NCBI Molecular Biology Resources for Toxicologists

Peter Cooper
National Library of Medicine
National Center for Biotechnology Information
Bethesda, MD
Phone: (301)-435-5951 Fax: (301) 480-9241
cooper@ncbi.nlm.nih.gov

Course Outline

- About NCBI
  - Affiliation and Purpose
  - NCBI Web Services
- Molecular Biology Databases
  - Primary vs. Derivative Databases
  - Sequences
    - GenBank
    - RefSeq
    - Protein
  - Other Databases

Course Outline

- Using Entrez as a Discovery System
  - Entrez Integration
  - Finding Paraoxonase 1 information
    - Global query results
    - Entrez results
    - Advanced search options
    - Finding homologs
    - Finding structural template
    - Mapping SNPs onto Structure
    - Genomic information
- NCBI BLAST services
  - Genome specific BLAST

Abbreviations

NCBI Glossary: .www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=handbook.glossary.1237

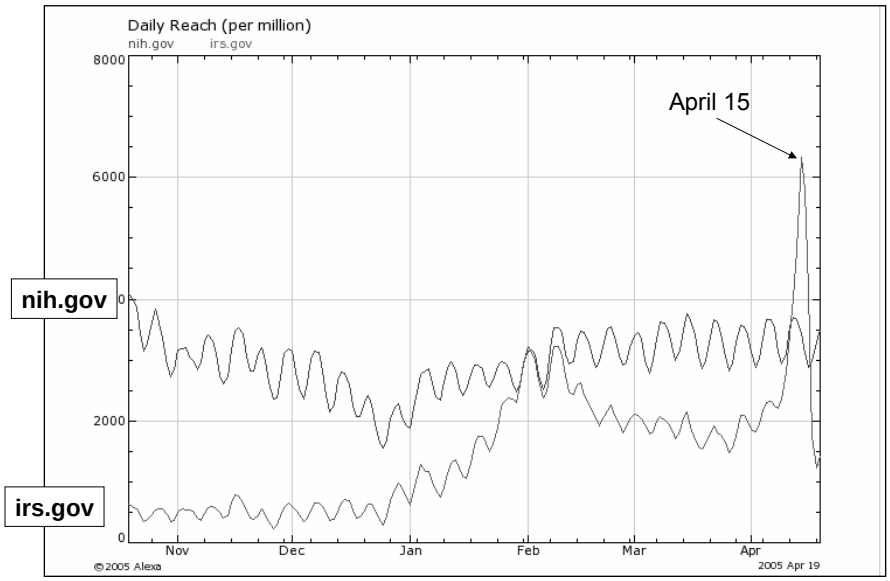
- BLAST – Basic Local Alignment Search Tool
- DDBJ – DNA databank of Japan
- EMBL – European Molecular Biology Laboratory
- EST – expressed sequence tag
- GEO – gene expression omnibus
- GSS – genome survey sequence
- HTG – high throughput genome sequence
- OMIM – online Mendelian inheritance in man
- RefSeq – reference sequence
- SNP – single nucleotide polymorphism
- WGS – whole genome shotgun sequence

The National Center for Biotechnology Information

Created in 1988 as a part of the National Library of Medicine at NIH

- Establish public databases
- Research in computational biology
- Develop software tools for sequence analysis
- Disseminate biomedical information

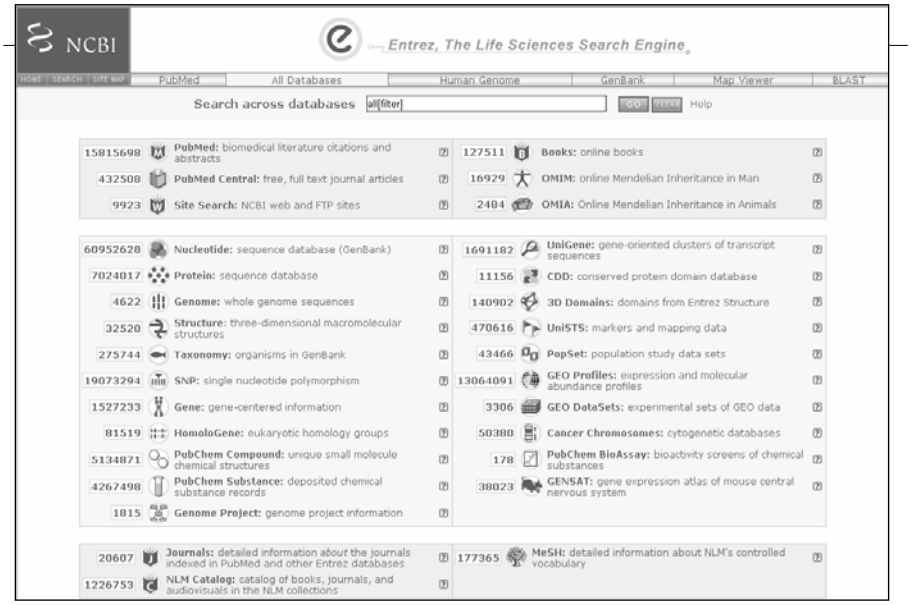
Government Web Traffic



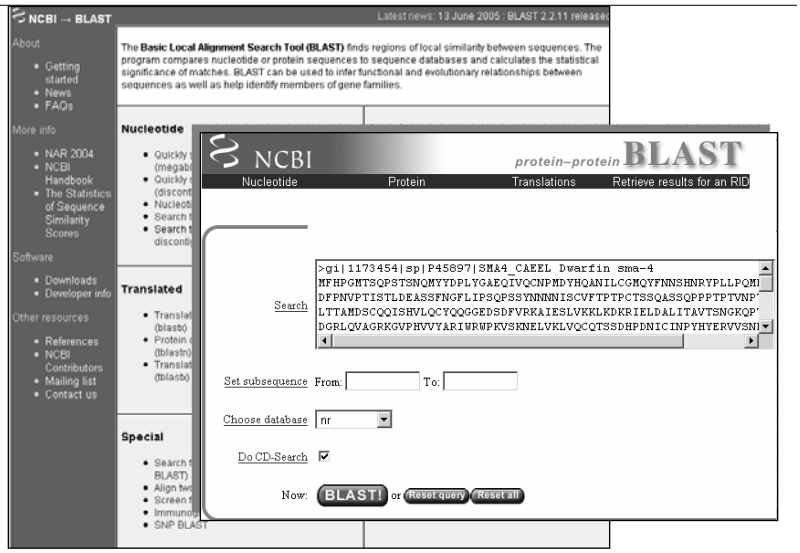
NCBI Databases and Services

- Free public access to biomedical literature
  - PubMed
  - PubMed Central full text online access
- GenBank largest sequence database
- Entrez integrated molecular and literature databases
- BLAST highest volume sequence search service

The Entrez System: Text Searches



BLAST: Sequence Similarity Searches



Types of Databases

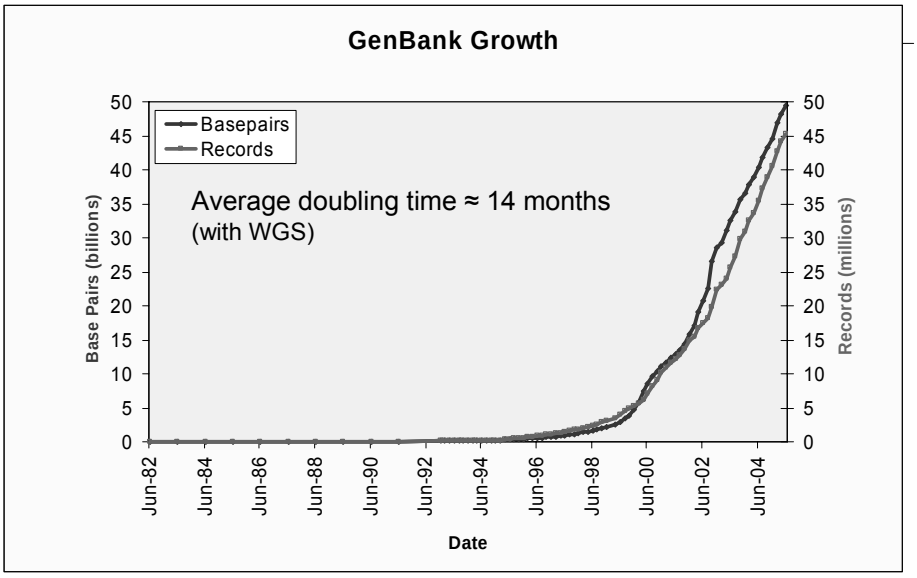
- Primary Databases
  - Original submissions by experimentalists
  - Content controlled by the submitter
    - Examples: GenBank, SNP, GEO
- Derivative Databases
  - Built from primary data
  - Content controlled by third party (NCBI)
    - Examples: RefSeq, TPA, RefSNP, UniGene, NCBI Protein, Structure, Conserved Domain

What is GenBank?

NCBI's Primary Sequence Database

- Nucleotide only sequence database
- Archival in nature
  - Historical
  - Reflective of submitter point of view (subjective)
  - Redundant
- GenBank Data
  - Direct submissions (traditional records)
  - Batch submissions (EST, GSS, STS)
  - ftp accounts (genome data)
- Three collaborating databases
  - GenBank
  - DNA Database of Japan (DDBJ)
  - European Molecular Biology Laboratory (EMBL) Database

The Growth of GenBank



RefSeq: NCBI's Derivative Sequence Database

- **Curated transcripts and proteins (NM_n...)**
  - **reviewed**
  - **human, mouse, rat, pig, cow, dog, possum, fruit fly, zebrafish, *Arabidopsis* ...**
    microbial genomes (proteins), and more
- **Model transcripts and proteins (XM_n...)**
- **Assembled Genomic Regions (contigs) (NT_n..., NW_n...)**
  - **Human, mouse, rat, chicken, sea urchin, honeybee**
- **Chromosome records (NC_n...)**
  - Human genome
  - microbial
  - organelle

`srcdb_refseq[Properties]`

ftp://ftp.ncbi.nih.gov/refseq/release/

RefSeq Benefits

- non-redundancy
- explicitly linked nucleotide and protein sequences
- updates to reflect current sequence data and biology
- data validation
- format consistency
- distinct accession series
- stewardship by NCBI staff and collaborators

Protein Database: Derivative Sequences

- GenPept
  - GenBank, EMBL, DDBJ, CDS Features
- NCBI RefSeqs
- Outside Protein-only DBs
  - Swiss-Prot www.expasy.org/sprot
    - Manually Curated Proteins
  - Protein Information Resource (PIR)
    - US protein database pir.georgetown.edu
  - Protein Research Foundation (PRF)
- Proteins from Structures
  - Protein Databank (PDB)
    www.rcsb.org/pdb

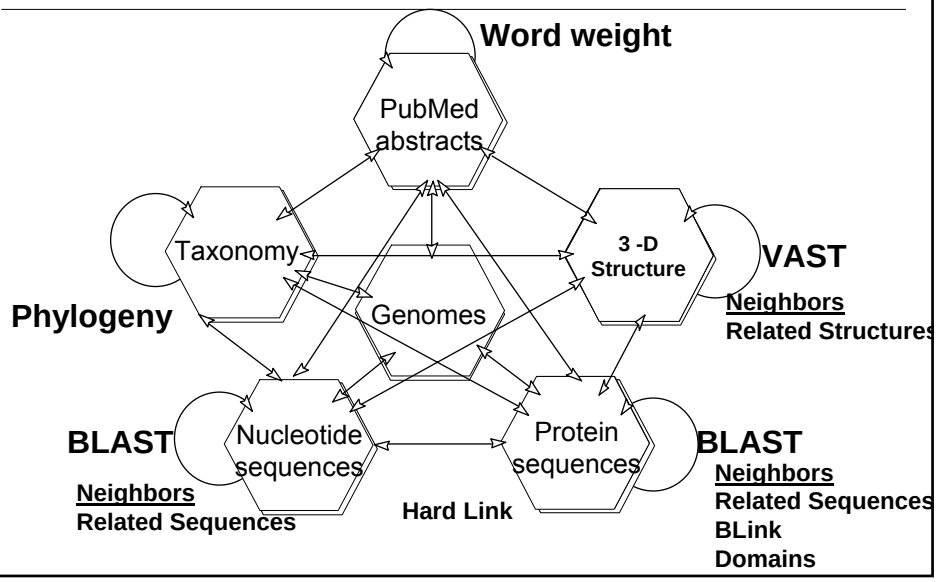
Other NCBI Databases

•dbSNP:	nucleotide polymorphism
•Geo:	Gene Expression Omnibus microarray and other expression data
•Gene:	gene records Unifies LocusLink and Microbial Genomes
•Structure:	imported structures (PDB) Cn3D viewer, NCBI curation
•CDD:	conserved domain database Protein families (COGs and KOGs) Single domains (PFAM, SMART, CD)

Using the NCBI System: Paraoxonase

- Entrez integration
- Finding human RefSeq
- Online Mendelian Inheritance in Man
- Finding Homologs
- Finding a Structure
- Mapping Polymorphisms
- Genomic Information
- Genomic BLAST Search

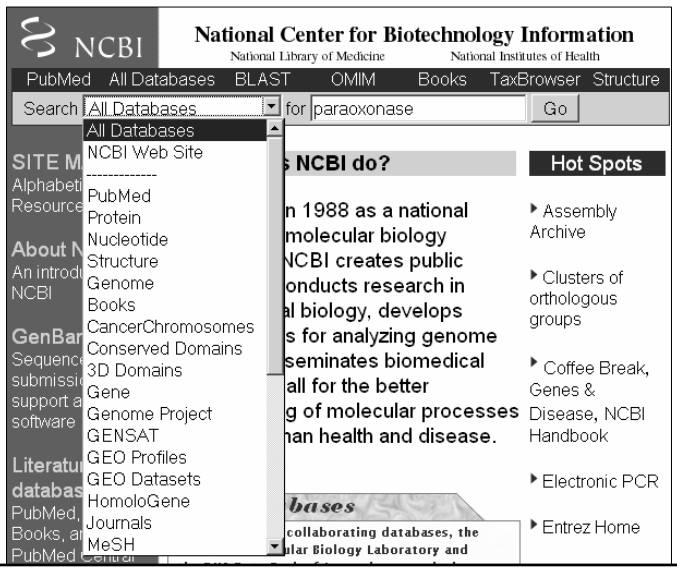
Entrez: Database Integration



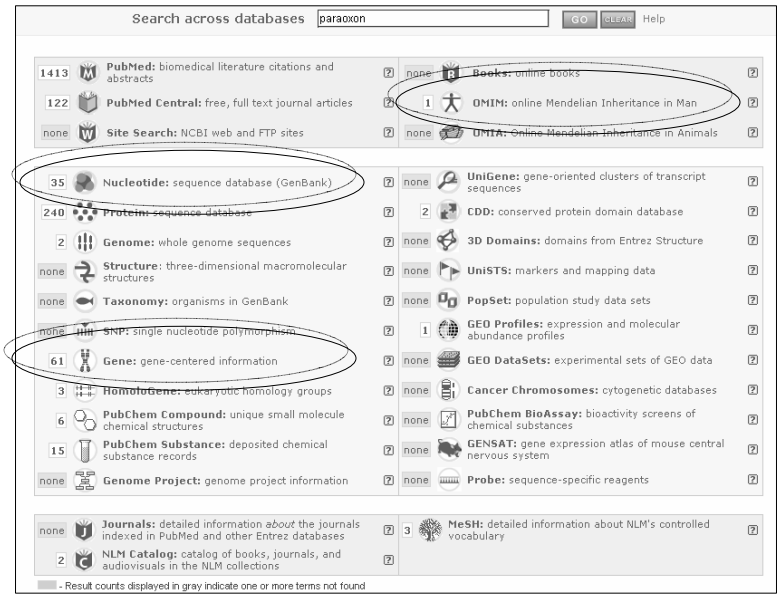
Entrez Searches: Paraoxonase

- Live Web Demonstration beginning at the NCBI homepage: http://www.ncbi.nlm.nih.gov
- Slides 20 – 47 Web pages from live Entrez searches

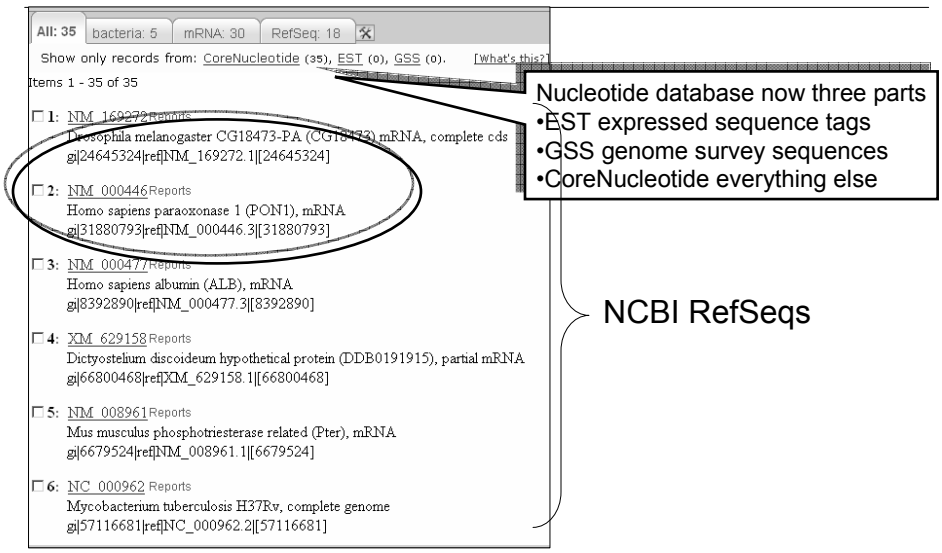
NCBI Homepage: Global Entrez Query



Global Query: paraoxon



Nucleotide Results



Nucleotide database now three parts
- EST expressed sequence tags
- GSS genome survey sequences
- CoreNucleotide everything else

NCBI RefSeqs

Nucleotide Results (cont.)

All: 35 | bacteria: 5 | mRNA: 30 | RefSe

Show only records from: CoreNucleotide

Items 1 - 35 of 35

1: NM_169272 Reports
Drosophila melanogaster CG18473-PA (
gi|24645324|ref|NM_169272.1|[246453

2: NM_000446 Reports
Homo sapiens paraoxonase 1 (PON1), m
gi|31880793|ref|NM_000446.3|[318807

3: NM_000477 Reports
Homo sapiens albumin (ALB), mRNA
gi|8392890|ref|NM_000477.3|[8392890

4: XM_629158 Reports
Dictyostelium discoideum hypothetical pr
gi|66800468|ref|XM_629158.1|[668004

5: NM_008961 Reports
Mus musculus phosphotriesterase related
gi|6679524|ref|NM_008961.1|[6679524

6: NC_000962 Reports
Mycobacterium tuberculosis H37Rv, com
gi|57116681|ref|NC_000962.2|[571166

30: BC044294 Reports Links
Xenopus laevis Serum paraoxonase/arylesterase 2, mRNA (cDNA clone MGC:53915 IMAGE:5569890), complete cds
gi|28422436|gb|BC044294.1|[28422436]

31: BC046160 Reports Links
Homo sapiens paraoxonase 2, mRNA (cDNA clone MGC:57708 IMAGE:5393700), complete cds
gi|28279806|gb|BC046160.1|[28279806]

32: BC021887 Reports Links
Mus musculus paraoxonase 2, mRNA (cDNA clone MGC:11614 IMAGE:3154583), complete cds
gi|18255249|gb|BC021887.1|[18255249]

33: AY243367 Reports Links
Enterobacter aerogenes putative zinc-type alcohol dehydrogenase (adh) gene, partial cds; and putative phosphohydrolase operon, complete sequence
gi|37935738|gb|AY243367.1|[37935738]

34: AF212237 Reports Links
Homo sapiens HPHRP mRNA, complete cds
gi|13182754|gb|AF212237.1|AF212237[13182754]

35: U28016 Reports Links
Mus musculus parathion hydrolase (phosphotriesterase)-related protein mRNA, complete cds
gi|881498|gb|U28016.1|MMU28016[881498]

GenBank

Protein Results

All: 240 | bacteria: 188 | RefSeq: 96

Items 1 - 240 of 240

1: NP_853900 Reports
PROBABLE PHOSPHOTRIESTERASE PHP (PARATHI
(ARYLDIALKYLPHOSPHATASE) (PARAOXONASE)
(ARYLTRIPHOSPHATASE) (PARAOXON HYDROLAS
AF2122/97]
gi|31791407|ref|NP

2: NP_214744 Reports
PROBABLE PHOSPHOTRIESTERASE PHP (PARATHI
(ARYLDIALKYLPHOSPHATASE) (PARAOXONASE)
(ARYLTRIPHOSPHATASE) (PARAOXON HYDROLAS
tuberculosis H37Rv]
gi|15607371|ref|NP_214744.1|[15607371]

3: CAD93099 Reports
PROBABLE PHOSPHOTRIESTERASE PHP (PARATHIO
(ARYLDIALKYLPHOSPHATASE) (PARAOXONASE)
(ARYLTRIPHOSP N HYDROLASE
AF2122/97]
gi|31616992|emb|CA

RefSeqs

GenPept

7: NP_000437 Reports
paraoxonase 1 [Homo sapiens]
gi|19923106|ref|NP_000437.3|[19923106]

8: P0A433 Reports
Parathion hydrolase precursor (Phosphotriesterase) (PTE)
gi|61229328|sp|P0A433|OPD_FLAS2[61229328]

9: P45548 Reports
Phosphotriesterase homology protein
gi|1176259|sp|P45548|PHP_ECOLI[1176259]

10: NP_822616 Reports
putative phosphotriesterase-family protein [Streptomyces avermitilis MA-4680]
gi|29827982|ref|NP_822616.1|[29827982]

Swiss-Prot

233: 1PSCA Reports
Chain A, Phosphotriesterase From Pseudomonas Diminuta
gi|2098312|pdb|1PSC|A[2098312]

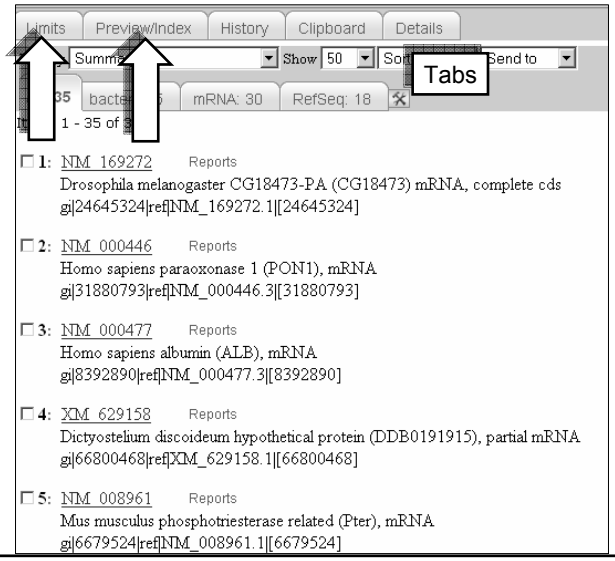
234: 1PTA Reports
Chain , Mol_id: 1; Molecule: Phosphotriesterase; Chain: Null; Ec: 3.5.-.-
gi|1310974|pdb|1PTA| [1310974]

PDB

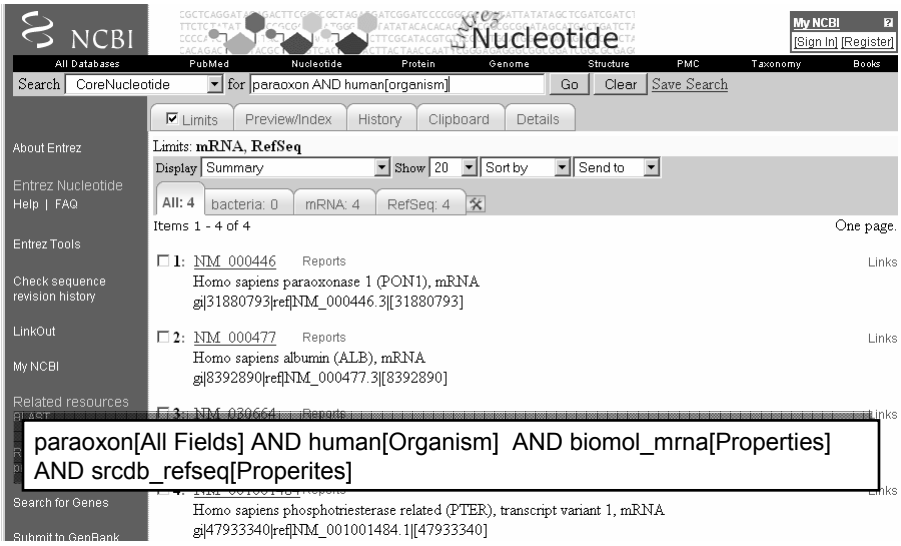
235: 1601316A Reports
parathion hydrolase
gi|226530|prf||1601316A[226530]

PRF

Advanced Search Options



More Precise Nucleotide Search



paraoxon[All Fields] AND human[Organism] AND biomol_mrna[Properties] AND srcdb_refseq[Properites]

Useful Field Restrictions

[Title]: Definition line in GenBank / GenPept format shown in Summary format

glyceraldehyde 3 phosphate dehydrogenase[Title]

[Organism]: NCBI's taxonomy. Organizing system for molecular databases

mouse[organism]; green plants[organism]; Streptomyces coelicolor[organism]

[Properties]: molecule type, location, database source

biomol_mrna[properties]; biomol_genomic[properties];
gene_in_mitochondrion[properties]; srcdb pdb[properties]

[Filter]: subsets of data, Entrez links

all[filter]; nucleotide mapview[Filter]; nucleotide omim[Filter]

Human MutL RefSeq

1: NM_000446. Reports Homo sapiens para...[gi:31880793] Links

Jump to Comment Features Sequence

```
LOCUS       NM_000446               2395 bp    mRNA    linear   PRI 06-NOV-2005
DEFINITION  Homo sapiens paraoxonase 1 (PON1), mRNA.
ACCESSION   NM_000446
VERSION     NM_000446.3  GI:31880793
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2395)
```

```
COMMENT     VALIDATED REFSEQ: This record has undergone preliminary review of
            the sequence, but has not yet been subject to final review. The
            reference sequence was derived from CB122285.1 and Z70723.1.
            On Jun 18, 2003 this sequence version replaced gi:19923105.
```

```
  REMARK    GeneRIF: Overall, our findings suggest that protecting the
            esterolytic activity of PON1 by antioxidants might be important in
            preserving its action on phospholipid peroxides.
REFERENCE   2  (bases 1 to 2395)
  AUTHORS   Furlong,C.E., Cole,T.B., Jarvik,G.P., Rettan-Brewer,C., Geiss,G.K.,
            Richter,R.J., Shih,D.M., Tward,A.D., Lusis,A.J. and Costa,L.G.
  TITLE     Role of paraoxonase (PON1) status in pesticide sensitivity: genetic
  JOURNAL   [illegible] (2005)
   PUBMED   [illegible]
  REMARK    Review article
            GeneRIF: Individual differences in detoxication capacities for
            specific organophosphorous compounds are due largely to differences
            in catalytic efficiency or abundance of the HDL-associated enzyme,
```

GenBank Records

NM_000446: Links

1: NM_000446. Reports Homo sapiens para...[gi:31880793]

Jump to Comment Features Sequence

```
LOCUS       NM_000446                2395 bp    mRNA    linear   PRI
DEFINITION  Homo
ACCESSION   NM_000
VERSION     NM_000
KEYWORDS    .
SOURCE      Homo
  ORGANISM  Homo
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut
            Mammalia; Eutheria; Euarchontoglires; Primates;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2395)
  AUTHORS   Karabina,S.A., Lehner,A.N., Frank,E., Parthasar
            Santanam,N.
  TITLE     Oxidative inactivation of paraoxonase--implications in
            mellitus and atherosclerosis
  JOURNAL   Biochim. Biophys. Acta 1725 (2), 213-221 (2005)
   PUBMED   16112460
  REMARK    GeneRIF: Overall, our findings suggest that protecting
            esterolytic activity of PON1 by antioxidants might be important in
            preserving its action on phospholipid peroxides.
REFERENCE   2  (bases 1 to 2395)
  AUTHORS   Furlong,C.E., Cole,T.B., Jarvik,G.P., Pettan-Brewer,C., Geiss,G.K.,
            Richter,R.J., Shih,D.M., Tward,A.D., Lusis,A.J. and Costa,L.G.
  TITLE     Role of paraoxonase (PON1) status in pesticide sensitivity: genetic
            and temporal determinants
  JOURNAL   Neurotoxicology 26 (4), 651-659 (2005)
   PUBMED   16112327
  REMARK    Review article
```

Links

- Gene
- HomoloGene
- Genome
- Genome Project
- Master
- Full text in PMC
- Probe
- Gene Genotype
- GeneView in dbSNP
- Related Sequences
- Map Viewer
- OMIM
- GEO Profiles
- Protein
- PubMed
- SNP
- Taxonomy
- UniGene
- UniSTS
- LinkOut

Nucleotide Neighbors

- Same gene sequences
- Highly stringent

OMIM

Online Mendelian Inheritance in Man

1: +168820

PARAOXONASE 1; PON1

Alternative titles; symbol

PON
PAROXONASE, PLASMA
ARYLESTERASE
ESTERASE A; ESA
ORGANOPHOSPHATE F
PARATHION POISONING
CORONARY ARTERY DI
CORONARY SPASMS, S

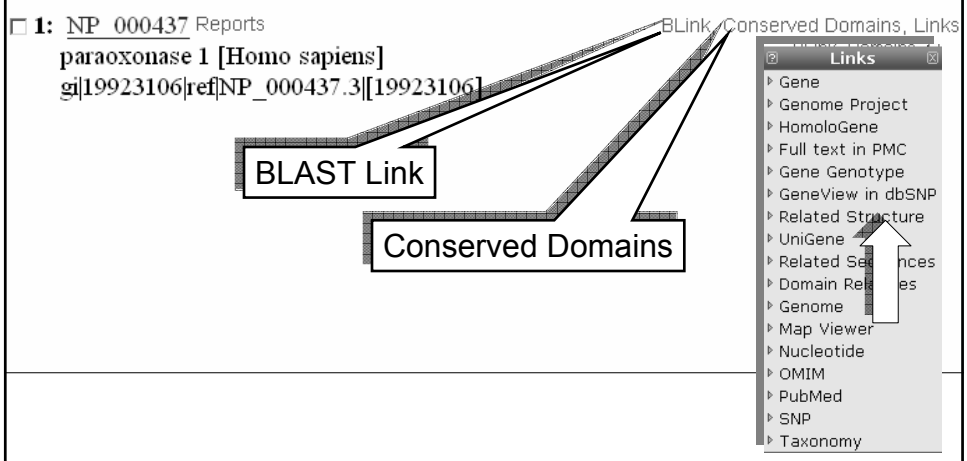
.0001 CORONARY ARTERY DISEASE, SUSCEPTIBILITY TO [PON1, GLN192ARG]

CORONARY SPASMS, SUSCEPTIBILITY TO, INCLUDED

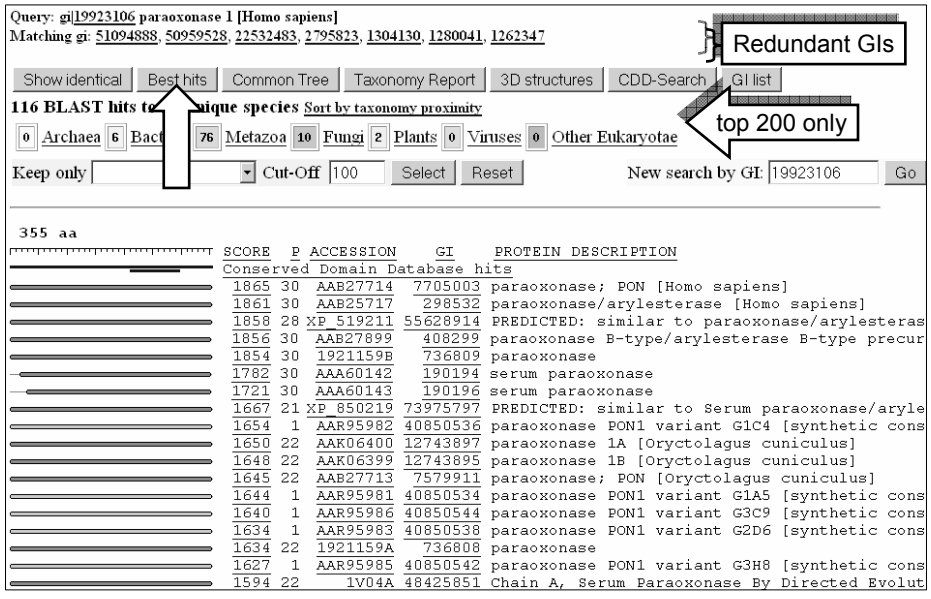
Hassett et al. (1991) isolated and sequenced PON cDNAs from humans and rabbits. Humbert et al. (1993) sequenced these clones and determined that arginine at position 192 specifies high activity plasma paraoxonase, whereas glutamine at this position specifies a low activity variant. They showed that allele-specific probes or restriction enzyme analysis of amplified DNA could be used for genotyping of individuals. This polymorphism is also referred to as GLN191ARG. Adkins et al. (1993) also demonstrated that glutamine or arginine at amino acid 191 determines the A and B allozymes, respectively, of PON1.

The arg192 isoform of paraoxonase hydrolyzes paraoxon more rapidly than the gln192 isoform. However, with respect to hydrolysis of toxic nerve agents, such as sarin, the arg192 isoform displays a lower activity than the other isoform. Yamasaki et al. (1997) found that the arg192 allele is more common in the Japanese (allele frequency, 0.66) than in people of other races (range, 0.24 to 0.31). In Japanese, 41.4% of subjects were homozygous for the arg192 allele, which shows a very low hydrolysis activity for sarin. Thus, there seems to be a racial difference in vulnerability to toxic nerve agents, such as sarin. The dominance of the arg192 allele in the Japanese population probably worsened the tragedy of March 1995 in the Tokyo subway.

Protein Link



BLink: BLAST Link



BLink: RefSeqs by Species

Query: gi|19923106 paraoxonase 1 [Homo sapiens]
Matching gi: 51094888, 50959528, 22532483, 2795823, 1304130, 1280041, 1262347

Show identical | All hits | Common Tree | Taxonomy Report | 3D structures | CDD-Search | GI list

13 BLAST hits to 7 selected species Sort by taxonomy proximity

0 Archaea | 0 Bacteria | 13 Metazoa | 0 Fungi | 0 Plants | 0 Viruses | 0 Other Eukaryotae

Keep only REFSEQ Cut-Off 100 Select Reset

New search by GI: 19923106 Go

mouse homolog

355 aa

BLAST

SCORE	P	ACCESSION	GI	N	ORGANISM
Conserved Domain Database hits					
1581	22	NP_035264	7242183	- 3	*Mus musculus*
1547	22	NP_114466	54292130	- 3	*Rattus norvegicus*
1277	30	NP_000296	66529294	- 2	*Homo sapiens*
1228	21	NP_001...	55741451	- 1	*Canis familiaris*
1156	17	NP_001...	55742304	- 1	*Xenopus tropicalis*
932	15	NP_001...	66773108	- 3	*Danio rerio*

Related Structures

NCBI **Related Structures**

HOME | SEARCH | SITE MAP | PubMed | Blast | Entrez Structure | Help

Query: paraoxonase 1 [Homo sapiens]
[gi: 19923106]

List | All MMDB | sequences, sort by | Blast E value | and display as | Graphic

Show Page 1 at 50 structures per page

1 hits with known structures found

Page 1 of 1

0 75 150 225 300 355

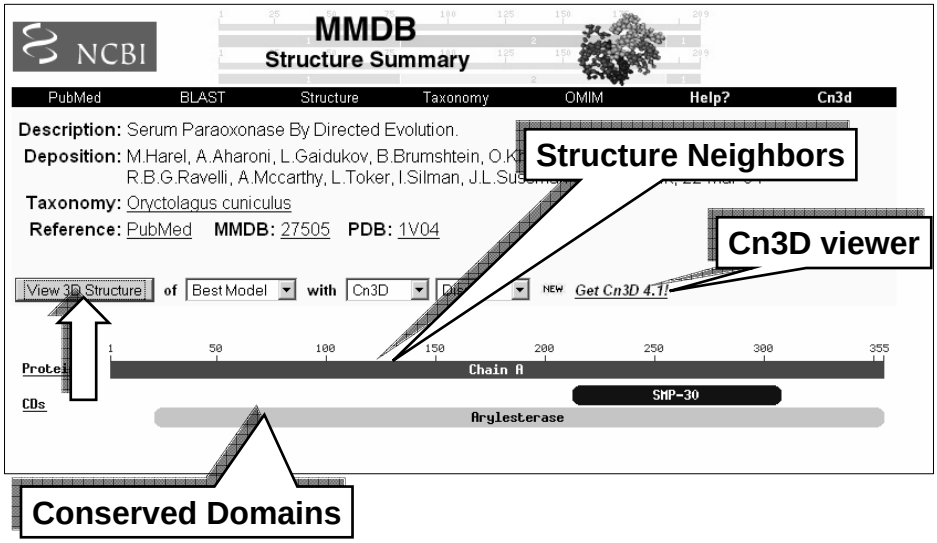
Arylesterase

SMP-30

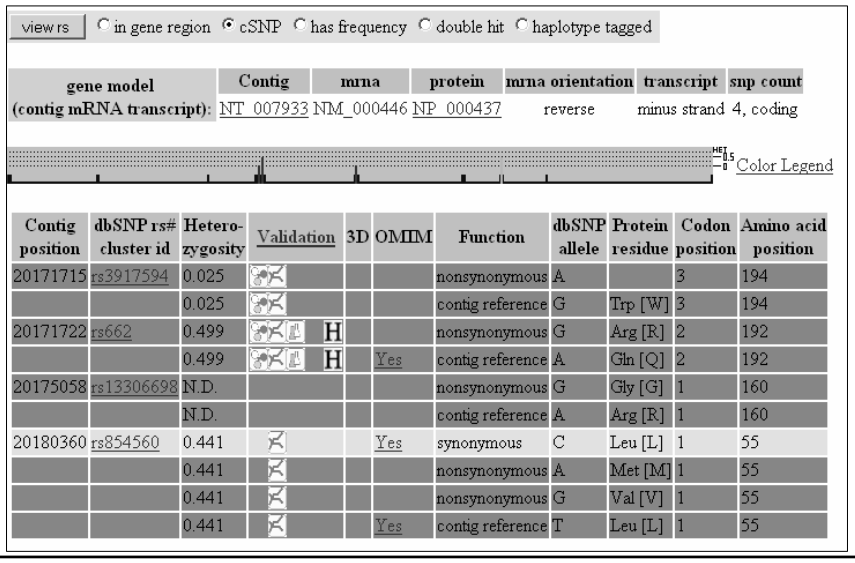
Structure | E Value

1V04 A | 1e-125

Rabbit Structure



GeneView: Variations Human PON1



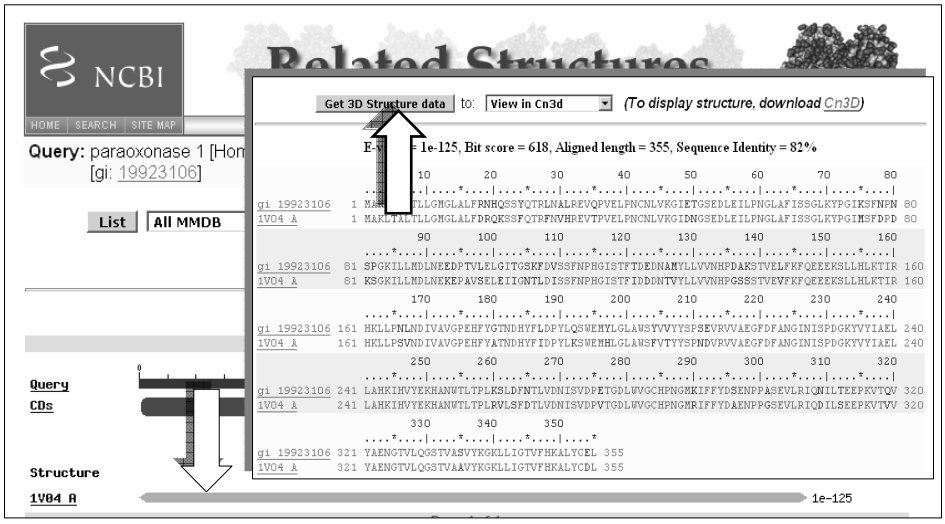
view rs | in gene region | cSNP | has frequency | double hit | haplotype tagged

gene model (contig mRNA transcript):	Contig	mrna	protein	mrna orientation	transcript	snp count
	NT_007933	NM_000446	NP_000437	reverse	minus strand	4, coding

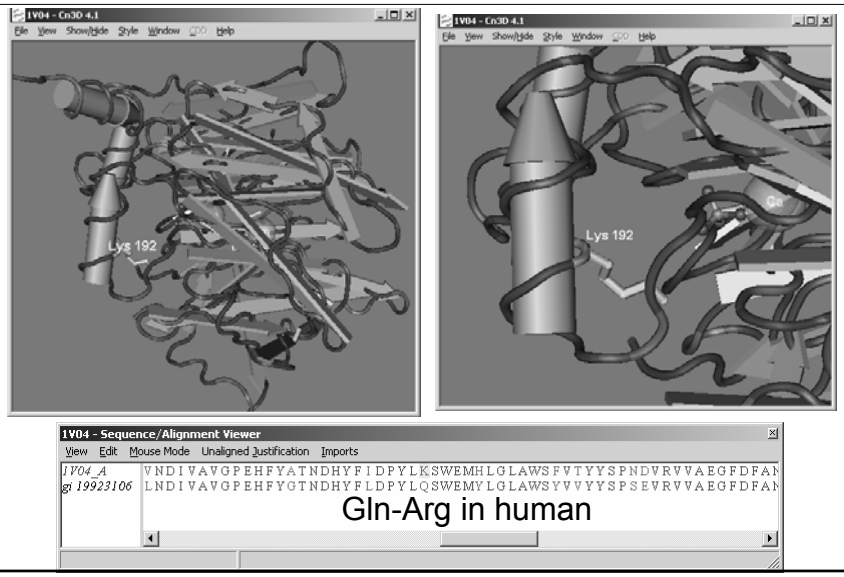
Color Legend

Contig position	dbSNP rs# cluster id	Hetero-zygosity	Validation	3D	OMIM	Function	dbSNP allele	Protein residue	Codon position	Amino acid position
20171715	rs3917594	0.025				nonsynonymous	A		3	194
		0.025				contig reference	G	Trp [W]	3	194
20171722	rs662	0.499	H			nonsynonymous	G	Arg [R]	2	192
		0.499	H		Yes	contig reference	A	Gln [Q]	2	192
20175058	rs13306698	N.D.				nonsynonymous	G	Gly [G]	1	160
		N.D.				contig reference	A	Arg [R]	1	160
20180360	rs854560	0.441			Yes	synonymous	C	Leu [L]	1	55
		0.441				nonsynonymous	A	Met [M]	1	55
		0.441				nonsynonymous	G	Val [V]	1	55
		0.441			Yes	contig reference	T	Leu [L]	1	55

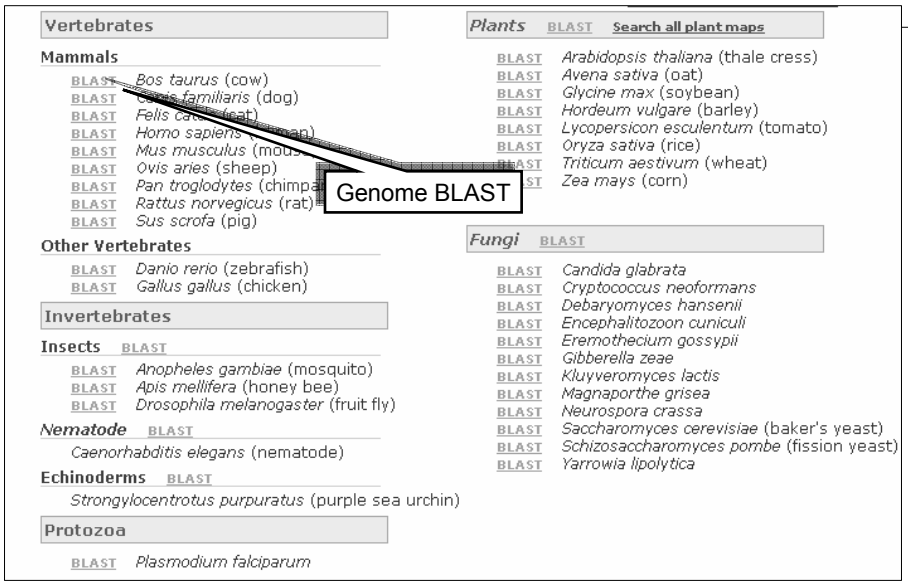
Related Structures



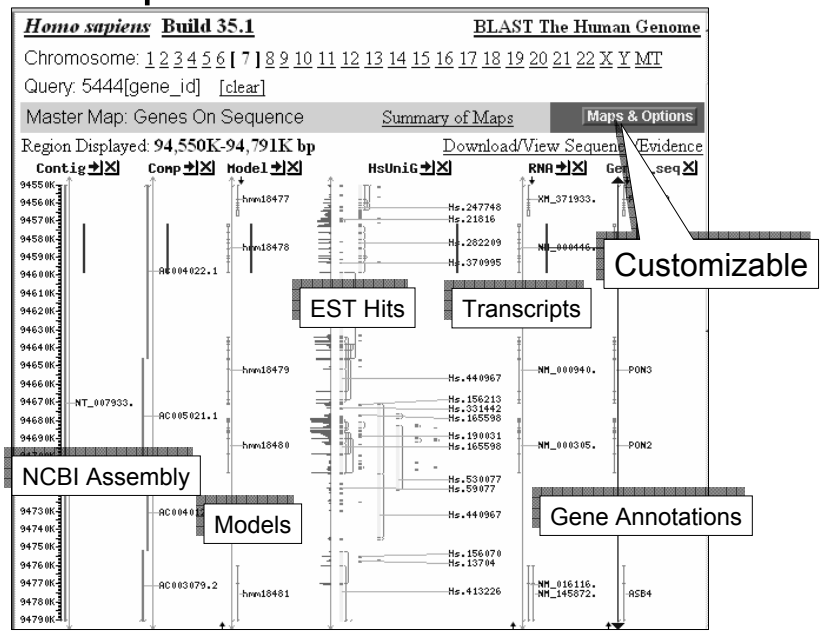
Mapping Variation Onto Structure



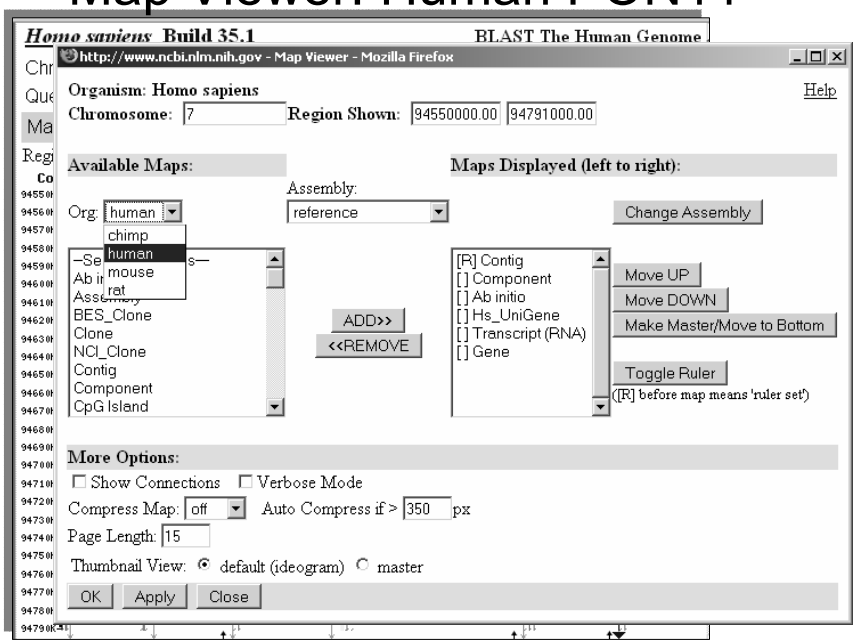
The Map Viewer



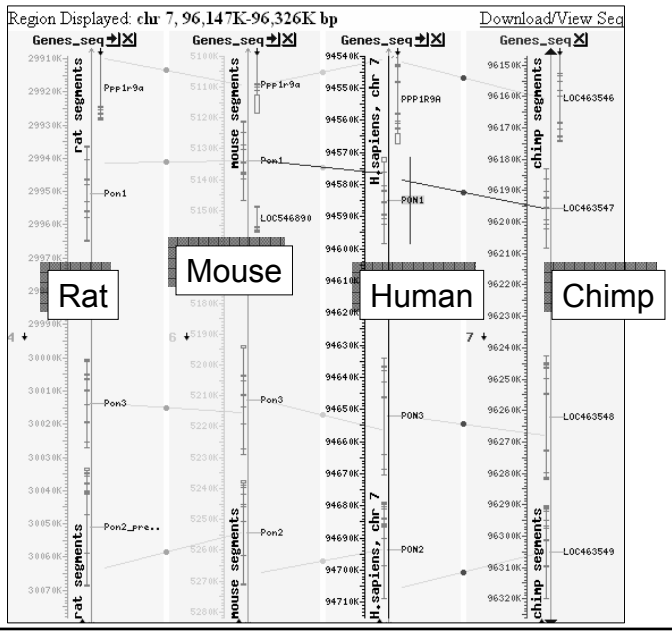
Map Viewer: Human PON11



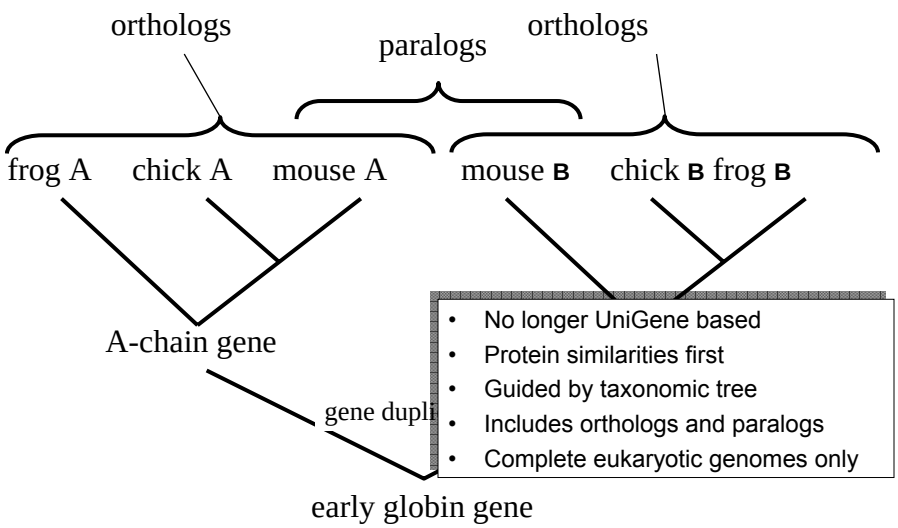
Map Viewer: Human PON11



Synteny: Mammalian Genomes

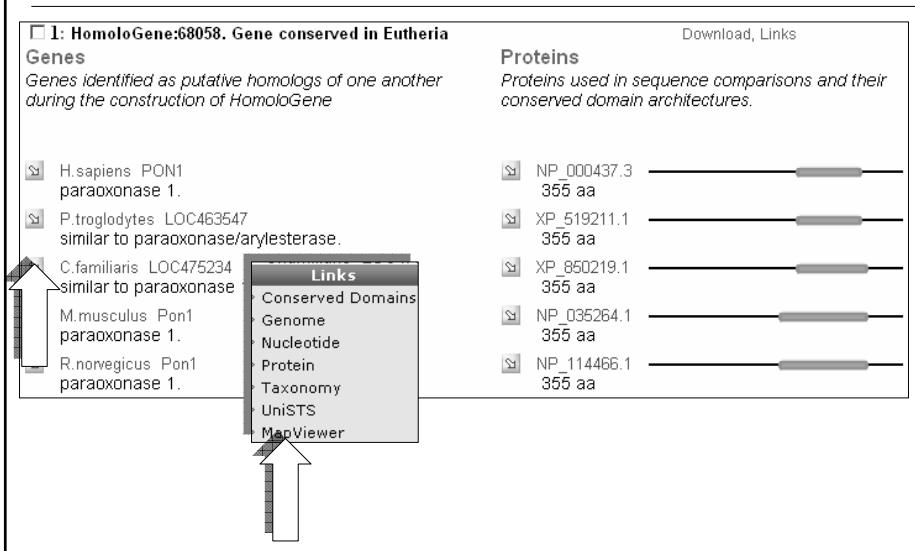


The New HomoloGene

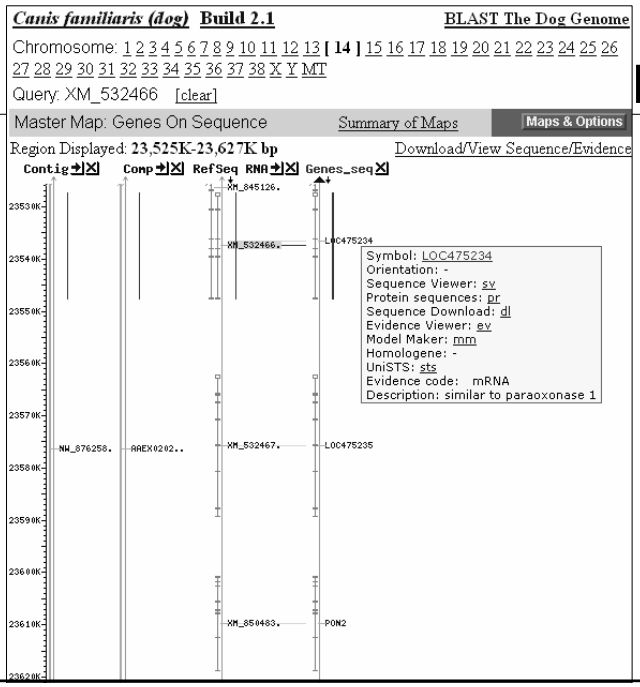


- No longer UniGene based
- Protein similarities first
- Guided by taxonomic tree
- Includes orthologs and paralogs
- Complete eukaryotic genomes only

The New HomoloGene



Dog Homolog



The Gene Database

- Gene Centered Information
- Unifies LocusLink and microbial Genomes
- 1.5 million records for 3,038 taxa

Human	32,924	Chicken	18, 444
Chimpanzee	24,158	Zebrafish	37, 715
Mouse	47,988	Sea Urchin	20,636
Rat	28,584	Mosquito	13,897
Dog	20,153	Fruit Fly	20,763
Cow	23, 647	*C. elegans*	20,846

Genes PON1: One Stop Shopping

☐ 1: **PON1 paraoxonase 1** [*Homo sapiens*]
GeneID: 5444 Locus tag: HGNC:9204; MIM: 168820 updated 07-Nov-2005

Summary

Official Symbol: PON1 **and Name:** paraoxo
Committee
Gene type: protein coding
Gene name: PON1
Gene description: paraoxonase 1
RefSeq status: Validated

Genomic regions, transcripts, and products

(minus strand) RefSeq below

NC_000007

94598495 | 94571639
5' | 3'
NM_000446 | NP_000437
■ - coding region ■ - untranslated region

Bibliography Gene References into Function (GeneRIF): Submit

PubMed links
GeneRIFs:
1. significant difference in the male genotype distribution of the CLA-1/SR-BI exon 8 (C8/T8) variant between groups and this significant risk was restricted to those subject carriers of Arg (R) and Leu (L) allele of the PON1 PubMed
2. PON1 direc
decreased supe
for attenuated a

See PON1 in MapViewer

94708946
PON3
PON1 PON2

Interactions

Description .....

PON1 Product	Interactant	Other Gene	Complex	Source	Pubs
NP_000437.3	NP_001822.1	CLU		HPRD	PubMed
NP_000437.3	NP_001822.2	CLU		HPRD	PubMed
Paraoxonase 1	NP_000468.1	ALB		HPRD	PubMed

Genes PON1: One Stop Shopping

General gene information

Markers
STS-M63012(e-PCR) (Links: UniSTS: 46466)
Alternate names: RH40691; sts-M63012
RH12859(e-PCR) (Links: UniSTS: 63408)
Alternate name: stSG8066
G44327(e-PCR) (Links: UniSTS: 95116)
Alternate name: WIAF-1100-STS
GDB:194296(e-PCR) (Links: UniSTS: 155789)

GeneOntology
Provided by GOA

Function	Evidence	
aryldialkylphosphatase activity	IEA	
arylesterase activity	IEA	
arylesterase activity	NAS	PubMed
hydrolase activity	IEA	
lipid transporter activity	IEA	
Process		
response to external stimulus	NAS	PubMed
Component		
extracellular region	IEA	
extracellular region	NAS	PubMed

Homology:
Mouse, Rat
Map Viewer

Phenotypes
Coronary artery disease, susceptibility to MIM: 168820
Coronary artery spasm, susceptibility to MIM: 168820

Pathways
KEGG pathway: Bisphenol A degradation 00363
KEGG pathway: gamma-Hexachlorocyclohexane degradation 0036

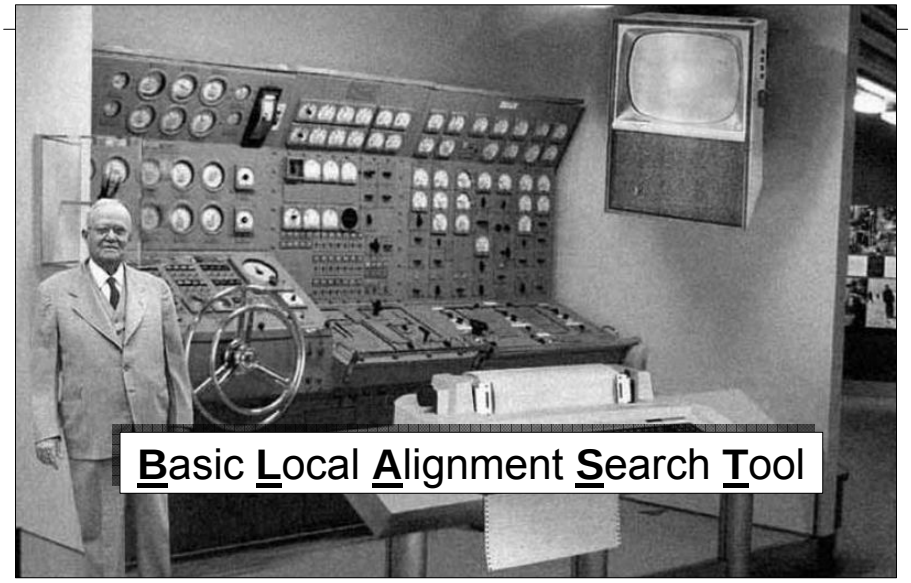
NCBI Reference Sequences (RefSeq)

mRNA Sequence NM_000446
Source Sequence CB122285,Z70723
Product NP_000437 paraoxonase 1
Consensus CDS (CCDS) CCDS5638.1
Conserved Domains (2) summary
pfam03758: SMP-30; Senescence marker protein-30 (SMP-30)
Location: 212 - 298 Blast Score: 97
pfam01731: Arylesterase; Arylesterase
Location: 2 - 355 Blast Score: 1316

Related Sequences

Nucleotide		Protein
Genomic	AC004022	AAC35293
Genomic	AF402963	AAL04502
Genomic	AF402965	AAL04504
Genomic	AF539592	AAM97935
Genomic	CH236949	EAL24133
Genomic	S56555	AAB25717
Genomic	S64696	AAB27899
Genomic	U55885	AAB41835
mRNA	BC074719	AAH74719
mRNA	CB122285	None
mRNA	D84371	BAA12327
mRNA	M63012	AAB59538
mRNA	M63013	AAA60142
mRNA	M63014	AAA60143
mRNA	S64615	AAB27714
mRNA	U53784	AAA97957
mRNA	Z70723	CAA94728
None		P27169
		Q6B0J6
		Q96P89
		Q96P91
		Q9UCB1

Sequence Similarity Searching



What BLAST tells you

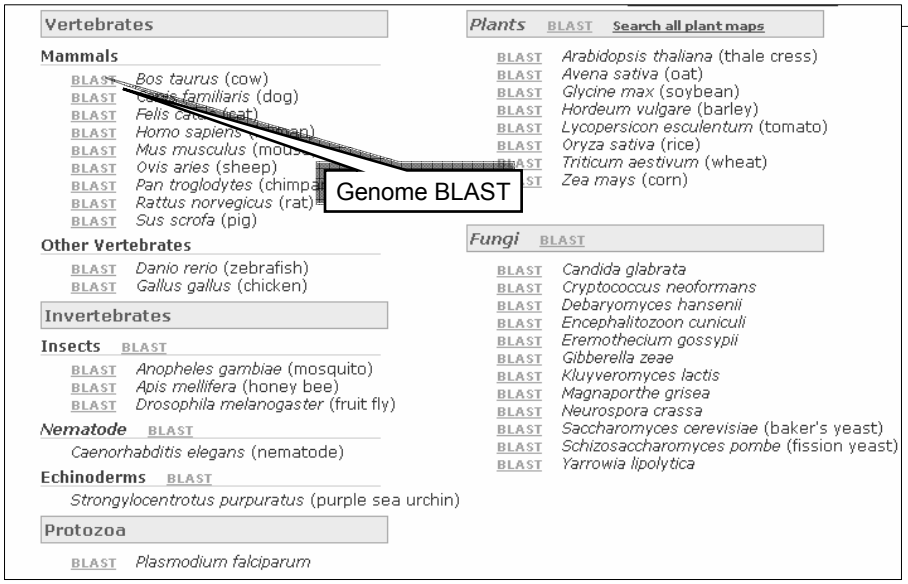
- BLAST reports surprising alignments
  - Different than chance
- Assumptions
  - Random sequences
  - Constant composition
- Conclusions
  - Surprising similarities imply **evolutionary homology**

Evolutionary Homology: descent from a common ancestor
Does not always imply similar function

Finding Cow Paraoxonase Genes: Genomic BLAST

- Live Web Demonstration linking to genomic BLAST from the Map Viewer
- http://www.ncbi.nlm.nih.gov/mapview
- Slides 51 – 54 Web pages from live BLAST searches

The Map Viewer



Cow Genome BLAST

NCBI Home ▸ Genomic Biology ▸ Cow Genome Resources ▸ BLAST

Search Map Viewer Go Clear

BLAST
overview
FAQs
news
manual
references

BLAST Cow Sequences

Blast your sequence against Cow specific sequences

Database genome (reference only)
genome (reference only)

Program

TBLASTN: Compare a protein sequence against a nucleotide database
megaBLAST: Compare highly related nucleotide sequences
cross-species megaBLAST: Compare distantly related nucleotide sequences
BLASTN: Compare nucleotide sequences
BLASTP: Compare protein sequences
BLASTX: Compare a nucleotide sequence against a protein database
TBLASTN: Compare a protein sequence against a nucleotide database

ESTs
Clone end sequences
Traces-WGS
Traces-Other
WGS contigs
SNPs

Begin S

Enter an

Optional p

Expect	Filter	Descriptions	Alignments
0.01	low complexity	100	100

Advanced options:

Begin Search Clear Input

BLAST Results

Database: ref_contig
102,340 sequences; 3,100,376,884 total letters

If you have any problems or questions with the results of this search please refer to the **BLAST FAQs**

Taxonomy reports

Genome View Show positions of the BL using the Entrez Genomes

Query= gi|19923106|ref|NP_000437.3| paraox (355 letters)

Distribution of 16 Blast Hits

Mouse-over to show defline and scores, click to

16 hits from three cont

Query 0 70 140

Sequences producing significant alignments

ref|NW 931245.1|Bt4 WGA715 2	Bos taurus chromosome 4 genomic con	123	2e-26
ref|NW 931246.1|Bt4 WGA716 2	Bos taurus chromosome 4 genomic con	113	2e-23
ref|NW 934662.1|BtUn WGA102181 2	Bos taurus chromosome Un genomi	83.6	2e-14

>ref|NW 931245.1|Bt4 WGA715 2 Bos taurus chromosome 4 genomic contig
Length=543455

Features in this part of subject sequence:
similar to paraoxonase 1

Score = 123 bits (308), Expect = 2e-26
Identities = 54/70 (77%), Positives = 63/70 (90%), Gaps = 0/70 (0%)
Frame = -3

Query 166 NLNDIVAVGPEHFYGTNDHYFLDPYLQSWEMYLGLAWSYVVYYSPSEVRVVAEGFDFANG 225
+LND+VAVG EHFY TNDHYF+D Y++SWE+YLGLAWS VVYYSP +VRVVA GFDFANG
Sbjct 530670 SLNDLVAVGREHFYATNDHYFVDQYMRSWELYLGLAWSNVVYYSPDDVRVVASGFDFANG 530491

Query 226 INISPDGKYV 235
I++SPD KYV
Sbjct 530490 ISNSPDRKYV 530461

•Protein alignments
•Sorted by e-value

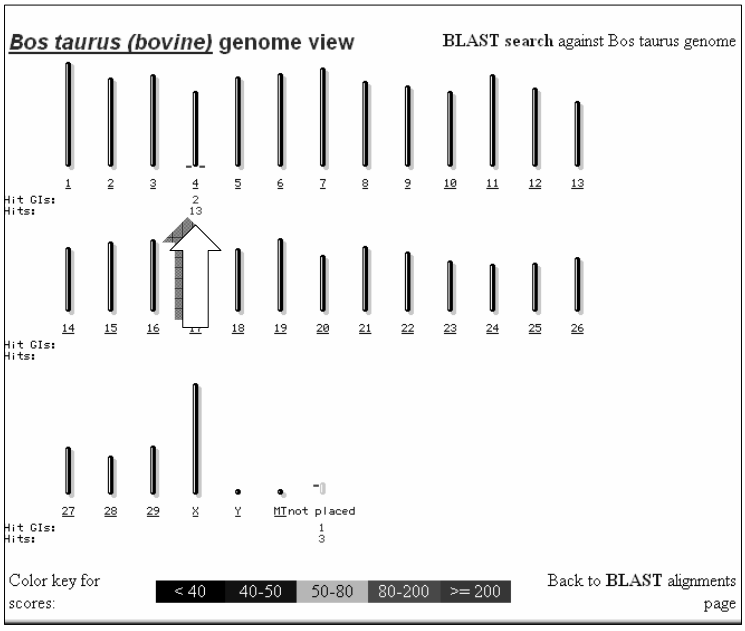
Features in this part of subject se
similar to paraoxonase 1

Score = 103 bits (257), Expect = 2e-20
Identities = 47/58 (81%), Positives = 52/58 (89%), Gaps = 0/58 (0%)
Frame = -3

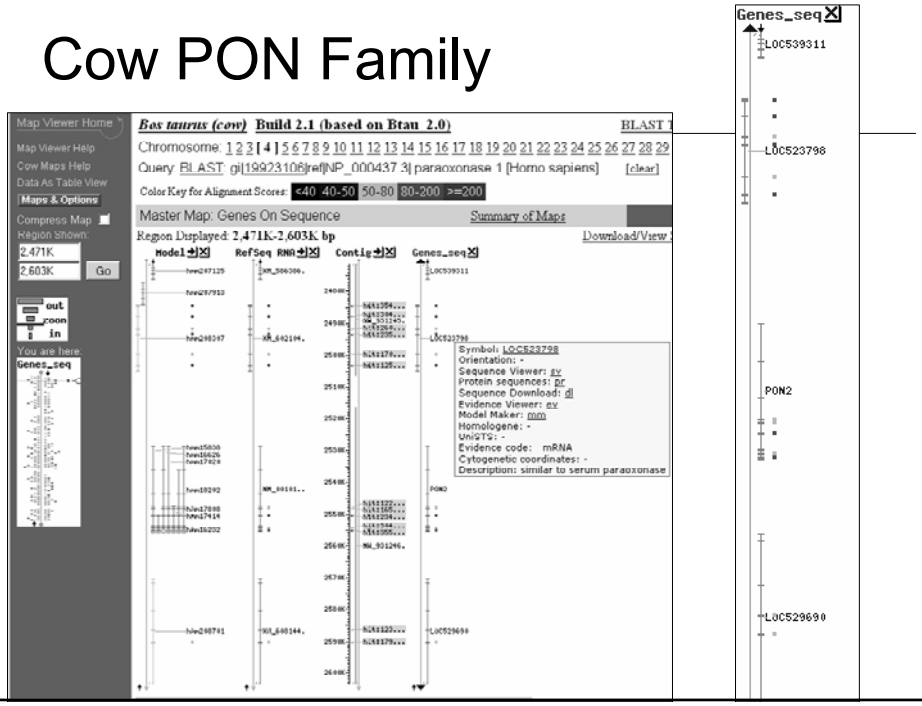
Query 68 GLKYPGIKSFNPNSPGKILLMDLNEEDPTVLELGITGSKFDVSSFNPHGISTFTDEDN 125
GL+YPG+KSF P+ PGKILLMDLN+EDPTVLEL +TGS FD SSFNPHGISTFTDE N
Sbjct 540504 GLRYPGVKSFEPDKPGKILLMDLNKEDPTVLELKMTGSNFDASSFNPHGISTFTDEGN 540331

Features in this part of subject sequence:
similar to paraoxonase 1

Genomic Context of BLAST Hits



Cow PON Family



Online Resources

- Power Point Slides: ftp.ncbi.nih.gov/pub/FieldGuide/Slides/Current/SOT
- NCBI Homepage: www.ncbi.nlm.nih.gov
- News: ./About/newsletter.html
- Handbook: ./books/bv.fcgi?rid=handbook

Contact Us

- General Help: info@ncbi.nlm.nih.gov
- BLAST Help: blast-help@ncbi.nlm.nih.gov
- Hotline: 301-496-2475

Selected References

- Wheeler DL, et al. 2005. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **33**(Database issue):D39-45. PMID: 15608222
- Benson DA, *et al*. 2005. Genbank. *Nucleic Acids Res.* **33**(Database issue):D34-38 PMID: 15608212
- Pruitt KD *et al.,.* 2005 NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **33**(Database issue):D501-4. PMID: 15608248.
- McGinnis S, Madden TL. 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.* **32**(Web Server issue):W20-5. PMID: 15215342
- Maglott D, Ostell J, Pruitt KD, Tatusova T. 2005. Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res*. **33**(Database issue):D54-8. PMID: 15608257