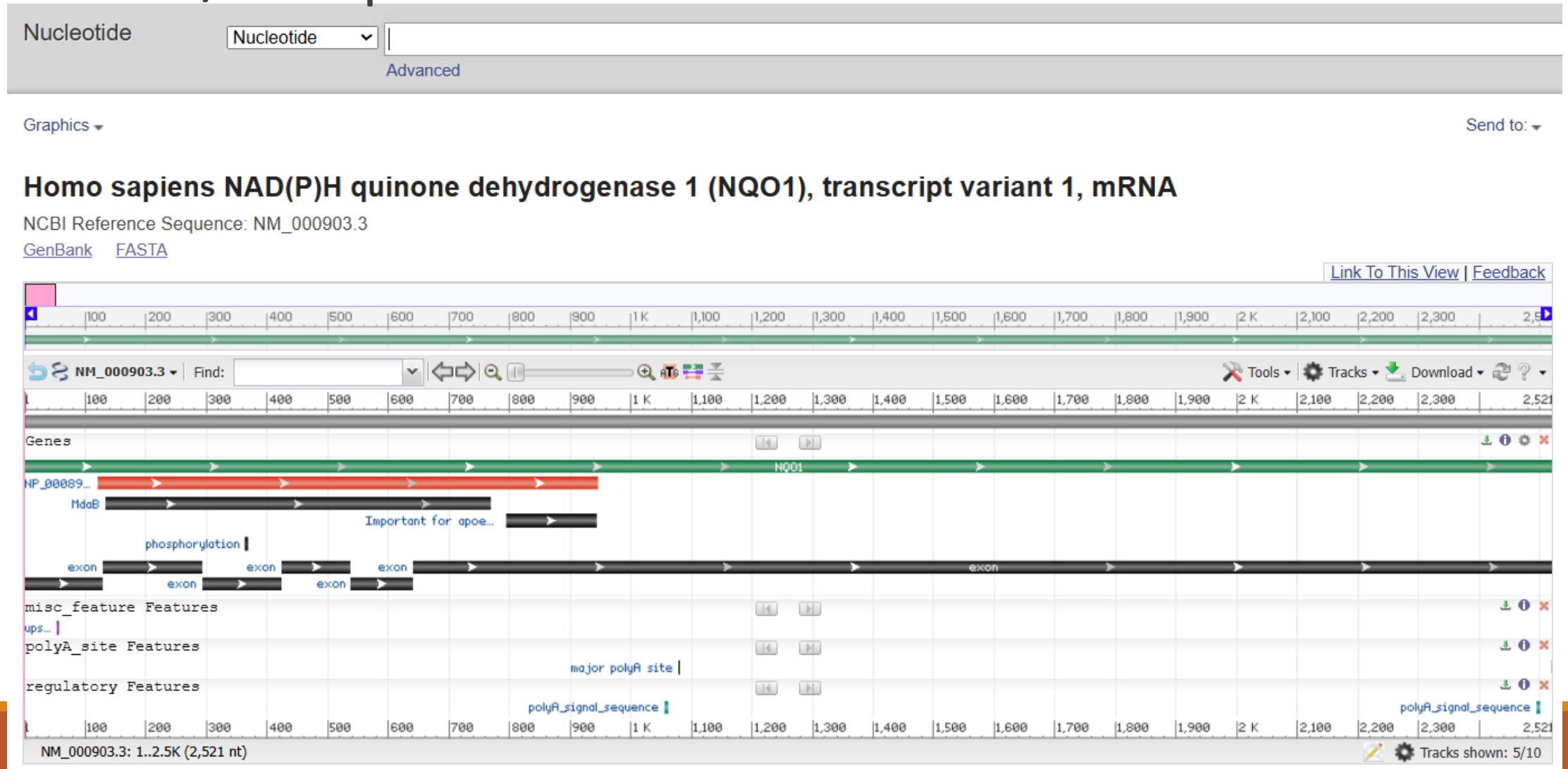
Analyze this sequence
Run BLAST
Pick Primers
Highlight Sequence Features
Find in this Sequence

Articles about the Nqo1 gene
Mitochondrial targeting of mouse NQO1 and CYP1B1 [Biochem Biophys Res Commun. 2013]
The protein level of PGC-1α, a key metabolic regulator, is controlled by NA [Mol Cell Biol. 2013]
NADPH:quinone oxidoreductase 1 regulates host susceptibility to ozone via iso [J Biol Chem. 2013]
See all...

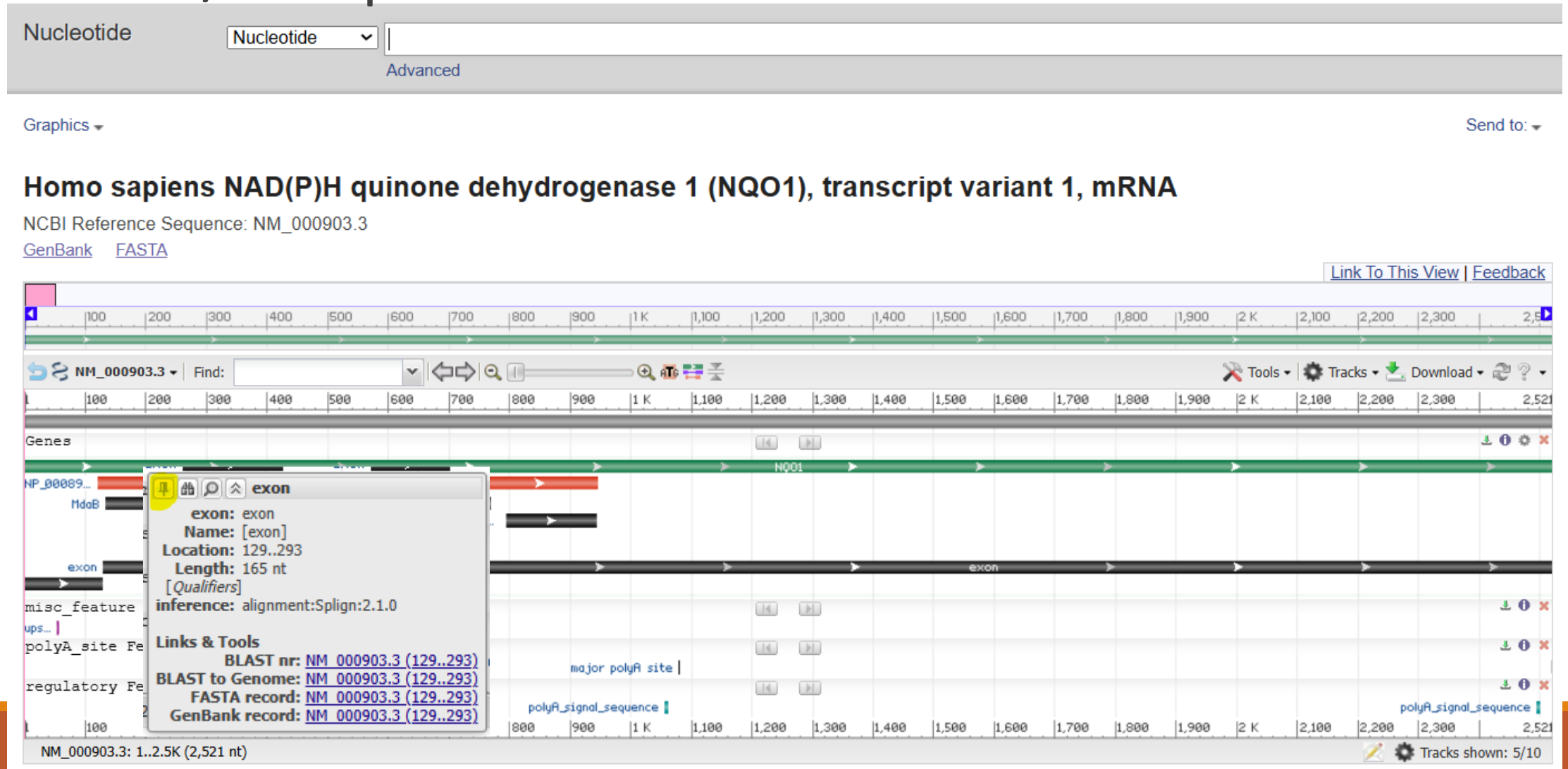
Pathways for the Nqo1 gene
Regulation of ornithine decarboxylase (ODC)
Metabolism of amino acids and derivatives
Metabolism
See all...

Reference sequence information

NCBI/Graphic



NCBI/Graphic



Ukázka možností NCBI/flatfile

organismus

mRNA, CDS, region

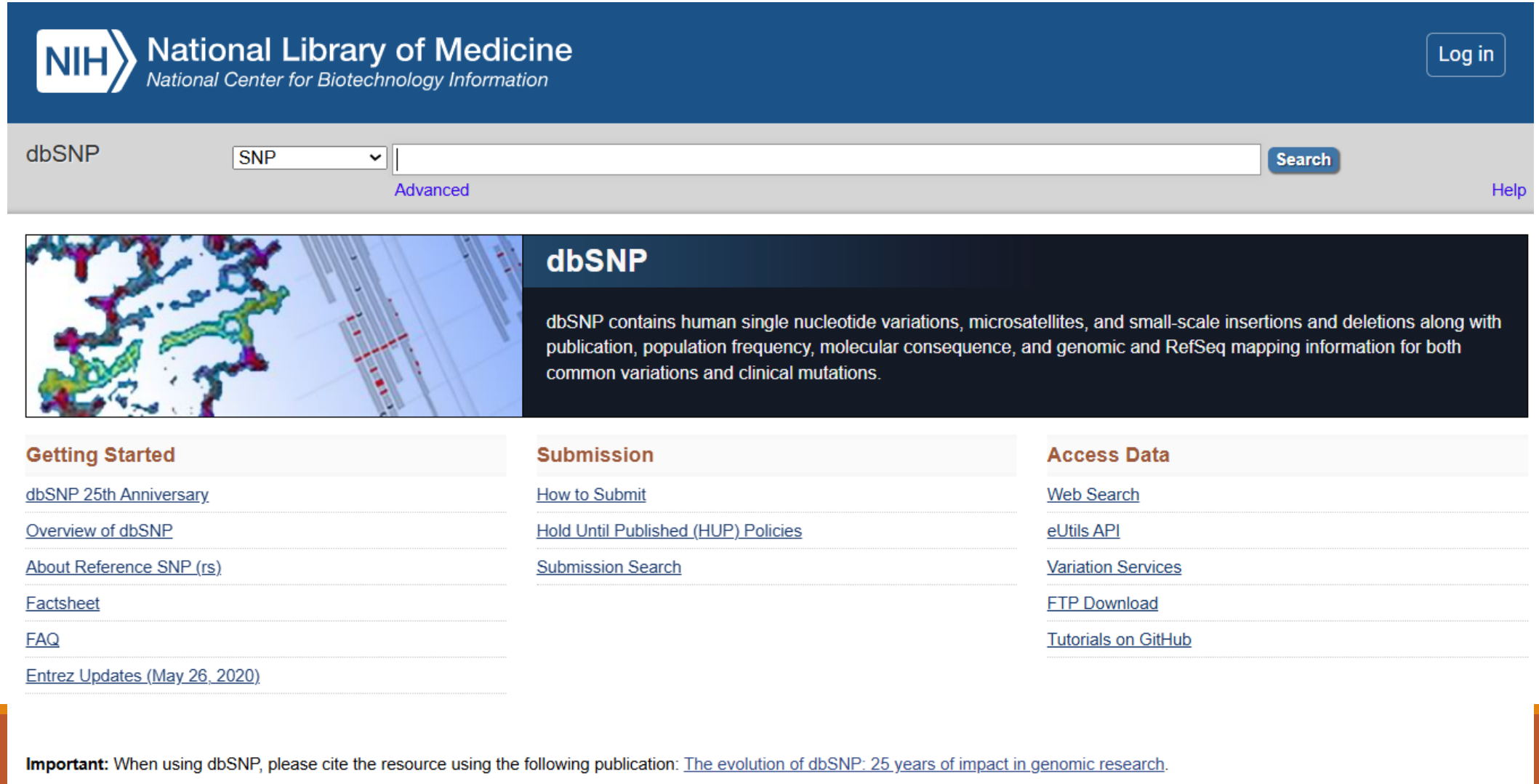
Find in sequence

Vyhledejte si „vlastní“ sekvenci

vyzkoušejte možnosti

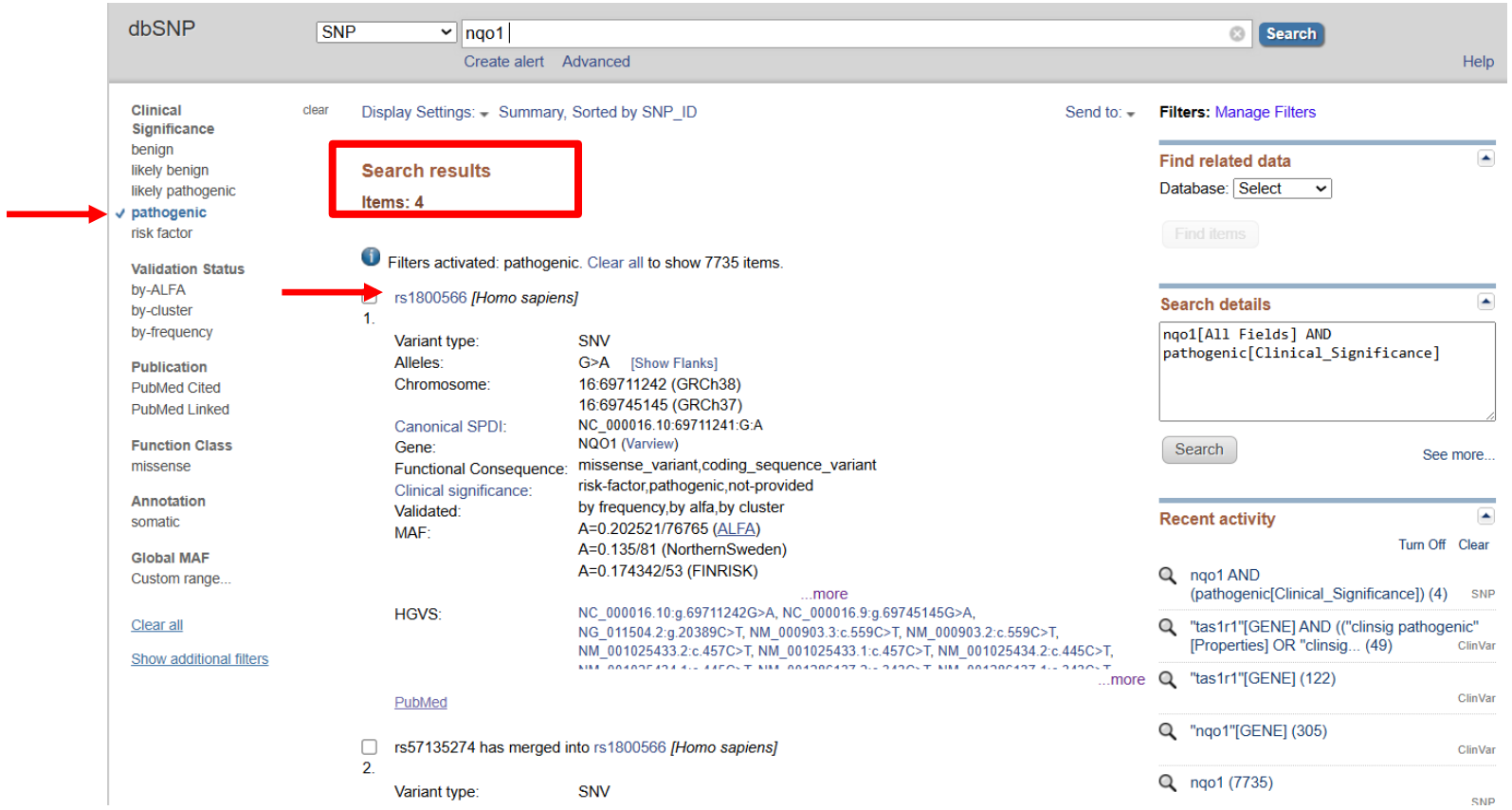
-podívejte se GRAPHIC zobrazení

dbSNP-varianty



dbSNP-varianty

dbSNP-varianty



dbSNP-varianty

dbSNP Short Genetic Variations

Search for terms | Search

Examples: rs268, BRCA1 and more | Advanced search

Welcome to the Reference SNP (rs) Report

All alleles are reported in the Forward orientation. Click on the Variant Details tab for details on Genomic Placement, Gene, and Amino Acid changes. HGVS names are in the HGVS tab.

Reference SNP (rs) Report

Download

rs1800566

Current Build 156
Released September 21, 2022

Organism	*Homo sapiens*	**Clinical Significance**	Reported in ClinVar
Position	chr16:69711242 (GRCh38.p14)	**Gene : Consequence**	NQO1 : Missense Variant
Alleles	G>A	**Publications**	172 citations LitVar 2 796
Variation Type	SNV Single Nucleotide Variation		
Frequency	A=0.202521 (76765/379048, ALFA) A=0.227477 (60211/264690, TOPMED) A=0.252052 (62641/248524, GnomAD_exome) (+ 28 more)	**Genomic View**	See rs on genome

dbSNP-variantv

Frequency | Variant Details | Clinical Significance | HGVS | Submissions | History | Publications | Flanks

ALFA Allele Frequency

The ALFA project provide aggregate allele frequency from dbGaP. More information is available on the project page including descriptions, data access, and terms of use.

Release Version: 20250407153717

Search:

Population	Group	Sample Size	Ref Allele	Alt Allele	Ref HMOZ	Alt HMOZ	HTRZ	HWEP
Total	Global	757684	G=0.798296	A=0.201704	0.639987	0.043395	0.316617	56
European	Sub	614020	G=0.811856	A=0.188144	0.659304	0.035592	0.305104	0
African	Sub	58734	G=0.81208	A=0.18792	0.659073	0.034903	0.306024	0
African Others	Sub	2098	G=0.8127	A=0.1873	0.658723	0.033365	0.307912	0
African American	Sub	56636	G=0.81206	A=0.18794	0.659086	0.03496	0.305954	0
Asian	Sub	17122	G=0.58223	A=0.41777	0.345871	0.181404	0.472725	4
East Asian	Sub	12886	G=0.57217	A=0.42783	0.334782	0.190439	0.474779	4
Other Asian	Sub	4236	G=0.6128	A=0.3872	0.379603	0.153919	0.466478	0

dbSNP-variantv

Frequency | Variant Details | **Clinical Significance** | HGVS | Submissions | History | Publications | Flanks

Allele: A (allele ID: 31848)

ClinVar Accession	Disease Names	Clinical Significance
RCV000018300.10	Benzene toxicity, susceptibility to	Risk-Factor
RCV000018301.10	Leukemia, post-chemotherapy, susceptibility to	Risk-Factor
RCV000018302.31	Breast cancer, post-chemotherapy poor survival in	Pathogenic
RCV000434090.9	Lung carcinoma	Not-Provided
RCV003974839.1	NQO1-related disorder	Benign

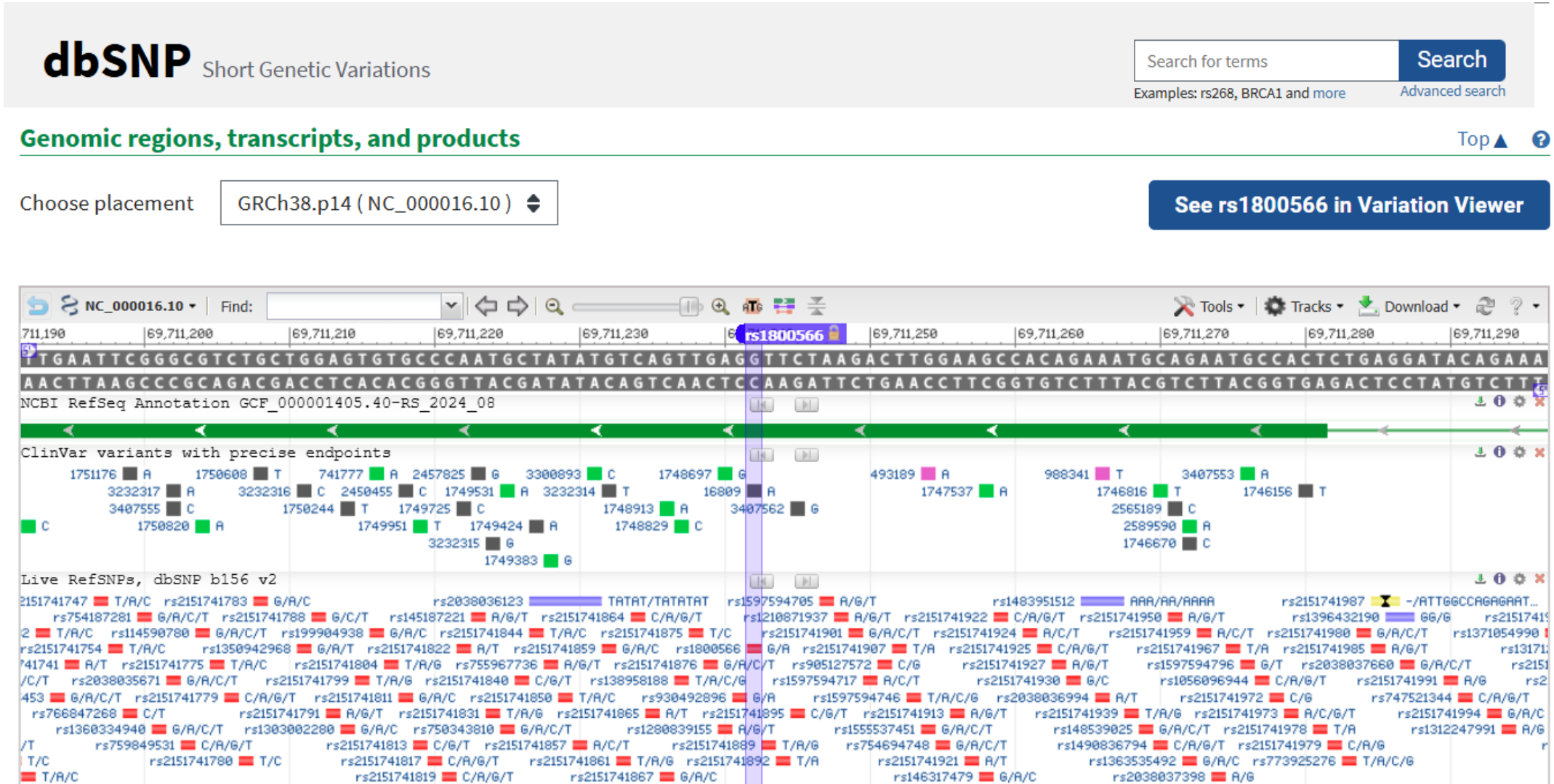
Genomic regions, transcripts, and products

Top▲

Choose placement GRCh38.p14 (NC_000016.10)

See rs1800566 in Variation Viewer

dbSNP-varianty



dbSNP-varianty

dbSNP Short Genetic Variations

Search for terms | Search

Examples: rs268, BRCA1 and more | Advanced search

Welcome to the Reference SNP (rs) Report

All alleles are reported in the Forward orientation. Click on the Variant Details tab for details on Genomic Placement, Gene, and Amino Acid changes. HGVS names are in the HGVS tab.

Reference SNP (rs) Report

Download

rs1800566

Current Build 156

Released September 21, 2022

Organism	*Homo sapiens*	**Clinical Significance**	Reported in ClinVar
Position	chr16:69711242 (GRCh38.p14)	**Gene : Consequence**	NQO1 : Missense Variant
Alleles	G>A	**Publications**	172 citations
Variation Type	SNV Single Nucleotide Variation		LitVar2 796
Frequency	A=0.202521 (76765/379048, ALFA) A=0.227477 (60211/264690, TOPMED) A=0.252052 (62641/248524, GnomAD_exome) (+ 28 more)	**Genomic View**	See rs on genome

ClinVar
Genomic variation as it relates to human health
((31848[AlleleID]))
Search ClinVar
Advanced search
About
Access
Submit
Stats
FTP
Help
NM_000903.3(NQO1):c.559C>T (p.Pro187Ser)
Cite
Follow
Print
Download
Germline
Classification
Conflicting classifications of pathogenicity; risk factor
Pathogenic(1); Benign(1)
4 out of 4 submissions contributed to this classification
Germline
Classification
Uncertain significance
1 out of 3 submissions contributed to this classification
Variant Details
Identifiers:
NM_000903.3(NQO1):c.559C>T (p.Pro187Ser)
Variation ID: 16809
Accession: VCV000016809.5

ClinVar Genomic variation as it relates to human health

((31848[AlleleID])) Search ClinVar

Protein change:	P187S, P115S, P149S, P153S
Other names:	-
Canonical SPDI:	NC_000016.10:69711241:G:A
Functional consequence :	-
Global minor allele frequency (GMAF) :	0.28894 (A)
Allele frequency :	The Genome Aggregation Database (gnomAD) 0.20869
	Trans-Omics for Precision Medicine (TOPMed) 0.22748
	Exome Aggregation Consortium (ExAC) 0.24516
	The Genome Aggregation Database (gnomAD), exomes 0.25205
	1000 Genomes Project 30x 0.28232
	1000 Genomes Project 0.28894
Links :	PharmGKB Clinical Annotation: 1183491004 ClinGen: CA126904 UniProtKB: P15559#VAR_008384 OMIM: 125860.0001 dbSNP: rs1800566 VarSome

OMIM®

An Online Catalog of Human Genes and Genetic Disorders

Updated March 7th, 2025

Search OMIM for clinical features, phenotypes, genes, and more...

Advanced Search : OMIM, Clinical Synopses, Gene Map | Search History

Need help? : Example Searches, OMIM Search Help, OMIM Video Tutorials

Mirror site : https://mirror.omim.org

OMIM is supported by a grant from NHGRI, licensing fees, and generous contributions from people like you.

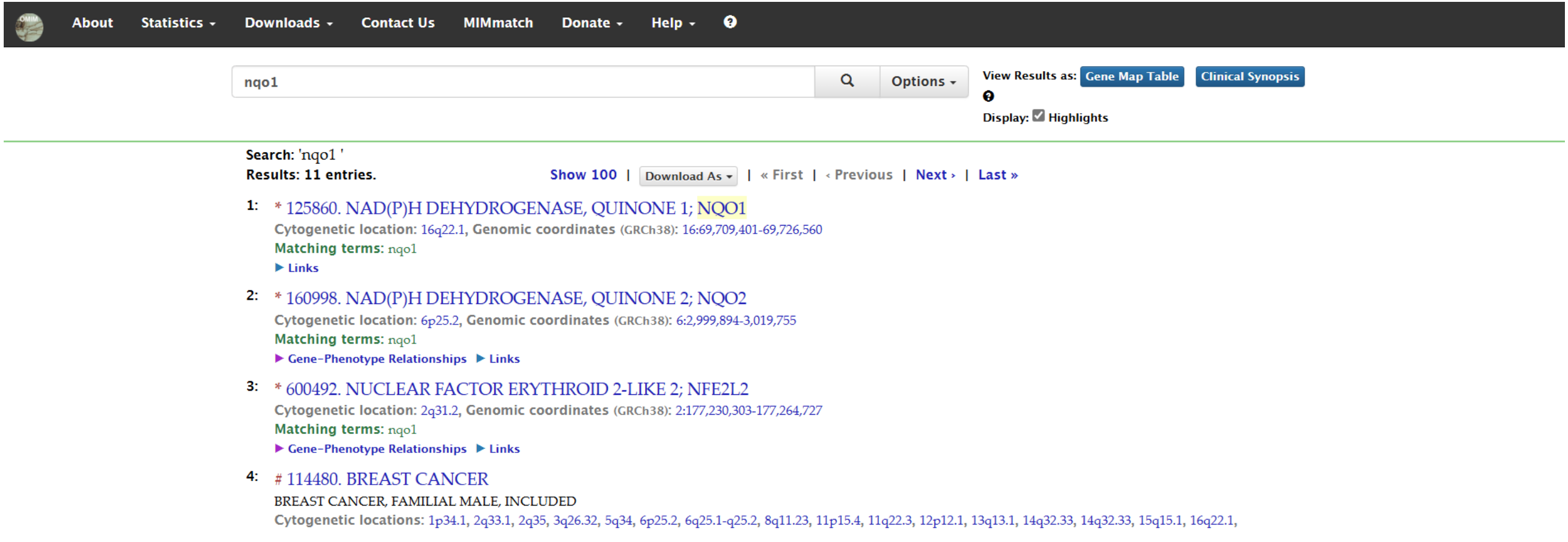
Make a donation!

McKUSICK-NATHANS Department of Genetic Medicine

JOHNS HOPKINS MEDICINE

IRDiRC RECOGNIZED RESOURCES

OMIM

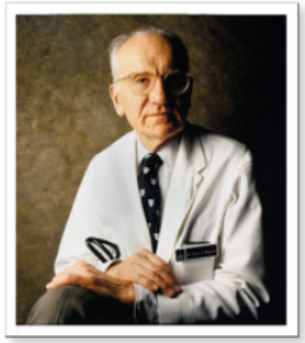

About
Statistics
Downloads
Contact Us
MIMmatch
Donate
Help
nqo1
Options
View Results as:
Gene Map Table
Clinical Synopsis
Display:
Highlights
Search: 'nqo1 '
Results: 11 entries.
Show 100 | Download As | « First | ‹ Previous | Next › | Last »
1: * 125860. NAD(P)H DEHYDROGENASE, QUINONE 1; NQO1
Cytogenetic location: 16q22.1, Genomic coordinates (GRCh38): 16:69,709,401-69,726,560
Matching terms: nqo1
▶ Links
2: * 160998. NAD(P)H DEHYDROGENASE, QUINONE 2; NQO2
Cytogenetic location: 6p25.2, Genomic coordinates (GRCh38): 6:2,999,894-3,019,755
Matching terms: nqo1
▶ Gene-Phenotype Relationships ▶ Links
3: * 600492. NUCLEAR FACTOR ERYTHROID 2-LIKE 2; NFE2L2
Cytogenetic location: 2q31.2, Genomic coordinates (GRCh38): 2:177,230,303-177,264,727
Matching terms: nqo1
▶ Gene-Phenotype Relationships ▶ Links
4: # 114480. BREAST CANCER
BREAST CANCER, FAMILIAL MALE, INCLUDED
Cytogenetic locations: 1p34.1, 2q33.1, 2q35, 3q26.32, 5q34, 6p25.2, 6q25.1-q25.2, 8q11.23, 11p15.4, 11q22.3, 12p12.1, 13q13.1, 14q32.33, 14q32.33, 15q15.1, 16q22.1,

OMIM

About Statistics Downloads Contact Us MIMmatch Donate Help

Search OMIM... Options

OMIM® – Online Mendelian Inheritance in Man®



Welcome to OMIM®, Online Mendelian Inheritance in Man®. OMIM is a comprehensive, authoritative compendium of human genes and genetic phenotypes that is freely available and updated daily. The full-text, referenced overviews in OMIM contain information on all known mendelian disorders and over 16,000 genes. OMIM focuses on the relationship between phenotype and genotype. It is updated daily, and the entries contain copious links to other genetics resources.

This database was initiated in the early 1960s by Dr. Victor A. McKusick as a catalog of mendelian traits and disorders, entitled Mendelian Inheritance in Man (MIM). Twelve book editions of MIM were published between 1966 and 1998. The online version, OMIM, was created in 1985 by a collaboration between the National Library of Medicine and the William H. Welch Medical Library at Johns Hopkins. It was made generally available on the internet starting in 1987. In 1995, OMIM was developed for the World Wide Web by NCBI, the National Center for Biotechnology Information.

OMIM is authored and edited at the McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, under the direction of Dr. Ada Hamosh.

NLM's Profiles in Science -- The McKusick Papers

OMIM

*125860
Table of Contents

- Title
- Text
  - Description
  - Cloning and Expression
  - Gene Function
  - Gene Structure
  - Mapping
  - Molecular Genetics
  - Animal Model
- Allelic Variants
  - Table View
- References
- Contributors
- Creation Date
- Edit History

▼ ALLELIC VARIANTS (1 Selected Example):

Table View | ClinVar

.0001 RECLASSIFIED - NQO1 POLYMORPHISM

NQO1, PRO187SER • rs1800566 ▾ RCV000018300...

This variant, formerly titled BENZENE TOXICITY, SUSCEPTIBILITY TO, with included titles of LEUKEMIA, POST-CHEMOTHERAPY, SUSCEPTIBILITY TO and BREAST CANCER, POST-CHEMOTHERAPY POOR SURVIVAL IN, has been reclassified as a polymorphism. Hamosh (2025) noted that the P187S allele was present in the gnomAD database in 69,078 of 279,866 alleles, in 9,942 homozygotes, for an allele frequency of 0.2468.

Traver et al. (1992, 1997) characterized a polymorphism in NQO1, a C-to-T substitution at position 609 of NQO1 cDNA, which codes for a proline-to-serine change at residue 187 (P187S). In cells with a T/T genotype, NQO1 activity was not detected, and lack of activity corresponded to a lack of NQO1 protein.

Kelsey et al. (1997) found that the prevalence of the T/T genotype varied among ethnic groups from 4 to 20%, with the highest prevalence occurring in Asian populations.

Susceptibility to Benzene Toxicity

Rothman et al. (1997) reported a case-control study of benzene-exposed workers in China showing an increased risk of hematotoxicity in individuals with the T/T NQO1 genotype. A protective role for NQO1 against benzene-derived quinones in bone marrow in situ required reconciling with the observation that freshly isolated bone marrow mononuclear and progenitor cells failed to express

▼ External Links
- ▶ Genome
- ▶ DNA
- ▶ Protein
- ▶ Gene Info
- ▶ Clinical Resources
- ▼ Variation
  - ClinVar
  - gnomAD
  - GWAS Catalog
  - GWAS Central
  - HGMD
  - NHLBI EVS
  - PharmGKB
- ▶ Animal Models
- ▶ Cell Lines
- ▶ Cellular Pathways

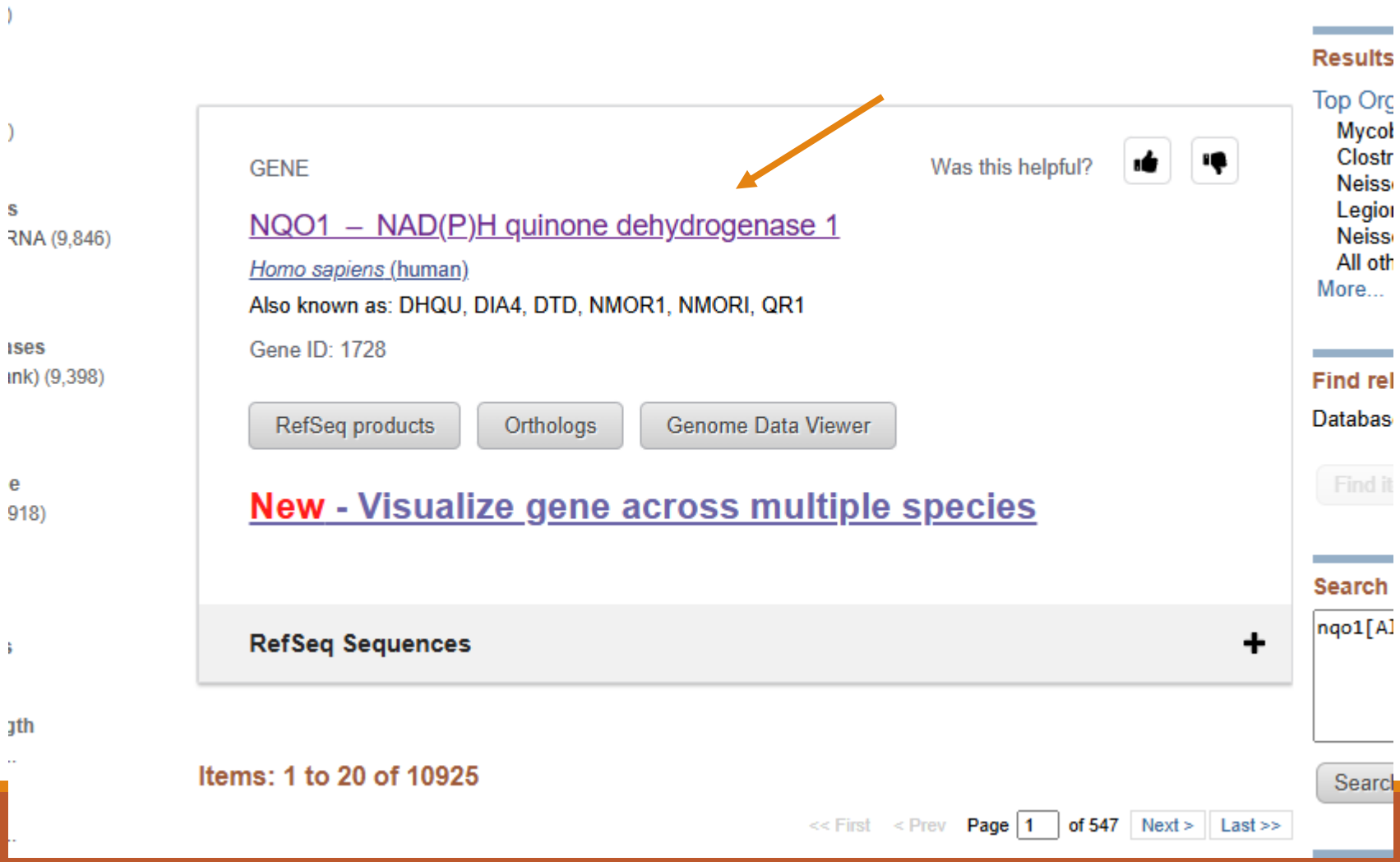
Vyzkoušejte

- Podívejte se, jestli váš gen má nějaké varianty v databázi dbSNP, jsou některé patologické?

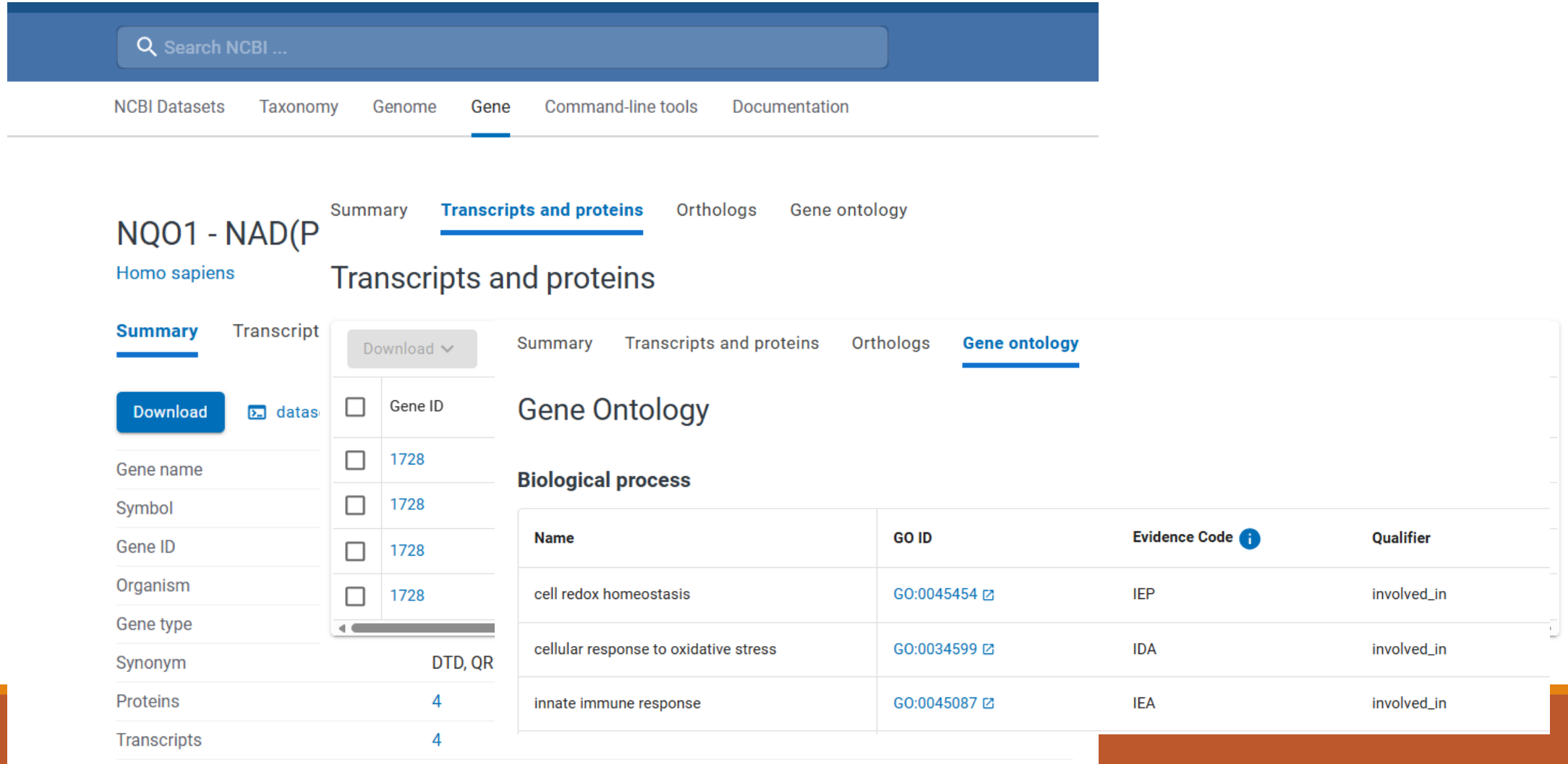
  - ve vztahu k čemu?

- Podívejte se na váš gen do databáze OMIM.

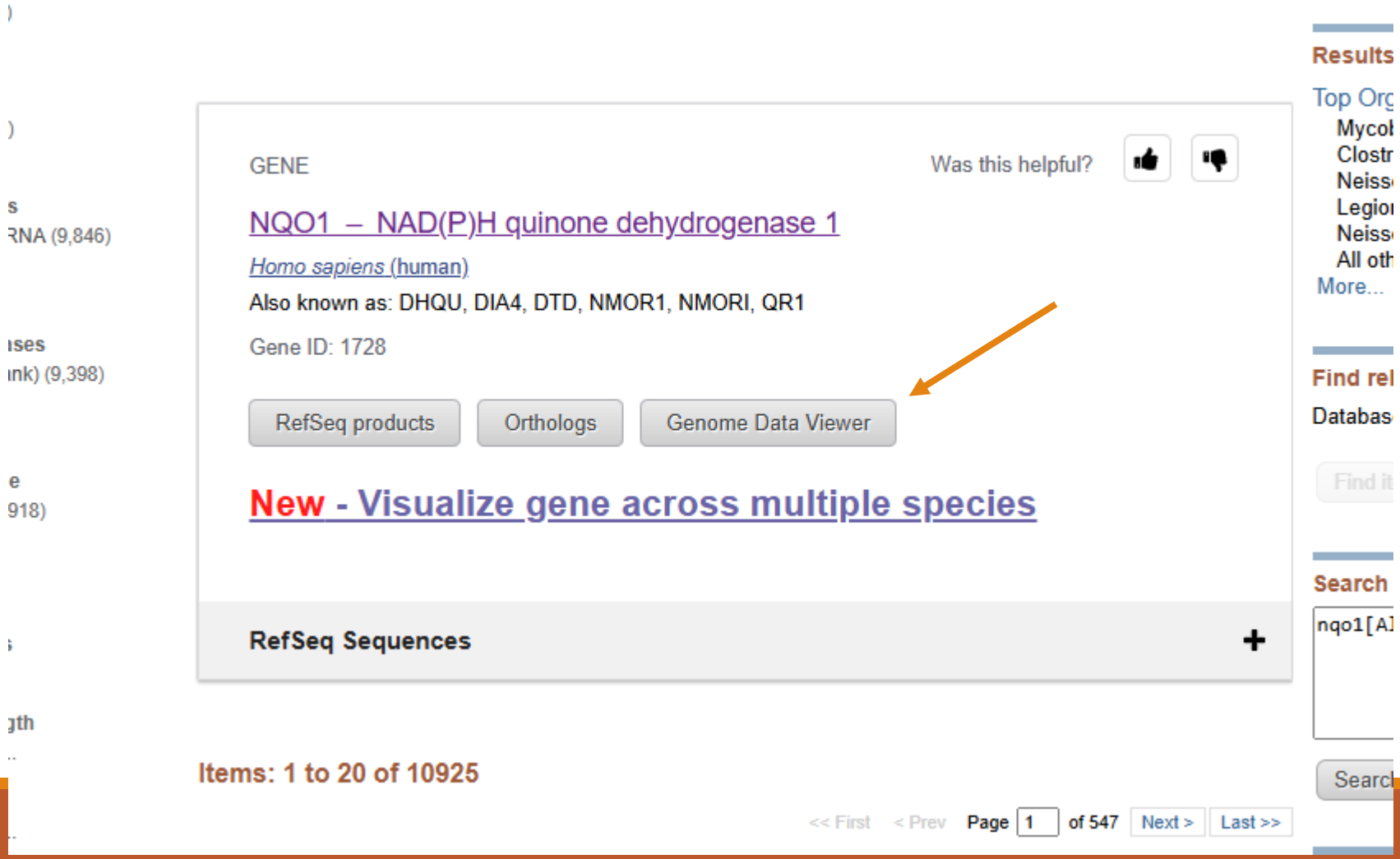
Genová databáze



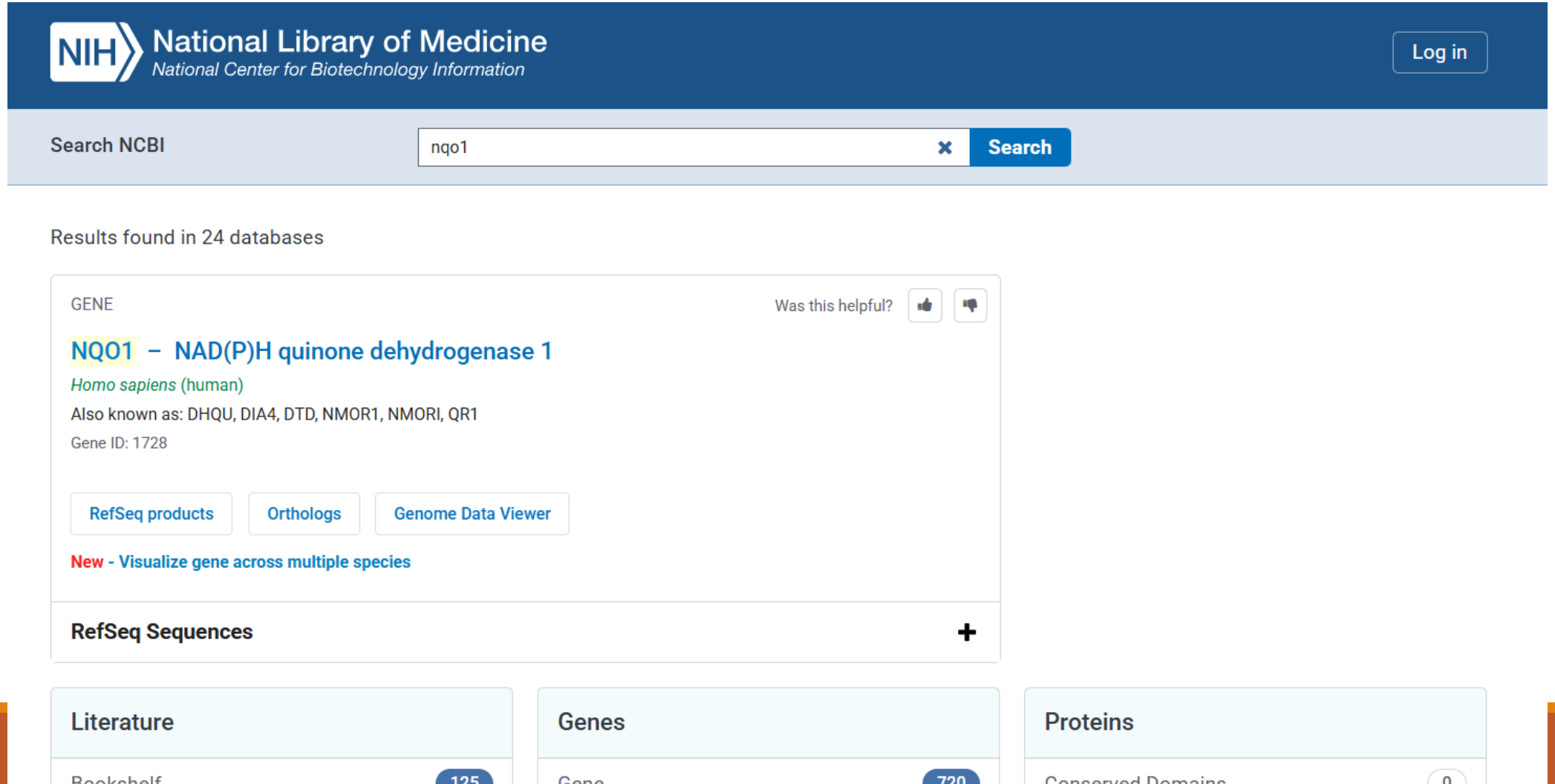
Genová databáze (nové zobrazení)



Genová databáze

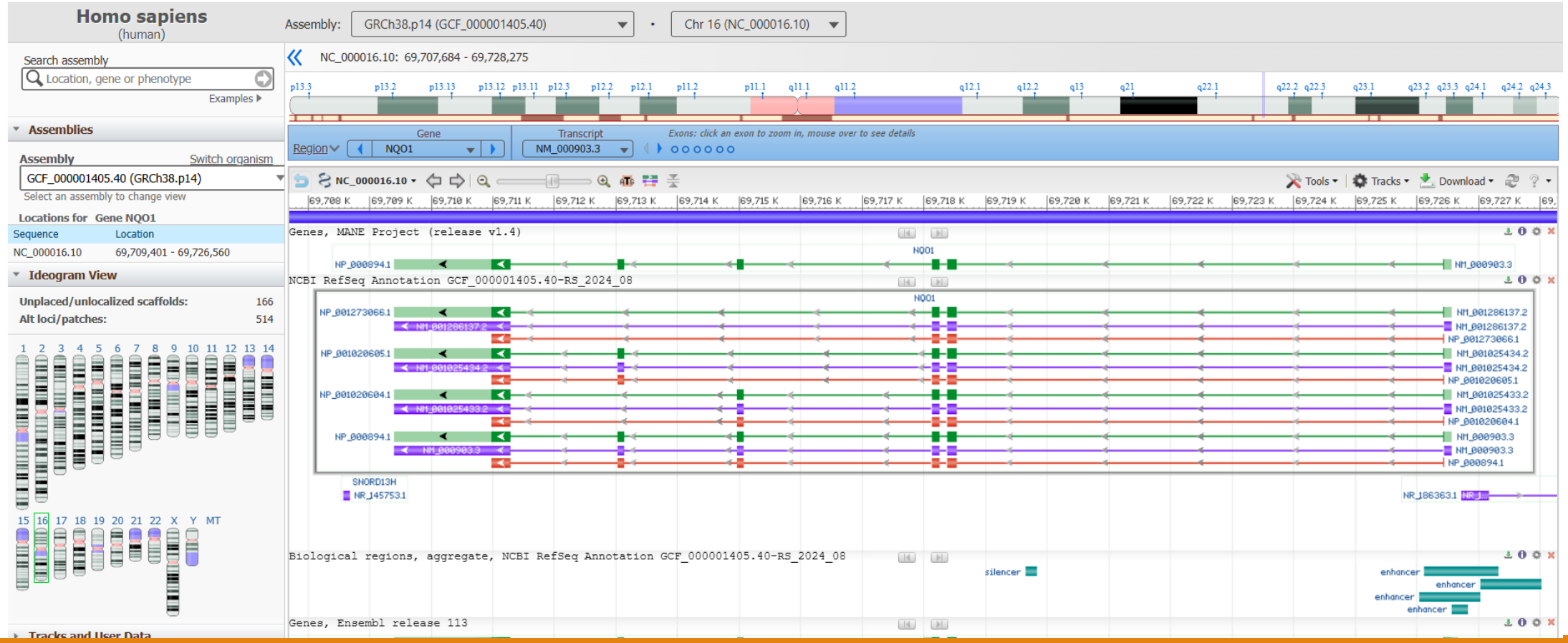


Genome Data Viewer

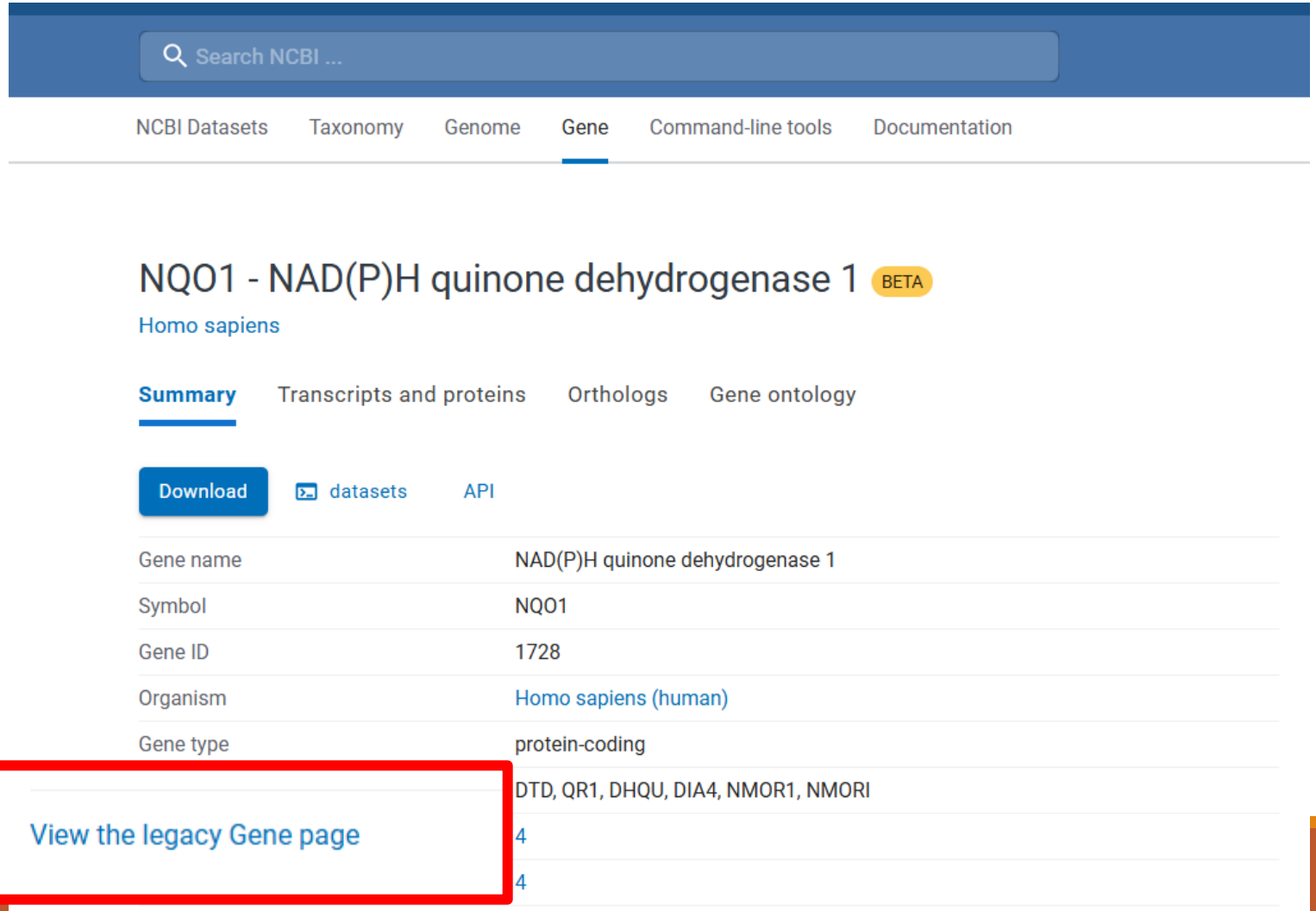


Genome Data Viewer

Genome Data Viewer



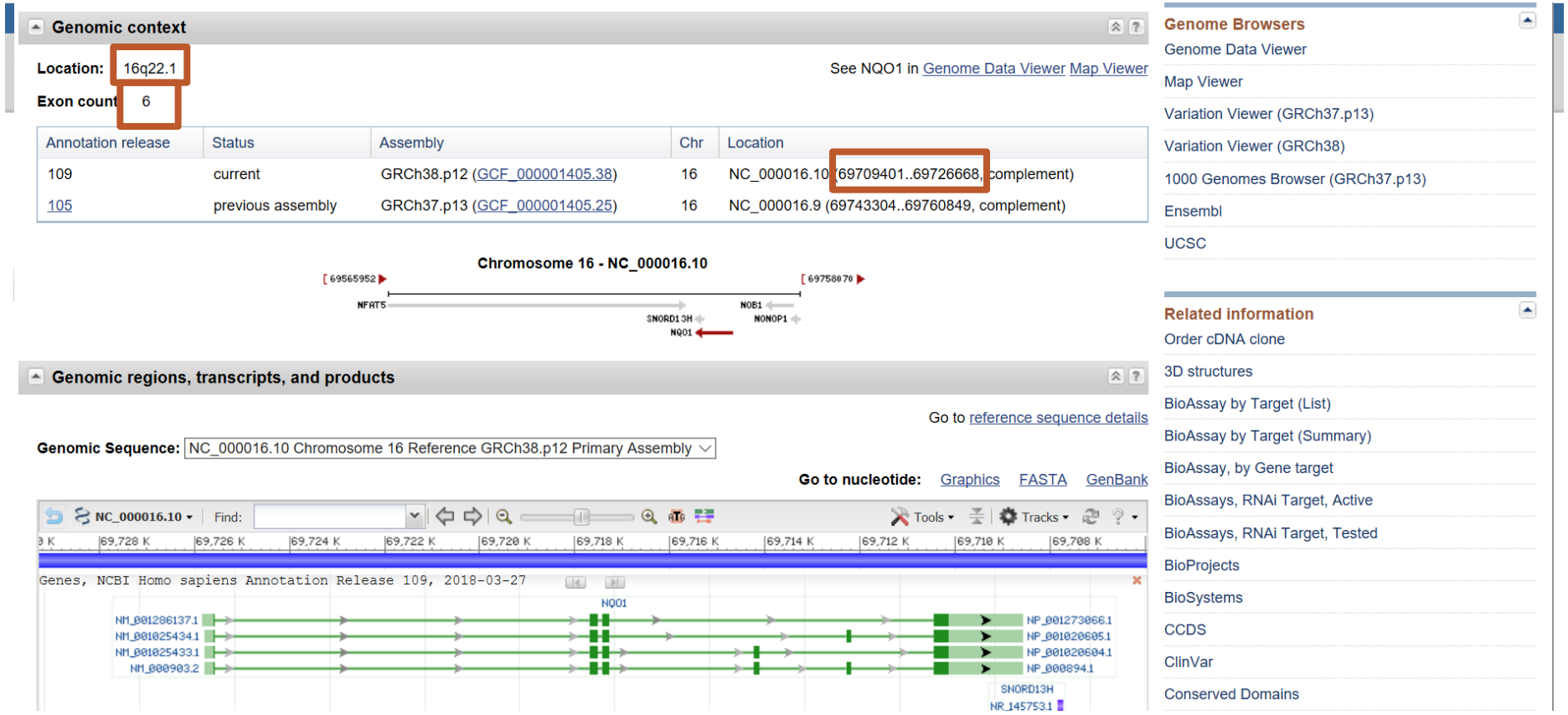
Genová databáze (nové zobrazení)



NCBI/Gene

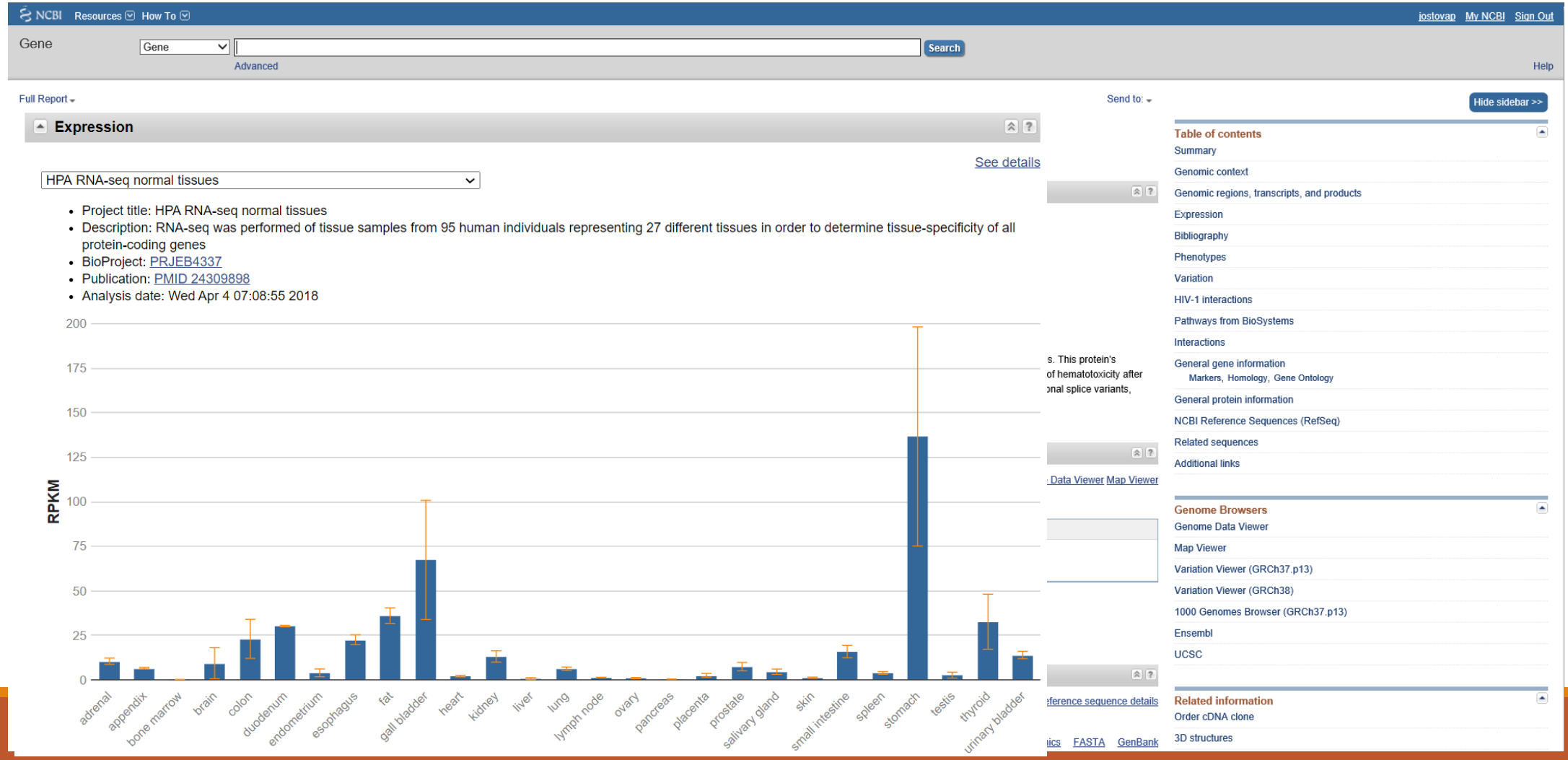
NCBI Resources How To Sign in to NCBI
Gene
Gene
Search
Advanced
Help
Full Report
Send to:
Hide sidebar >>
NQO1 NAD(P)H quinone dehydrogenase 1 [Homo sapiens (human)]
Gene ID: 1728, updated on 29-Mar-2018
Summary
Official Symbol NQO1 provided by HGNC
Official Full Name NAD(P)H quinone dehydrogenase 1 provided by HGNC
Primary source HGNC:HGNC:2874
See related Ensembl:ENSG00000181019 MIM:125860; Vega:OTTHUMG00000137575
Gene type protein coding
RefSeq status REVIEWED
Organism Homo sapiens
Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo
Also known as DTD; QR1; DHQU; DIA4; NMOR1; NMORI
Summary This gene is a member of the NAD(P)H dehydrogenase (quinone) family and encodes a cytoplasmic 2-electron reductase. This FAD-binding protein forms homodimers and reduces quinones to hydroquinones. This protein's enzymatic activity prevents the one electron reduction of quinones that results in the production of radical species. Mutations in this gene have been associated with tardive dyskinesia (TD), an increased risk of hematotoxicity after exposure to benzene, and susceptibility to various forms of cancer. Altered expression of this protein has been seen in many tumors and is also associated with Alzheimer's disease (AD). Alternate transcriptional splice variants, encoding different isoforms, have been characterized. [provided by RefSeq, Jul 2008]
Expression Biased expression in stomach (RPKM 136.7), gall bladder (RPKM 67.3) and 11 other tissues See more
Table of contents
Summary
Genomic context
Genomic regions, transcripts, and products
Expression
Bibliography
Phenotypes
Variation
HIV-1 interactions
Pathways from BioSystems
Interactions
General gene information
Markers, Homology, Gene Ontology
General protein information
NCBI Reference Sequences (RefSeq)
Related sequences
Additional links

NCBI/Gene (obsáhlejší verze)



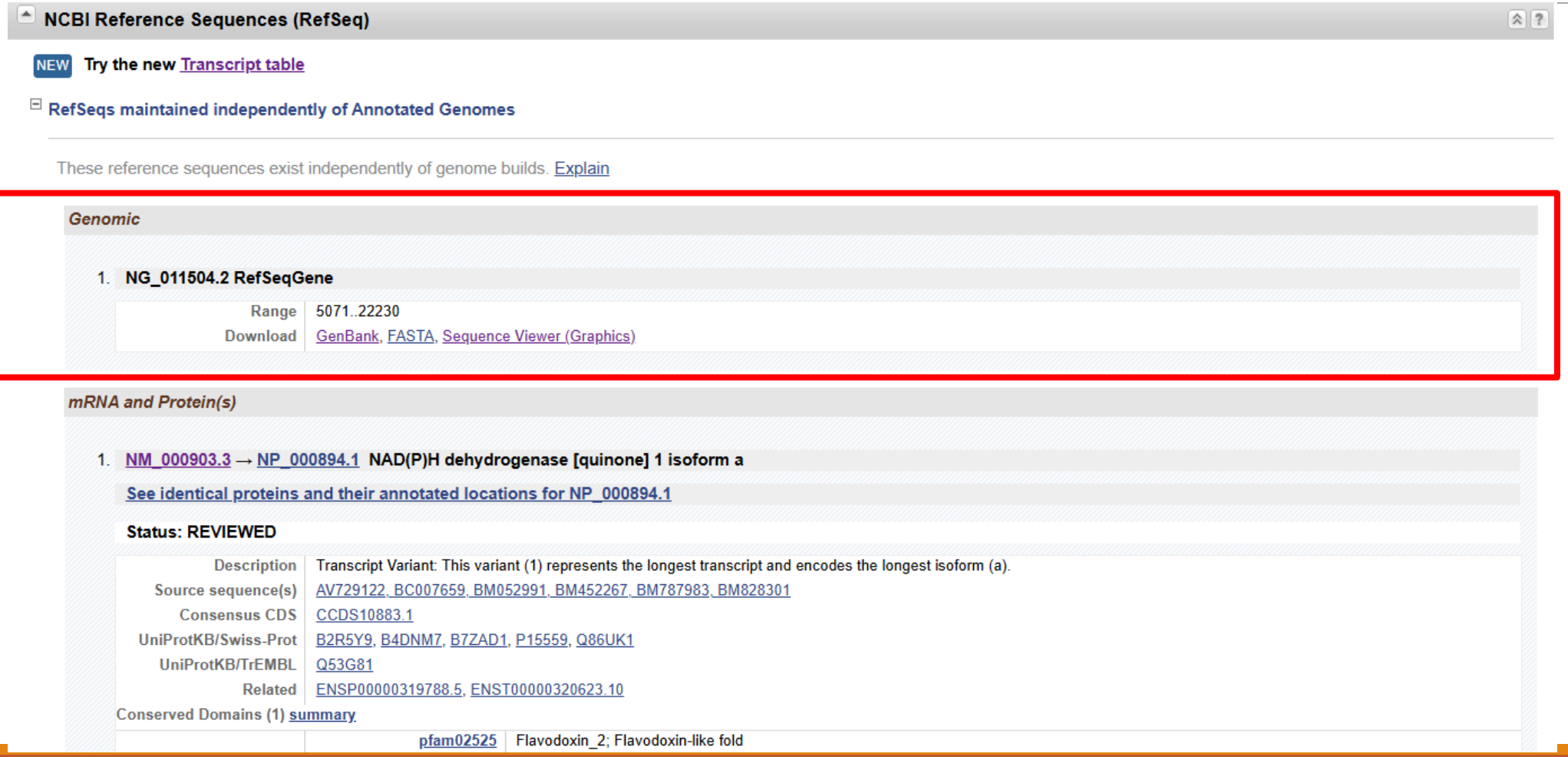
NCBI/Gene-pseudogeny?

NCBI/Gene - exprese?



NCBI/Gene – sekvence celého genu

NCBI/Gene – sekvence celého genu NG_



NCBI Reference Sequences (RefSeq)

NEW Try the new Transcript table

RefSeqs maintained independently of Annotated Genomes

These reference sequences exist independently of genome builds. Explain

Genomic

1. **NG_011504.2 RefSeqGene**

Range	5071..22230
Download	GenBank, FASTA, Sequence Viewer (Graphics)

mRNA and Protein(s)

1. NM_000903.3 → NP_000894.1 **NAD(P)H dehydrogenase [quinone] 1 isoform a**

See identical proteins and their annotated locations for NP_000894.1

Status: REVIEWED

Description	Transcript Variant: This variant (1) represents the longest transcript and encodes the longest isoform (a).
Source sequence(s)	AV729122, BC007659, BM052991, BM452267, BM787983, BM828301
Consensus CDS	CCDS10883.1
UniProtKB/Swiss-Prot	B2R5Y9, B4DNM7, B7ZAD1, P15559, Q86UK1
UniProtKB/TrEMBL	Q53G81
Related	ENSP00000319788.5, ENST00000320623.10

Conserved Domains (1) summary

pfam02525	Flavodoxin_2; Flavodoxin-like fold

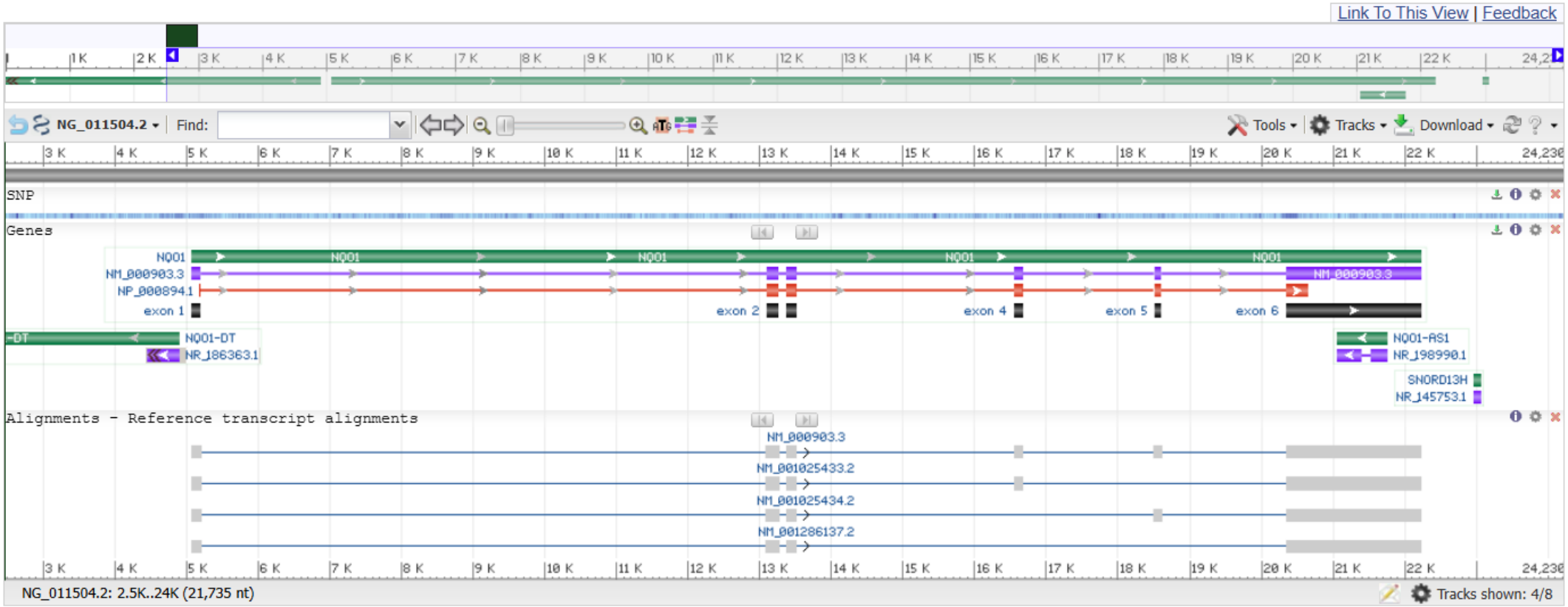
NCBI/Gene – sekvence celého genu / **Graphics**

Graphics ▾ Send to: ▾

Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1), RefSeqGene on chromosome 16

NCBI Reference Sequence: NG_011504.2

GenBank FASTA



Někdo nemá NG_...?

NCBI/Gene – sekvence celého genu

NCBI Reference Sequences (RefSeq)

NEW **Try the new Transcript table**

RefSeqs maintained independently of Annotated Genomes

These reference sequences exist independently of genome builds. Explain

Genomic

1. **NG_011504.2 RefSeqGene**

Range	5071..22230
Download	GenBank, FASTA, Sequence Viewer (Graphics)

mRNA and Protein(s)

1. NM_000903.3 → NP_000894.1 **NAD(P)H dehydrogenase [quinone] 1 isoform a**

See identical proteins and their annotated locations for NP_000894.1

Status: REVIEWED

Reference GRCh38.p14 Primary Assembly

Genomic

1. **NC_000016.10 Reference GRCh38.p14 Primary Assembly**

Range	69709401..69726560 complement
Download	GenBank, FASTA, Sequence Viewer (Graphics)

NCBI/Gene – sekvence celého genu z chromosomu NC_...

GenBank ▾ Send to: ▾

Homo sapiens chromosome 16, GRCh38.p14 Primary Assembly

NCBI Reference Sequence: NC_000016.10

FASTA Graphics

Go to:

```
LOCUS       NC_000016              17160 bp    DNA     linear   CON 06-APR-2022
DEFINITION  Homo sapiens chromosome 16, GRCh38.p14 Primary Assembly.
ACCESSION   NC_000016 REGION: 69709401..69726560
VERSION     NC_000016.10
DBLINK      BioProject: PRJNA168
            Assembly: GCF_000001405.40
KEYWORDS    RefSeq.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 17160)
  AUTHORS   Martin,J., Han,C., Gordon,L.A., Terry,A., Prabhakar,S., She,X.,
            Xie,G., Hellsten,U., Chan,Y.M., Altherr,M., Couronne,O., Aerts,A.,
            Bajorek,E., Black,S., Blumer,H., Branscomb,E., Brown,N.C.,
            Bruno,W.J., Buckingham,J.M., Callen,D.F., Campbell,C.S.,
            Campbell,M.L., Campbell,E.W., Caoile,C., Challacombe,J.F.,
            Chasteen,L.A., Chertkov,O., Chi,H.C., Christensen,M., Clark,L.M.,
            Cohn,J.D., Denys,M., Detter,J.C., Dickson,M.,
            Dimitrijevic-Bussod,M., Escobar,J., Fawcett,J.J., Flowers,D.,
            Fotopulos,D., Glavina,T., Gomez,M., Gonzales,E., Goodstein,D.,
```

Change region shown

- ○ Whole sequence (abbreviated view)
- ◉ Selected region

from: 69709401 to: 69726560

Update View

Customize view

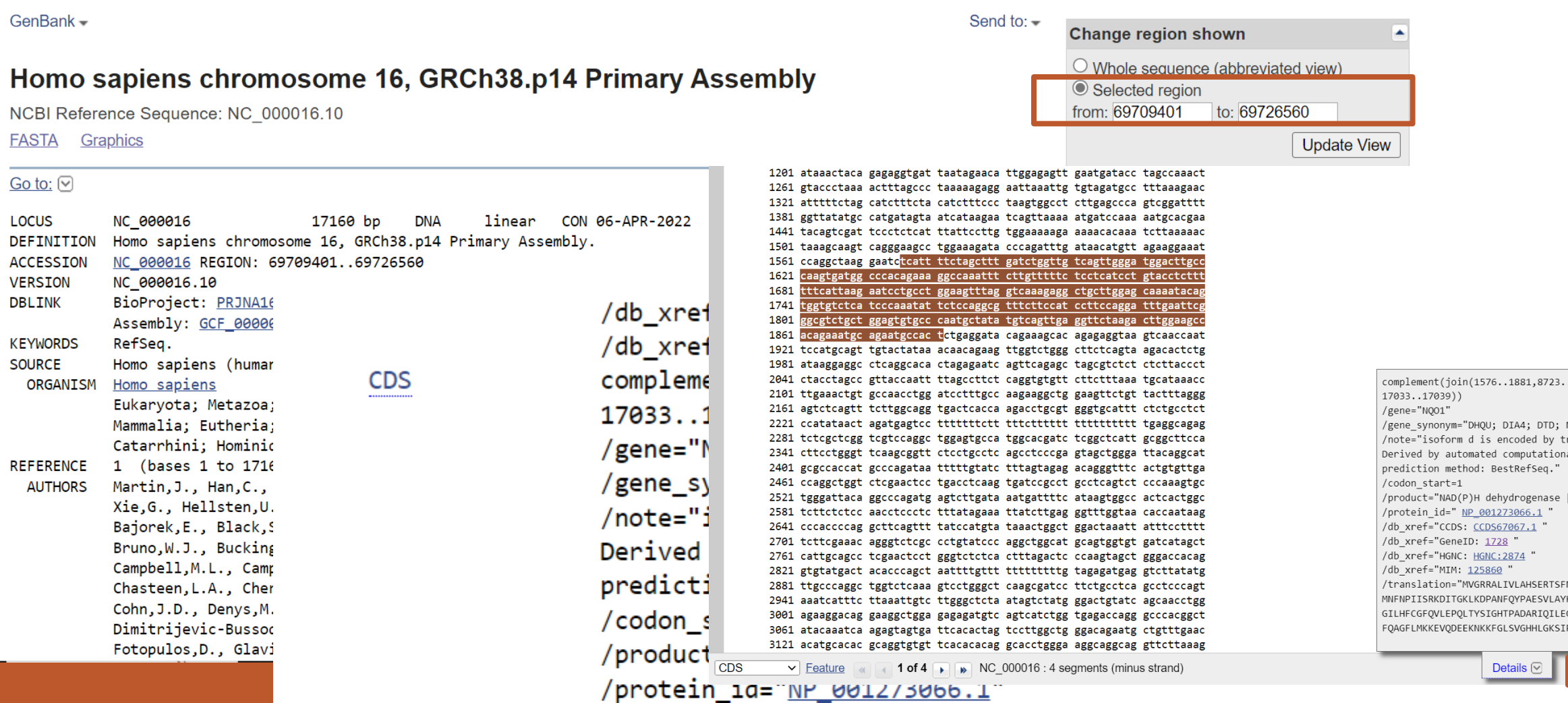
Analyze this sequence

- Run BLAST
- Pick Primers
- Highlight Sequence Features
- Find in this Sequence

Related information

- Assembly
- BioProject
- Protein
- PubMed
- Taxonomy

NCBI/Gene – sekvence celého genu z chromosomu



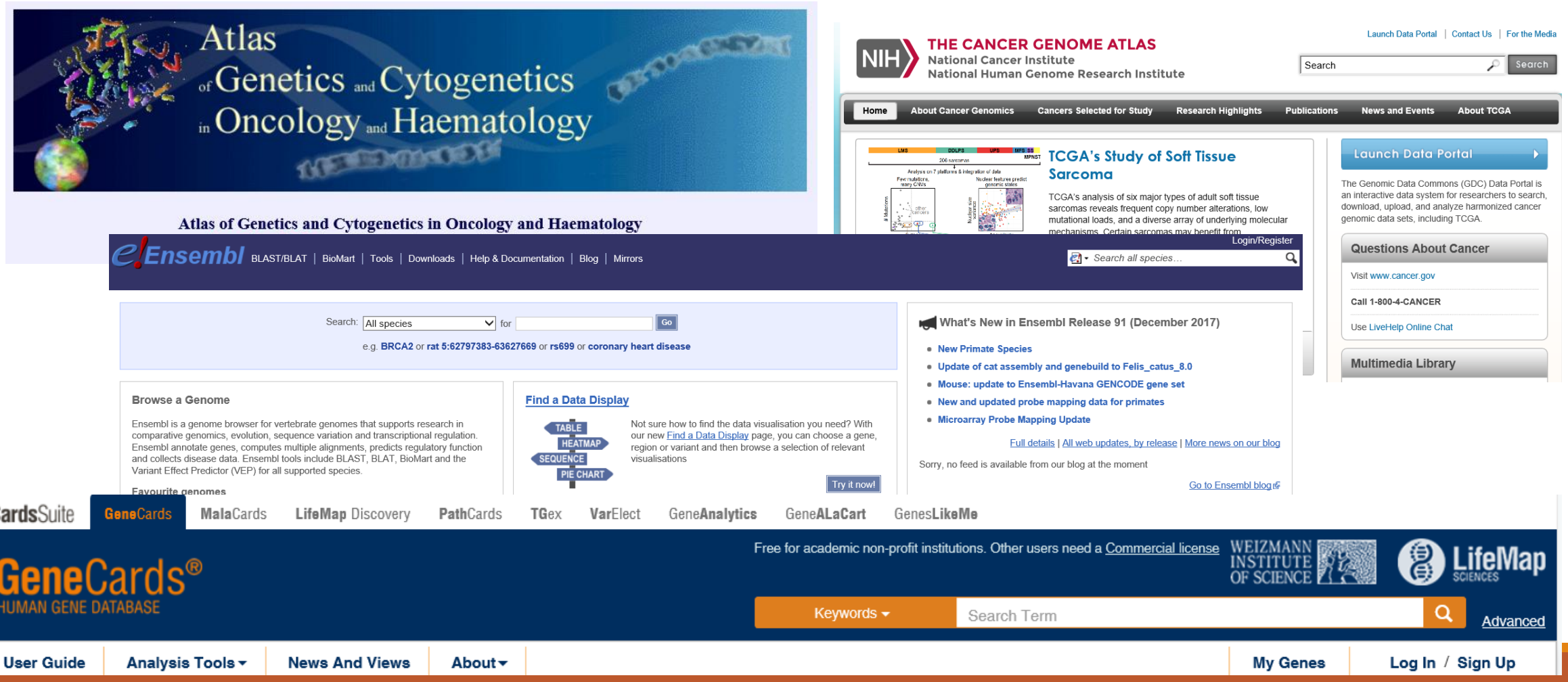
praktická ukázka GENE

Vyhledejte si „vlastní“ gen

prohlédněte si databázi GENE

Vyzkoušejte různé „linky“

Specializované databáze (sekundární)



Specializované databáze (sekundární)

GeneCards®: The Human Gene Database

GeneCards is a searchable, integrative database that provides comprehensive, user-friendly information on all annotated and predicted human genes. It automatically integrates gene-centric data from ~125 web sources, including genomic, transcriptomic, proteomic, genetic, clinical and functional information.

Cvičení 1-**Vyhledávání v molekulárně genetických databázích**
(prof.Beránek)

Vyhledejte si „vlastní“ gen
v jiných databázích

Vyzkoušejte různé „linky“

„Bioinformatika nukleových kyselin I“

Vyhledávání NK sekvencí, **specializované databáze**

Analýza vlastností sekvencí-složení, „reverse complement“, identifikace restrikčních míst (Palindromy)

Práce s kódující DNA=práce s proteiny / překlad DNA sekvence-otvírání čtecího rámce

Návrh primerů pro PCR, rt-PCR

Předpověď sekundárních struktur

Porovnávání sekvencí

Vyhledání SNPs

„čtení“ sekvenačních dat a spojování fragmentů

Vyhledávání hladin expresí jednotlivých genů

mikroRNA

Celé genomy

….

Charakteristiky sekvencí

Filter DNA

DNA stats

Range Extraktor DNA

Reverse complement

SMS

Format Conversion
-Combine FASTA
-EMBL to FASTA
-EMBL Feature Extractor
-EMBL Trans Extractor
-Filter DNA
-Filter Protein
-GenBank to FASTA
-GenBank Feature Extractor
-GenBank Trans Extractor
-One to Three
-Range Extractor DNA
-Range Extractor Protein
-Reverse Complement
-Split Codons
-Split FASTA
-Three to One
-Window Extractor DNA
-Window Extractor Protein
Sequence Analysis
-Codon Plot

Sequence Manipulation Suite:

Filter DNA

Filter DNA removes non-DNA characters from text. Use this program when you wish to remove digits and blank spaces from a sequence to make it suitable for other applications.

Paste the text into the text area below. Input limit is 500000 characters.

```
     1 ttaagatttg cgctttgcca actgtacacc caacctcggt
    41 ttattgtcga acctcccgct tgtgccgcca tctgcatata
    81 gatcccggtc agtccgtcac attctgccaa ttgagtatcc
   121 tcgaagtctt attccacgtg ctcaaagcaa gggtatcgta
   161 cagtgataac cgcctcgtgc agatccaaat tctcgattaa
   201 cactcaagta ctgatttta tcatcaggta actaaaaact
```

Please check the browser compatibility page before using this program.

Submit Clear Reset

Filter DNA results

>filtered DNA sequence consisting of 1000 bases.
ttaagatttgcgctttgccaactgtacacccaacctcggtttattgtcgaacctcccgct
tgtgccgccatctgcatatagatcccggtcagtccgtcacattctgccaattgagtatcc
tcgaagtcttattccacgtgctcaaagcaagggtatcgtacagtgataaccgcctcgtgc
agatccaaattctcgattaacactcaagtactgatttttatcatcaggtaactaaaaact
cacaatttgaagcaccagcgagaatcgttctattctctagcttcgcaacatcgacagttg
taatggcataacttcggcattcatagtggctgagtttagcggactaagcgaaaaactggt
cgttagatcttcctcaccatgattttacaagaaaggtgaactcaatttgacggcggtaaa
gttagatggctacgcgcgacaagtctccgtatcgtcatgaaattagcgaagaggtaatgg
caaagcttggctacgaatacaggagcgcgctgtgattacagtagggttaggatagcgaaa
acgttcaacgtggatagactcttatcggcacacgatcatatgcttccaaggttcccaagg
cgaattactagggtgcacagagctacgagtacgctgtccggcttgattcgctcgtacatc
cactgttcaaaaagctccgataccgacgatcactctcgatctctgtgtgggacgcactta

Reverse Complement

Reverse Complement converts a DNA sequence into its reverse, complement, or reverse-complement counterpart. The entire IUPAC DNA alphabet is supported, and the case of each input sequence character is maintained. You may want to work with the reverse-complement of a sequence if it contains an ORF on the reverse strand.

Paste the raw sequence or one or more FASTA sequences into the text area below.

```
>Sample sequence 1
aaaaaaaaaaaa

>Sample sequence 2
ctctctctc
```

Please check the browser compatibility page before using this program.

Submit Clear Reset

Reverse Complement results

>Sample sequence 1 reverse complement
ttttttttttt

>Sample sequence 2 reverse complement
gagagagag

attttttttttc

Reverse complement

Paste the raw sequence or one or more FASTA sequences into the text are

```
>Sample sequence 3
ccccccccccga
```

Submit Clear Reset

Reverse Complement results

```
>Sample sequence 3 reverse complement
tcgggggggggg
```

Vyzkoušejte si analyzovat „vlastní“ gen

Vyzkoušejte SMS suite.

DÚ5: nukleotidová sekvence

Pracujte s „vaším“ genem/proteinem/enzymem (př. NQO1)

1) Vyhledejte si vaši nukleotidovou sekvenci transkriptu (mRNA=RefSeq_transcript) a genu (RefSeq_gene).
   - Zapište si přístupové kódy. (NM_... a NG_...)
   - Jak dlouhé jsou tyto sekvence ? (stačí 1. transkript, máte-li více)
2) Vyberte pouze kódující sekvenci (CDS), uložte formát fasta.
3) Na kterém chromozomu leží? Kolik má exonů? Jak je dlouhý 2. exon? Stáhněte sekvenci 2. exonu.
4) Existuje nějaká souvislost vašeho genu v OMIM databázi?
5) Kolik obsahuje 2. exon Cytosinů? Kolik dinukleotidů CG?

DÚ5: Příklad řešení

DÚ5

1) NQO1:
Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1), transcript variant 1, mRNA

NM_000903.2

Transkript (mRNA): 2 521bp

2) >CDS

```
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTATGCCATGAAGG
AGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCGGACCTCTATGCCATGAACTT
CAATCCCATCATTTCCAGAAAGGACATCACAGGTAAACTGAAGGACCCTGCGAACTTTCAGTATCCTGCC
GAGTCTGTTCTGGCTTATAAAAGAAGGCCATCTGAGCCCAGATATTGTGGCTGAACAAAAGAAGCTGGAAG
CCGCAGACCTTGTGATATTCCAGTTCCCCCTGCAGTGGTTTGGAGTCCCTGCCATTCTGAAAGGCTGGTT
TGAGCGAGTGTTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCTTCCGGAGT
AAGAAGGCAGTGCTTTCCATCACCACTGGTGGCAGTGGCTCCATGTACTCTCTGCAAGGGATCCACGGGG
ACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCATTTCTGTGGCTTCCAAGTCTTAGAACC
TCAACTGACATATAGCATTGGGCACACTCCAGCAGACGCCCGAATTCAAATCCTGGAAGGATGGAAGAAA
CGCCTGGAGAATATTTGGGATGAGACACCACTGTATTTTGCTCCAAGCAGCCTCTTTGACCTAAACTTCC
AGGCAGGATTCTTAATGAAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCAACTGACAACCAGATCAAAGCTAGAAAATGA
```

3) Chromozome 16 (map="16q22.1)
4) 6 exonů (2:164nt)
>NM_000903.2:**199-363** Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1), transcript variant 1, mRNA

```
GCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTATGCCATGAAGGAGGCTGC
TGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCGGACCTCTATGCCATGAACTTCAATCCC
ATCATTTCCAGAAAGGACATCACAG
```

ANO, jeden (rs689453)

5) Exon obsahuje: 38 cytosinů a 4 dinukleotidy CG

DNA Stats results

Results for 165 residue sequence "Untitled" starting "GCAGAAGAGC"

Pattern:	Times found:	Percentage:
g	47	28.48
a	47	28.48
t	33	20.00
c	38	23.03
cg	4	2.44
ca	15	9.15

Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1), RefSeqGene on chromosome 16

NCBI Reference Sequence: NG_011504.2

Gen: 24 230 bp