

DNA配列の中から遺伝子を探す

BLAST finds regions of similarity between biological sequences. [more...](#)

**New** Designing or Testing PCR Primers? Try your search in **Primer-BLAST**.

#### Align Sequences with BLAST

A new BL2seq functionality has been added to the standard BLAST pages that allows you to align a query against a set of subject sequences.  
2008-09-04 12:56:52

[More BLAST news...](#)

#### Tip of the Day

How to Search Custom Databases  
[ntrez Queries.](#)

## BLAST Assembled Genomes

Choose a species genome to search, or [list all genomic BLAST databases.](#)

- [Human](#)
- [Mouse](#)
- [Rat](#)
- [Arabidopsis thaliana](#)
- [Oryza sativa](#)
- [Bos taurus](#)
- [Danio rerio](#)
- [Drosophila melanogaster](#)
- [Gallus gallus](#)
- [Pan troglodytes](#)
- [Microbes](#)
- [Apis mellifera](#)

## Basic BLAST

Choose a BLAST program to run.

- ▶ nucleotide blast** Search a **nucleotide** database using a **nucleotide** query  
*Algorithms: blastn, megablast, discontinuous megablast*
- protein blast** Search **protein** database using a **protein** query  
*Algorithms: blastp, psi-blast, phi-blast*
- blastx** Search **protein** database using a **translated nucleotide** query
- tblastn** Search **translated nucleotide** database using a **protein** query
- tblastx** Search **translated nucleotide** database using a **translated nu**

Query

データベース

DNA塩基配列	→	DNA塩基配列
アミノ酸配列	→	アミノ酸配列
DNA塩基配列をアミノ酸配列に変換	→	アミノ酸配列
アミノ酸配列	→	DNA塩基配列をアミノ酸配列に変換
DNA塩基配列をアミノ酸配列に変換	→	DNA塩基配列をアミノ酸配列に変換

## Specialized BLAST

Choose a type of specialized search (or database name in parentheses.)

- Make specific primers with [Primer-BLAST](#)
- Search [trace archives](#)
- Find [conserved domains](#) in your sequence (cds)
- Find sequences with similar [conserved domain architecture](#) (cdart)
- Search sequences that have [gene expression profiles](#) (GEO)
- Search [immunoglobulins](#) (IgBLAST)
- Search for [SNPs](#) (snp)
- Screen sequence for [vector contamination](#) (vecscreen)
- [Align](#) two sequences using BLAST (bl2seq)
- Search [protein](#) or [nucleotide](#) targets in PubChem BioAssay

1. TATGGCTTA----  
T G L

2. TATGGCTTA----  
M A -

3. TATGGCTTA----  
W L -

cDNA配列やアミノ酸配列があれば、Blast検索で有用な情報が得られるが、遺伝子があるかも分からない長いDNA配列を検索しても有用な情報が得られない。

そこで、長いDNA配列の中にどこに遺伝子が隠されているかをまず探し出す。

# DNA配列の中から遺伝子を予測する

## 代表的な遺伝子予想ソフト

Softberry/FGENESH: <http://linux1.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind>

GeneMark: <http://opal.biology.gatech.edu/GeneMark/eukhmm.cgi>

GENSCAN : <http://genes.mit.edu/GENSCAN.html>

## リンク集

### データベース検索:

1. [PubMed](#): 論文検索
2. [Entrez](#): 総合データベース
3. [Google Scholar](#): 文献データベース
4. [特許データベース](#)

### ホモロジー検索:

1. [BLAST](#) [GenomeNET]
2. [FASTA](#) [GenomeNET]

### 配列解析:

1. [DNA → AA](#): DNA配列をアミノ酸配列に変換
2. [Spidey](#): DNAとcDNAのマッチング
3. [BLAST\(NCBI\)](#): 遺伝子のホモロジー検索
4. [BLAST\(DDBJ\)](#): 遺伝子のホモロジー検索
5. [Softberry/FGENESH](#): 遺伝子の予測
6. [GeneMark](#): 遺伝子の予測
7. [GENESCAN](#): 遺伝子の予測
8. [SSPN](#): スプライシングの予測
9. [ORF Finder](#): ORFの予測
10. [clustalW](#): アライメント
11. [TFSEARCH](#): 転写因子結合部位の予測
12. [MOTIF](#): タンパク質モチーフ検索
13. [pI/Mw](#): 等電点、分子量の予測

### 解析用配列

[ある遺伝子AのgenomeDNA配列](#)

[ある遺伝子AのcDNA配列](#)

[ある遺伝子XのgenomeDNA配列](#)

[ある遺伝子XのcDNA配列](#)

[ある遺伝子YのgenomicDNA](#)

[ある遺伝子YのcDNA](#)



TEST ON LINE

- GENE FINDING in Eukaryota
- GENE FINDING WITH SIMILARITY
- OPERON AND GENE FINDING IN BACTERIA
- GENE FINDING IN VIRUSES
- ALIGNMENT /Sequences&genomes
- GENOME EXPLORER /Infogene
- SEARCH FOR MOTIFS /promoters&functional
- PROTEIN LOCATION /patterns/Epitops
- RNA STRUCTURE COMPUTING
- PROTEIN STRUCTURE
- PROTEIN / DNA 3D-Visual Works
- SEQMAN
- MULTIPLE ALIGNMENTS
- CLUSTERING ESTs
- ANALYSIS OF EXPRESSION DATA
- HUMAN-MOUSE-RAT SYNTENY
- PLANT PROMOTERS DATABASE
- REPEATS /find&map repeats
- SNP Extracting known SNPs

FGENESH

HMM-based gene structure prediction (multiple genes, both chains)

Paste nucleotide sequence here:

```
ccctgtaaaatagtcaccattcgaatcgaataaactactagcagtgtaggtactattgctggaactgaattgattgca  
gtaaagggtgaatagcagccagatgaaggtttacaattgaacatcgaacgaaatccacacgaaattcataaatt  
catcgcatcttgcgaagctaccagattctatcgattcagctcagccaataggtcagggg
```

ある遺伝子YのGenomicDNA

Alternatively, load a local file with sequence in Fasta format:

Local file name:  ファイル...いません

- Organism:  Bos taurus  Chicken  Fish  Frog (Xenopodinae)  Human  Mouse
- Anopheles gambiae  Culex  Drosophila  Honey Bee  Tribolium (red flour beetle)
- Brugia malayi (parasitic nematode)  C.elegans  Sea urchin
- Diatom  Plasmodium falciparum  Phytophthora  Monocot plants (Corn, Rice, Wheat, Barley)  Nicotiana tabacum
- Dicot plants (Arabidopsis)  Medicago (legume plant)  Tomato  Vitis vinifera
- Chlamydomonas (single celled green algae)
- Aspergillus  Batrachochytrium  Botrytis  Coccidioides immitis  Coprinopsis cinerea  Cryptococcus neoformans
- Fusarium graminearum  Histoplasma (fungus)  Leptosphaeria  Magnaporthe  Neurospora crassa  Paracocci
- Phanerochaete chrysosporium (white rot)  Pyrenophora  Rhizopus\_oryzae  Schizosaccharomyces pombe
- Sclerotinia sclerotiorum  Stagnospora nodorum  Uncinocarpus reesii  Ustilago

[\[Help\]](#) [\[Show advanced options\]](#)  
[\[Example: Homo sapiens genomic beta globin region \(HBB@\) on chromosome 11\]](#)  
[\[Example: Search in -chain\]](#)

Return to page with other programs of group: [Gene finding](#)

Most gene finding parameters presented here were trained by Softberry for its own use and distribution, using proprietary and publicly available data. Some of the parameters were

Show picture of predicted genes in PDF file

FGENSEH 2.6 Prediction of potential genes in Monocot genomic DNA  
 Time : Fri Nov 7 23:19:44 2008  
 Seq name: test sequence  
 Length of sequence: 3420  
 Number of predicted genes 1: in +chain 1, in -chain 0.  
 Number of predicted exons 10: in +chain 10, in -chain 0.  
 Positions of predicted genes and exons: Variant 1 from 1, Score:173.436011

G Str	Feature	Start	End	Score	ORF	Len
1 +	TSS	196		-1.98		
1 +	1 CDSf	337 -	370	6.03	337 -	369 33
1 +	2 CDSi	469 -	605	27.00	471 -	605 135
1 +	3 CDSi	718 -	764	9.62	718 -	762 45
1 +	4 CDSi	1373 -	1698	39.73	1374 -	1697 324
1 +	5 CDSi	1811 -	1893	22.50	1813 -	1893 81
1 +	6 CDSi	1996 -	2071	9.65	1996 -	2070 75
1 +	7 CDSi	2166 -	2227	16.29	2168 -	2227 60
1 +	8 CDSi	2326 -	2421	7.16	2326 -	2421 96
1 +	9 CDSi	2552 -	2713	29.19	2552 -	2713 162
1 +	10 CDSl	2850 -	2966	14.09	2850 -	2966 117
1 +	Pola	3224		0.44		

Predicted protein(s):

>FGENSEH: [mRNA] 1 10 exon (s) 337 - 2966 1140 bp. chain +

```

ATGGCGACAGCCGGGAAGGTGATCAAGTGC AAAGCGGCGGTGGCATGGGAGGCCGGGAAG
CCGCTGTTCGATCGAGGAGGTGGAGGTTGCGCCGCGCAGGCCATGGAGTCCGCGTCAAG
ATCCTCTACACCCCTCTGCCACACCGACGTCTACTTCTGGGAGGCCAAGGGGCAAACA
CCTGTTTTCCCTAGGATCTTGGGCCATGAAGCTGGAGGCATGCTGGAGAGTCTGGGAGAG
GGTGTGACCGAACTCGCGCCGGGCGACCATGTCTCCCGGTGTTACCCGGCGAGTGC AAG
GAGTGTGATCACTGCAAAATCGGAGGAGAGCAACATGTGTGACCTCCTCAGGATCAACGTC
GACCGCGCGCTCATGATCGCGGACGGCAAGTCCCGATTACCACATCAAGGGGAAGCCCATC
TTCCACTTTGTTGGCACATCCACCTTCAGCGAGTACACCGTCAATCATGTCGGCTGCCTC
CCGAAGATCAACCCGGAGGCGCCTCTCGACAAGGTCTGCATTCAGCTGGGTTTTCTCT
ACCGGTTTTGGTGC AACAGTGAATGTGCGGAAACCGAAAAGGGCCAGACCGTGGCTATT
TTCGGTCTTGGAGCTGTTGGTCTTGTCTGCTATGGAAGGTGCCAGGCTGCTGGAGCATCG
AGGATCATTTGGTGTGGACCTGAACCTTGC GAAATTCGAACAAGCTAAGAAGTTTGGCTGC
ACTGACTTGGTAAACCCCAAGGACCAAGCAAGCCAGTGCACGAGGTGCTGATGAGATG
ACCAACGGTGGACTCGACAGGCGCGTTGAATGCACGGGCAACATCAACGCCATGATATCC
TGCTTCGAATGTGCTCATGATGGATGGGGCGTTCGGGTGCTGGTGGCGTGC CGACCAAG
GATGATGTGTTCAAGACCCACCCGATGAACCTCCTCAACGAGAAGACGCTCAAGGGGACG
TTCTTCGGTAAC TACAAGCCCGCACCGACCTGCCCAACGTCGAGCTGTATCATGAAG
AAGGAGCTGGAGCTGGAGAAGTTCATCAGGCACAGCGTGCCGTTCTCGGAGATCAACAG
CGCTTCGATCTCATGCTCAAGGGGAGAGTCTCCGCTCGCTCATGAGGATGGATGAGTAG

```

Monocotのモデルで解析

予想されたATG-TAG(CDS)

Wordに保存

>FGENSEH: 1 10 exon (s) 337 - 2966 379 aa. chain +

```

MATAGKVIKCKAAVAWEAGKPLSIEEVEVAPPQAMEVVRKILY TALCHTDVYFWEAKGQT
PVFPRILGHEAGGIVESVGEVTE LAPGDHVLVFTGECCKEDHCKSEESNMCDLLRINV
DRGMVIGDGKSRFTIKGKPIHFVGTSTFSEYTVIHVGLAKINPEAPLDKVICILSCGFS
TGFATVNVAKPKKQTV AIFGLGAVGLAAMEGARLSCASRIIGVDLNPAPKFEQAKKFGK
TDFVNPKDHSKPVHEVLEIMTNGGLDRAVECTGNINAMISCFECVHDGCVAVLVGVPTK
DDVFKTHPMNFLNEKTLKGTFFGNYPRTDLPNVVELYMKKELELEKFI THSVPFSEINT
AFDMLKGESLRCVMRME

```

予想されたアミノ酸

## Organism; Monocot

MATAGKVIKCKAAVAWEAGKPLSIEEVEVAPPQAMEVRVKILYTALCHTDVYFWEAKGQT  
PVFPRILGHEAGGIVESVGEGVTELAPGDHVLVPVFTGECKECDHCKSEESNMCDLLRINV  
DRGVMIGDGKSRFTIKGKPIFHVGTSTFSEYTVIHVGCLAKINPEAPLDKVCILSCGFS  
TGFGATVNVAKPKKGQTVAI FGLGAVGLAAMEGARLSGASRIIGVDLNPAAKFEQAKKFGC  
TDFVNPKDHSKPVHEVLIEMTNGGLDRAVECTGNINAMISCFECVHDGWSVAVLVGVPTK  
DDVFKTHPMNFLNEKTLKGTFFGNYKPRTDLPNVVELYMKKELELEKFI THSVPFSEINT  
AFDLMLKGESLRCVMMRDE





TEST ON LINE

- GENE FINDING in Eukaryota
- GENE FINDING WITH SIMILARITY
- OPERON AND GENE FINDING IN BACTERIA
- GENE FINDING IN VIRUSES
- ALIGNMENT /Sequences&genomes
- GENOME EXPLORER /Infogene
- SEARCH FOR MOTIFS /promoters&functional
- PROTEIN LOCATION /patterns/Epitops
- RNA STRUCTURE COMPUTING
- PROTEIN STRUCTURE
- PROTEIN / DNA 3D-Visual Works
- SEQMAN
- MULTIPLE ALIGNMENTS
- CLUSTERING ESTs
- ANALYSIS OF EXPRESSION DATA
- HUMAN-MOUSE-RAT SYNTENY
- PLANT PROMOTERS DATABASE
- REPEATS



HMM-based gene structure prediction (multiple genes, both chains)

Paste nucleotide sequence here:

```
ccctgtaaattagtagaccattcgaatcgtataaaactagcagtgaggtagtactgctggaactgaattgattgca
gtaaaagggtgaatagcagccagtatgaaggttttacaattgaacatcgaacgaaatccacagaaattcataaatt
catcgcatcttgcgaagctaccagattctatcgaattcgaatcgaatccaataggtcgaagg
```

ある遺伝子YのGenomicDNA

Alternatively, load a local file with sequence in Fasta format:

Local file name:  ファイルを選択  ファイル...いません



Organism:  Bos taurus  Chicken  Fish  Frog (Xenopodinae)  Human  Mouse

- Anopheles gambiae  Culex  Drosophila  Honey Bee  Tribolium (red flour beetle)
- Brugia malayi (parasitic nematode)  C.elegans  Sea urchin
- Diatom  Plasmodium falciparum  Phytophthora
- Dicot plants (Arabidopsis)  Medicago (legume plant)  Monocot plants (Corn, Rice, Wheat, Barley)  Nicotiana tabacum
- Tomato  Vitis vinifera
- Chlamydomonas (single celled green algae)
- Aspergillus  Batrachochytrium  Botrytis  Coccidioides immitis  Coprinopsis cinerea  Cryptococcus neoformans
- Fusarium graminearum  Histoplasma (fungus)  Leptosphaeria  Magnaporthe  Neurospora crassa  Paracocci
- Phanerochaete chrysosporium (white rot)  Pyrenophora  Rhizopus\_oryzae  Schizosaccharomyces pombe
- Sclerotinia sclerotiorum  Stagnospora nodorum  Uncinocarpus reesii  Ustilago

[\[Help\]](#) [\[Show advanced options\]](#)  
[\[Example: Homo sapiens genomic beta globin region \(HBB@\) on chromosome 11\]](#)  
[\[Example: Search in -chain\]](#)

名古屋大学関係 サイエンス ニュース (321) お役立ち

Show picture of predicted genes in PDF file

FGENESH 2.6 Prediction of potential genes in Homo\_sapiens genomic DNA  
 Time : Sat Nov 8 23:13:31 2008  
 Seq name: test sequence  
 Length of sequence: 3420  
 Number of predicted genes 1: in +chain 1, in -chain 0.  
 Number of predicted exons 9: in +chain 9, in -chain 0.  
 Positions of predicted genes and exons: Variant 1 from 1, Score:73.703784

G Str	Feature	Start	End	Score	ORF	Len
1 +	TSS	196		-3.99		
1 +	1 CDSf	537 -	605	15.82	537 -	605 69
1 +	2 CDSi	718 -	764	-4.98	718 -	762 45
1 +	3 CDSi	1373 -	1698	35.16	1374 -	1697 324
1 +	4 CDSi	1811 -	1893	10.51	1813 -	1893 81
1 +	5 CDSi	1996 -	2071	-0.46	1996 -	2070 75
1 +	6 CDSi	2166 -	2227	5.23	2168 -	2227 60
1 +	7 CDSi	2326 -	2421	1.18	2326 -	2421 96
1 +	8 CDSi	2552 -	2713	16.54	2552 -	2713 162
1 +	9 CDSi	2850 -	2966	-0.52	2850 -	2966 117
1 +	PolA	3224		1.12		

Predicted protein(s):

>FGENESH:[mRNA] 1 9 exon (s) 537 - 2966 1038 bp, chain +

```

ATGGAGGTCGCGCTCAAGATCCTCTACACCGCCCTCTGCCACACCGACGCTACTTCTGG
GAGGCCAAGGGGGCAAACACCTGTTTTCCCTAGGATCTTGGGCCATGAAGCTGGAGGCATT
GTGGAGAGTGTGGGAGAGGTGTGACCGAACTCGCGCCGGGGCACCATTGCTCCCGGTG
TTCACCGGCGAGTGAAGGAGTGTGATCACTGCAAATCGGAGGAGAGCAACATGTGTGAC
CTCCTCAGGATCAACGTCGACCGCGCGCTCATGATCGGGCAGCGGCAAGTCCCGATTACCC
ATCAAGGGGAAGCCCATCTTCCACTTTGTGGCACATCCACCTTCAGCGAGTACACCGTTC
ATCCATGTGGCTGCCTCGCGAAGATCAACCCGGAGGCGCCTCTCGACAAGGTCTGCATT
CTCAGCTGCGGTTTCTCTACCGGTTTGGTGCAACAGTGAATGTGCGGAAACCGAAAAAG
GGCCAGACCGTGGTATTTTCGGTCTTGGAGCTGTGGTCTTGCCTATGGAAGGTGCC
AGGCTGTCTGGAGCATCGAGGATCATTTGGTGTGGACCTGAACCCGCGAAATTCGAACAA
GCTAAGAAGTTTGGCTGCACTGACTTCGTAACCCCAAGGACCACAGCAAGCCAGTGCAC
GAGGTGCTGATGAGATGACCAACGGTGGACTCGACAGGGCCGTTGAATGCACCGGCAAC
ATCAACGCCATGATATCCCTGCTTCCGAATGTGTCATGATGGATGGGGCGTTGGCGTGCTG
GTCGGCGTGCCGACCAAGGATGATGTGTTCAAGACCCACCCGATGAACCTTCCCAACGAG
AAGACGCTCAAGGGGACGTTCTTCGGTAACTACAAGCCGGCACCACCTGCCCAACGTC
GTCGAGCTGTACATGAAGAAGGAGCTGGAGCTGGAGAAAGTTCATCACGCACAGCGTGCCG
TTCCTGGAGATCAACACGGCGTTCGATCTCATGCTCAAGGGGGAGAGTCTCCGCTGCGTC
ATCAGGATGGATCAGTAC

```

Humanのモデルで解析

予想されたATG-TAG

```

>FGENESH: 1 9 exon (s) 537 - 2966 345 aa, chain +
MEVRVKILYALCHTDVYFWEAKGQTPVFPRIILGHEAGGIVESVGEVTELPDGHVLPV
FTGECKECDHCKSEESNMCDLLRINVDRGVMIGDKSRFTIKGKPIFHFVGTSTPSEYTV
IHVGLAKINPEAPLDKVCILSCGFSTGFGATVNVAKPKKQTVAFGLGAVGLAAMEGA
RLSGASRIIGVDLNPAPKFEQAKKFGCTDFVNPDKHSPVHEVLIEMTNGGLDRAVECTGN
INAMISCFECVHDGWSVAVLVGVPKDDVFKTHPMNLFNEKTLKGTFFGNYPKPRDLPNV
VELYMKKELELEKPIHSHVFPSEINTAFDMLLKGESLRCVVRMDE

```

予想されたアミノ酸

## Modellによって予想されるORFは多少異なる

Organism; Monocot

MATAGKVIKCKAAVAWEAGKPLSIEEVEVAPPQAMEVRVKILY TALCHTDVYFWEAKGQT  
PVFPRILGHEAGGIVESVGEGVTELAPGDHVLVPVFTGECKECDHCKSEESNMCDLLRINV  
DRGVMIGDGKSRFTIKGKPIHFVGTSTFSEYTVIHVGCLAKINPEAPLDKVCILSCGF  
TGFGATVNVAKPKKGQTVAIFGLGAVGLAAMEGARLSGASRIIGVDL NPAKFEQAKKFGC  
TDFVNP KDHSKPVHEVLIEMTNGGLDRAVECTGNINAMISCFECVHDGWGVAVLVGVPTK  
DDVFKTHPMNFLNEKTLKGTFFGNYKPRTDLPNVVELYMKKELELEK FITHSVPFSEINT  
AFDLMLKGESLRCVMRMDE

Organism; 人間

MEVRVKILY TALCHTDVYFWEAKGQTPVFPRILGHEAGGIVESVGEGVTELAPGDHVLVP  
FTGECKECDHCKSEESNMCDLLRINVDRGVMIGDGKSRFTIKGKPIHFVGTSTFSEYTV  
IHVGCLAKINPEAPLDKVCILSCGFSTGFGATVNVAKPKKGQTVAIFGLGAVGLAAMEGA  
RLSGASRIIGVDL NPAKFEQAKKFGCTDFVNP KDHSKPVHEVLIEMTNGGLDRAVECTGN  
INAMISCFECVHDGWGVAVLVGVPTKDDVFKTHPMNFLNEKTLKGTFFGNYKPRTDLPNV  
VELYMKKELELEK FITHSVPFSEINTAFDLMLKGESLRCVMRMDE

Organism; ショウジョウバエ

MFCGEEKLIRSGFGVFAAAVAWEAGKPLSIEEVEVAPPQAMEVRVKILY TALCHTDVYFW  
EAKGQTPVFPRILGHEAGGIVESVGEGVTELAPGDHVLVPVFTGECKECDHCKSEESNMCD  
LLRINVDRGVMIGDGKSRFTIKGKPIHFVGTSTFSEYTVIHVGCLAKINPEAPLDKVC  
LSCGFSTGFGATVNVAKPKKGQTVAIFGLGAVGLAAMEGARLSGASRIIGVDL NPAKFEQ  
AKKFGCTDFVNP KDHSKPVHEVLIEMTNGGLDRAVECTGNINAMISCFECVHDGWGVAVL  
VGVPTKDDVFKTHPMNFLNEKTLKGTFFGNYKPRTDLPNVVELYMKKELELEK FITHSVP  
FSEINTAFDLMLKGESLRCVMRMDE

Safari ファイル 編集 表示 履歴 ブックマーク ウィンドウ ヘルプ

BLAST: Basic Local Alignment Search Tool

http://blast.ncbi.nlm.nih.gov/Blast.cgi

名古屋大学関係 ▼ サイエンス ▼ ニュース (321) ▼ お役立ち ▼

**BLAST** Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

My NCBI [Sign In] [Register]

▶ NCBI/ BLAST Home

BLAST finds regions of similarity between biological sequences. [more...](#)

**New** Designing or Testing PCR Primers? Try your search in **Primer-BLAST**.

### BLAST Assembled Genomes

Choose a species genome to search, or [list all genomic BLAST databases](#).

- [Human](#)
- [Mouse](#)
- [Rat](#)
- [Arabidopsis thaliana](#)
- [Oryza sativa](#)
- [Bos taurus](#)
- [Danio rerio](#)
- [Drosophila melanogaster](#)
- [Gallus gallus](#)
- [Pan troglodytes](#)
- [Microbes](#)
- [Apis mellifera](#)

### Basic BLAST

Choose a BLAST program to run.

<a href="#">nucleotide blast</a>	Search a <b>nucleotide</b> database using a <b>nucleotide</b> query <i>Algorithms:</i> blastn, megablast, discontinuous megablast
<a href="#">protein blast</a>	Search <b>protein</b> database using a <b>protein</b> query <i>Algorithms:</i> blastp, psi-blast, phi-blast
<a href="#">blastx</a>	Search <b>protein</b> database using a <b>translated nucleotide</b> query
<a href="#">tblastn</a>	Search <b>translated nucleotide</b> database using a <b>protein</b> query
<a href="#">tblastx</a>	Search <b>translated nucleotide</b> database using a <b>translated nucleotide</b> query

### Specialized BLAST

Choose a type of specialized search (or database name in parentheses.)

- Make specific primers with [Primer-BLAST](#)
- Search [trace archives](#)
- Find [conserved domains](#) in your sequence (cds)
- Find sequences with similar [conserved domain architecture](#) (cdart)
- Search sequences that have [gene expression profiles](#) (GEO)
- Search [immunoglobulins](#) (IgBLAST)
- Search for [SNPs](#) (snp)
- Screen sequence for [vector contamination](#) (vecscreen)
- [Align](#) two sequences using BLAST (bl2seq)
- Search [protein](#) or [nucleotide](#) targets in PubChem BioAssay

### News

[Align Sequences with BLAST](#)

A new BL2seq functionality has been added to the standard BLAST pages that allows you to align a query against a set of subject sequences.  
2008-09-04 12:56:52

[More BLAST news...](#)

### Tip of the Day

**How to Search Custom Databases in Web-Blast Using Entrez Queries.**

A powerful feature of the BLAST Web interface is the ability to limit BLAST searches to a subset of any database using a standard Entrez query. Skillful use of Entrez queries allows the equivalent of on-the-fly construction of databases of exact composition.

[More tips...](#)

NCBI/BLAST/blastp suite: BLASTP programs search protein databases using a protein query. [more...](#) [Reset page](#) [Bookmark](#)

Enter Query Sequence

Enter accession number, gi, or FASTA sequence [Clear](#)  
TGFGATVNVAKPKKGQTVAIFGLGAVGLAAMEGARLSGASRIIGVDLNPQAKFEQAKKFGC  
TDFVNPDKHSKPVHEVLIEMTNGGLDRAVECTGNINAMISCFECVHDGWGVAVLVGVPTK  
DDVFKTHPMNFLNEKTLKGTFFGNYKPRDLPNVVELYMKKELELEKFITHSVPFSEINT  
AFDMLKGESLRCVMMRME

Query subrange  
From   
To

Or, upload file  ファイル... いません  
Job Title   
Enter a descriptive title for your BLAST search

Blast 2 sequences

Choose Search Set

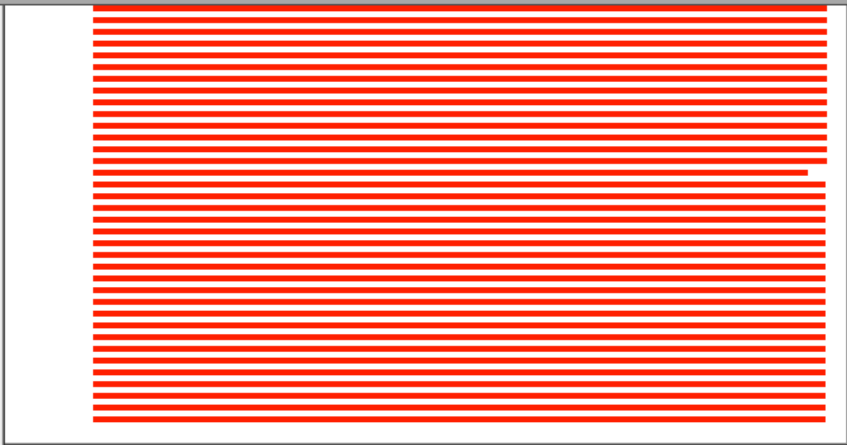
Database   
Organism   
Optional Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown.  
Entrez Query   
Optional

Program Selection

Algorithm  
 blastp (protein-protein BLAST)  
 PSI-BLAST (Position-Specific Iterated BLAST)  
 PHI-BLAST (Pattern Hit Initiated BLAST)  
[Choose a BLAST algorithm](#)

**BLAST** Search database nr using Blastp (protein-protein BLAST)  
 Show results in a new window

[Algorithm parameters](#)



▼ Descriptions

Sequences producing significant alignments:			Score	E
			(Bits)	Value
<a href="#">dbj BAE00035.1</a>	alcohol dehydrogenase [Oryza rufipogon]		778	0.0
<a href="#">sp Q4R1E8.1 ADH2_ORYSI</a>	RecName: Full=Alcohol dehydrogenase 1 ...		778	0.0
<a href="#">dbj BAE00029.1</a>	alcohol dehydrogenase [Oryza rufipogon] >dbj ...		777	0.0
<a href="#">dbj BAE00049.1</a>	alcohol dehydrogenase [Oryza meridionalis]		777	0.0
<a href="#">dbj BAE00044.1</a>	alcohol dehydrogenase [Oryza sativa Indica Gr...		776	0.0
<a href="#">dbj BAE00047.1</a>	alcohol dehydrogenase [Oryza glumipatula]		776	0.0
<a href="#">dbj BAE00048.1</a>	alcohol dehydrogenase [Oryza barthii]		776	0.0
<a href="#">dbj BAE00046.1</a>	alcohol dehydrogenase [Oryza sativa Japonica ...		775	0.0
<a href="#">dbj BAE00025.1</a>	alcohol dehydrogenase [Oryza rufipogon]		775	0.0
<a href="#">dbj BAE00042.1</a>	alcohol dehydrogenase [Oryza rufipogon]		775	0.0
<a href="#">dbj BAE00039.1</a>	alcohol dehydrogenase [Oryza rufipogon]		774	0.0
<a href="#">dbj BAE00043.1</a>	alcohol dehydrogenase [Oryza rufipogon]		774	0.0
<a href="#">qb EAZ17793.1</a>	hypothetical protein OsJ_032002 [Oryza sativa ...		773	0.0
<a href="#">dbj BAE00028.1</a>	alcohol dehydrogenase [Oryza rufipogon]		773	0.0
<a href="#">qb EAY80334.1</a>	hypothetical protein OsI_034293 [Oryza sativa ...		763	0.0
<a href="#">ref NP_001067485.1</a>	Os1lg0210500 [Oryza sativa (japonica cult...		761	0.0
<a href="#">qb ABY76197.1</a>	alcohol dehydrogenase 2 [Oryza coarctata]		739	0.0
<a href="#">qb ACG28464.1</a>	unknown [Zea mays]		736	0.0
<a href="#">ref NP_001105410.1</a>	alcohol dehydrogenase 2 [Zea mays] >sp P04...		732	0.0
<a href="#">qb ABL74253.1</a>	alcohol dehydrogenase ADH3D [Triticum aestivum]		728	0.0
<a href="#">qb ABL74256.1</a>	alcohol dehydrogenase ADH2B [Triticum aestivum]		724	0.0
<a href="#">qb ABL74255.1</a>	alcohol dehydrogenase ADH2D [Triticum aestivum]		724	0.0
<a href="#">sp P10848.1 ADH3_HORVU</a>	RecName: Full=Alcohol dehydrogenase 3 ...		723	0.0
<a href="#">qb ABL74260.1</a>	alcohol dehydrogenase ADH2 [Aegilops tauschii]		723	0.0
<a href="#">emb CAA34364.1</a>	alcohol dehydrogenase 2 [Oryza sativa (indica...		722	0.0
<a href="#">qb ABL74254.1</a>	alcohol dehydrogenase ADH2A [Triticum aestivum]		721	0.0
<a href="#">qb ABL74266.1</a>	alcohol dehydrogenase ADH2A [Triticum turgidum...		721	0.0
<a href="#">qb ABL74252.1</a>	alcohol dehydrogenase ADH3-1 [Triticum aestivum]		721	0.0
<a href="#">qb ABL74268.1</a>	alcohol dehydrogenase ADH3 [Aegilops tauschii]		719	0.0
<a href="#">qb ABL74267.1</a>	alcohol dehydrogenase ADH2B [Triticum turgidum...		718	0.0
<a href="#">qb ABL74265.1</a>	alcohol dehydrogenase ADH2 [Triticum monococcum]		718	0.0
<a href="#">qb ABL74257.1</a>	alcohol dehydrogenase ADH2H [Triticum aestivum]		715	0.0
<a href="#">qb ABL74270.1</a>	alcohol dehydrogenase ADH2 [Aegilops speltoides]		714	0.0

UG

UG

## 代表的な遺伝子予想ソフト

Softberry/FGENESH: <http://linux1.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind>

GeneMark: <http://opal.biology.gatech.edu/GeneMark/eukhmm.cgi>

GENSCAN : <http://genes.mit.edu/GENSCAN.html>

## リンク集

### データベース検索:

1. [PubMed](#): 論文検索
2. [Entrez](#): 総合データベース
3. [Google Scholar](#): 文献データベース
4. [特許データベース](#)

### ホモロジー検索:

1. [BLAST](#) [GenomeNET]
2. [FASTA](#) [GenomeNET]

### 配列解析:

1. [DNA → AA](#): DNA配列をアミノ酸配列に変換
2. [Spidey](#): DNAとcDNAのマッチング
3. [BLAST\(NCBI\)](#): 遺伝子のホモロジー検索
4. [BLAST\(DDBJ\)](#): 遺伝子のホモロジー検索
5. [Softberry/FGENESH](#): 遺伝子の予測
6. [GeneMark](#): 遺伝子の予測
7. [GENESCAN](#): 遺伝子の予測
8. [SSPN](#): スプライシングの予測
9. [ORF Finder](#): ORFの予測
10. [clustalW](#): アライメント
11. [TFSEARCH](#): 転写因子結合部位の予測
12. [MOTIF](#): タンパク質モチーフ検索
13. [pI/Mw](#): 等電点、分子量の予測

### 解析用配列

[ある遺伝子AのgenomeDNA配列](#)

[ある遺伝子AのcDNA配列](#)

[ある遺伝子XのgenomeDNA配列](#)

[ある遺伝子XのcDNA配列](#)

[ある遺伝子YのgenomicDNA](#)

[ある遺伝子YのcDNA](#)



## Eukaryotic GeneMark.hmm(1,2) [\(Reload this page\)](#)

### References:

- <sup>1</sup>Borodovsky M. and Lukashin A. (unpublished)
- <sup>2</sup>Lomsadze A., Ter-Hovhannisyan V., Chernoff Y. and Borodovsky M., "Gene identification in novel eukaryotic genomes by self-training algorithm", **Nucleic Acids Research**, 2005, Vol. 33, No. 20, 6494-6506

[Accuracy comparison](#)

**UPDATE October 2005.** Added pre-built models of eukaryotic GeneMark.hmm ES-3.0 (E - eukaryotic; S - self-training; 3.0 - the version)

[Listing of previous updates](#)

### Input Sequence

Title (optional):

Sequence:

Sequence File upload:

ファイルが選...ていません

Species:  [Model description](#)

### Output Options

Email Address: (required for graphical output or sequences longer than 400000 bp)

- Generate PDF graphics (screen)
- Generate PostScript graphics (email)
- Print GeneMark 2.4 predictions in addition to GeneMark.hmm predictions
- Translate predicted genes into protein

Run

Web pages maintained by [GeneMark administrator, alexl@amber.gatech.edu](mailto:alexl@amber.gatech.edu). Please send any suggestions for improvements or problems to the web page maintainer.

### Eukaryotic GeneMark.hmm(1,2) [\(Reload this page\)](#)

#### References:

- <sup>1</sup>Borodovsky M. and Lukashin A. (unpublished)
- <sup>2</sup>Lomsadze A., Ter-Hovhannisyanyan V., Chernoff Y. and Borodovsky M., "Gene identification in novel eukaryotic genomes by self-training algorithm", **Nucleic Acids Research**, 2005, Vol. 33, No. 20, 6494-6506

[Accuracy comparison](#)

**UPDATE October 2005.** Added pre-built models of eukaryotic GeneMark.hmm ES-3.0 (E - eukaryotic; S - self-training; 3.0 - the version)

[Listing of previous updates](#)

#### Input Sequence

Title (optional):

GenomicDNAY

Sequence:

```

cgtcgtcgagctgtacatgaagaaggtaaaaataattgcagaaacagaatcagattgcaacctgaatgctcattg
ttcagaaatctatggtcagtgattccaatgagccactgattgacactatttatgctaagtattctggttttgatg
ctgaaggagctggagctggagaaattcatcacgcacagcgtgctcgttctcgagatcaacacgcgcttctgatc
catgctcaagggggagagctcctcgtcgtcatgaggatgtagtagatgatgatctctctctctctctctctct
ctctctctgagtttcagtcatacaagaacaaatgaaagcaaaaatggatgagaagtttgaggttttcc
atgaaatttaataatggtcgtgagacagggctgtgtgagtcagctctgtttactcctcttcaatgctgatct
gtttattgtagcattgtgcttagcagattttccagtttaccctgtaaattagaccattctgaaatcgtataaa
actactagcagtgtaggactattgctggaactgaattgagtcagtaaaagaggtgaatagcagccagatg
aaggttttacaatttgaaatcgaaacgaaatccacgaaatcataaattcatcgactcttctgctgaagcttacc
gattctatcgaattcagatcagctccaataggtcagatggg

```

Sequence File upload:

ファイルを選択 ファイルが選...ていません

Species:  A.gambiae ES-3.0

- A.thaliana ES-3.0
- A.thaliana
- C.briggsae ES-3.0
- C.elegans ES-3.0
- C.elegans
- C.intestinalis ES-3.0
- C.reinhardtii ES-3.0
- C.remanei ES-3.0
- D.melanogaster ES-3.0
- D.melanogaster
- G.gallus
- H.sapiens
- H.vulgare
- M.musculus
- M.truncatula ES-3.0
- O.sativa
- T.aestivum
- T.gondii ES-3.0
- Z.mays

[Model description](#)

Output C

Email Ad  (graphical output or sequences longer than 400000 bp)

General (screen)

General (graphics (email))

Print (printouts in addition to GeneMark.hmm predictions)

Trans (translate into protein)

Run

Start GeneMark.hmm

## Eukaryotic GeneMark.hmm(1,2) [\(Reload this page\)](#)

### References:


- <sup>1</sup>Borodovsky M. and Lukashin A. (unpublished)
- <sup>2</sup>Lomsadze A., Ter-Hovhannisyan V., Chernoff Y. and Borodovsky M., "Gene identification in novel eukaryotic genomes by self-training algorithm", **Nucleic Acids Research**, 2005, Vol. 33, No. 20, 6494-6506

### [Accuracy comparison](#)

**UPDATE October 2005.** Added pre-built models of eukaryotic GeneMark.hmm ES-3.0 (E - eukaryotic; S - self-training; 3.0 - the version)

### [Listing of previous updates](#)

### Input Sequence

Title (optional): 

GenomicDNAY

### Sequence:


```
cgctcgcgagctgtacatgaagaaggtaaaataattgcagaacagaatcagattgcaacctgaatgctcattg
ttcagaatctatggtcagtgattccaatgagccaactgattgatcactatattatgctaagattctggtttgatg
ctgaaggagctggagctggagaagtcatcagcagcagcgtgcccgtctcggagatcaacacggcgttcgatct
catgctcaaggggagagatctcctgctgcatgaggatggatgagatgagatgatgatctatctctctctct
ctctctctgcaagtttcaagtcatacaagaacaaaaatgaaagcaaaaaatggatgagaagtggaggtttcc
attaaagtttaataatgctgtgagacagggcttggtggagtcagctctgtttactcctcttcatgcttgatc
tgtttattgtagcattggtgcttagcagattttcacagttttacccctgtaaattagtagcattctgaaatgtaataa
actactagcagtgtaggtagctattgctggaactgaattgattgtagtaagaggtgaatgagcagccagtagt
aaggtttacaattgaaacatcgaaacgaaatccacagaaattcataaattcatcgcattcttctggaagcttacc
gattctatcgaattcagatcagctccaataggtcagggg
```




### Sequence File upload:

[ファイルを選択](#) ファイルが選...ていません

Species:  [Model description](#)

### Output Options

Email Address: (required for graphical output or sequences longer than 400000 bp) 

- Generate PDF graphics (screen)
- Generate PostScript graphics (email) 
- Print GeneMark 2.4 predictions in addition to GeneMark.hmm predictions 
- Translate predicted genes into protein 

Run

Default

Start GeneMark.hmm



[Listing of previous updates](#)

## Result of last submission:

[View PDF Graphical Output](#)

### GeneMark.hmm Listing

[Go to: GeneMark.hmm Protein Translations](#)[Go to: Job Submission](#)


Eukaryotic GeneMark.hmm version bp 3.8h August 16, 2007  
Sequence name: GenomicDNAY  
Sequence length: 3420 bp  
G+C content: 45.06%  
Matrices file: /home/genmark/euk\_ghm.matrices/rice\_hmm3.0mod  
Fri Nov 7 22:35:27 2008

#### Predicted genes/exons


Gene #	Exon #	Strand	Exon Type	Exon Range	Exon Length	Start/End Frame
1	1	+	Initial	337 370	34	1 1 --
1	2	+	Internal	469 605	137	2 3 --
1	3	+	Internal	718 764	47	1 2 --
1	4	+	Internal	1373 1698	326	3 1 --
1	5	+	Internal	1811 1893	83	2 3 --
1	6	+	Internal	1996 2071	76	1 1 --
1	7	+	Internal	2166 2227	62	2 3 --
1	8	+	Internal	2326 2421	96	1 3 --
1	9	+	Internal	2552 2713	162	1 3 --
1	10	+	Terminal	2850 2966	117	1 3 --

```
>gene_1|GeneMark.hmm|379_aa
MATAGKVIKKAAVAWEACKPLSIEEVEVAPPQAMEVRVKILYALCHTDVYFWEAKGQT
PVFPRIILGHEAGGIVESVGEVTELPDGHVLPVFTGECKEDHCKSEESNMCDLLRINV
DRGVMIGDKSRFTIKGKPIHFVGTSTFSEYTVIHVGLAKINPEAPLDKVCILSCGFS
TCFGATVNVAKPKKQTVAFGLGAVGLAAMEGARLSCASRIIGVDLNPAKFEQAKKFGC
TDFVNPKDHSKPVHEVLIEMTNGGLDRAVECTGNINAMISCFECVHDGQVAVLVGVPTK
DDVFKTHPMNLFNEKTLKGTFFGNYKPRDLPNVVELYMKKELELEKFI THSVPFSEINT
AFDLMLKGESLRCVMMRME
```

#### Input Sequence

Title (optional): 

GenomicDNAY

Sequence: 

```
atccgcccgaaccggaaaaaccacccaagctcgaaccggcggaggcaagcagcagcatgcagcag
acgcagacctctctcttccaagcaaaaaatcagatgacgcctctcttccccctcactgctt
```

## 予想ソフトによっても予想されるORFは多少異なる

### Softberryによる予想、Organism; Monocot

MATAGKVIKCKAAVAWEAGKPLSIEEVEVAPPQAMEVRVKILYALCHTDVYFWEAKGQT  
PVFPRILGHEAGGIVESVGEGVTELAPGDHVLVPVFTGECKECDHCKSEESNMCDLLRINV  
DRGVMIGDGKSRFTIKGKPIHFVGTSTFSEYTVIHVGCLAKINPEAPLDKVCILSCGFS  
TGFGATVNVAKPKKGQTVAFGLGAVGLAAMEGARLSGASRIIGVDLNPAAKFEQAKKFGC  
TDFVNPDKHSPVHEVLIEMTNGGLDRAVECTGNINAMISCFECVHDGWGVAVLVGVPTK  
DDVFKTHPMNFLNEKTLKGTFFGNYKPRTDLPNVVELYMKKELELEKFITHSVPFSEINT  
AFDLMLKGESLRCVMMRDE

### GeneMarkによる予想、Organism; イネ

MATAGKVIKCKAAVAWEAGKPLSIEEVEVAPPQAMEVRVKILYALCHTDVYFWEAKGQTPVFPR  
ILGHEAGGIVESVGEGVTELAPGDHVLVPVFTGECKECDHCKSEESNMCDLLRINVDRGVMIGDGK  
SRFTIKGKPIHFVGTSTFSEYTVIHVGCLAKINPEAPLDKVCILSCGFSTGFGATVNVAKPKKGQ  
TVAFGLGAVGLAAMEGARLSGASRIIGVDLNPAAKFEQAKKFGCTDFVNPDKHSPVHEVLIEMT  
NGGLDRAVECTGNINAMISCFECVHDGWGVAVLVGVPTKDDVFKTHPMNFLNEKTLKGTFFGNY  
KPRTDLPNVVELYMKKELELEKFITHSVPFSEINTAFDLMLKGESLRCVMMRDE

## 代表的な遺伝子予想ソフト

Softberry/FGENESH: <http://linux1.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind>

GeneMark: <http://opal.biology.gatech.edu/GeneMark/eukhmm.cgi>

GENSCAN: <http://genes.mit.edu/GENSCAN.html>

## リンク集

### データベース検索:

1. [PubMed](#): 論文検索
2. [Entrez](#): 総合データベース
3. [Google Scholar](#): 文献データベース
4. [特許データベース](#)

### ホモロジー検索:

1. [BLAST](#) [GenomeNET]
2. [FASTA](#) [GenomeNET]

### 配列解析:

1. [DNA → AA](#): DNA配列をアミノ酸配列に変換
2. [Spidey](#): DNAとcDNAのマッチング
3. [BLAST\(NCBI\)](#): 遺伝子のホモロジー検索
4. [BLAST\(DDBJ\)](#): 遺伝子のホモロジー検索
5. [Softberry/FGENESH](#): 遺伝子の予測
6. [GeneMark](#): 遺伝子の予測
7. [GENESCAN](#): 遺伝子の予測
8. [SSPN](#): スプライシングの予測
9. [ORF Finder](#): ORFの予測
10. [clustalW](#): アライメント
11. [TFSEARCH](#): 転写因子結合部位の予測
12. [MOTIF](#): タンパク質モチーフ検索
13. [pI/Mw](#): 等電点、分子量の予測

### 解析用配列

[ある遺伝子AのgenomeDNA配列](#)

[ある遺伝子AのcDNA配列](#)

[ある遺伝子XのgenomeDNA配列](#)

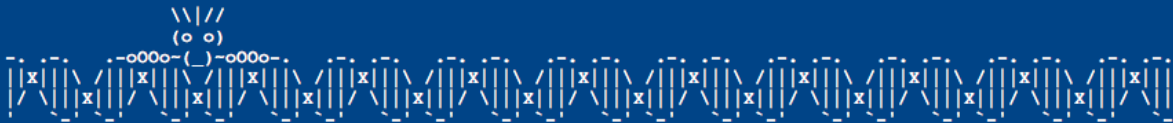
[ある遺伝子XのcDNA配列](#)

[ある遺伝子YのgenomicDNA](#)

[ある遺伝子YのcDNA](#)

# The New GENSCAN Web Server at MIT

## Identification of complete gene structures in genomic DNA



[For information about Genscan, click here](#)

This server provides access to the program Genscan for predicting the locations and exon-intron structures of genes in genomic sequences from a variety of organisms.

This server can accept sequences up to 1 million base pairs (1 Mbp) in length. If you have trouble with the web server or if you have a large number of sequences to process, request a local copy of the program (see bottom of this page) or use the [GENSCAN email server](#). If your browser (*e.g.*, Lynx) does not support file upload or multipart forms, use the [older version](#).

Organism:  Suboptimal exon cutoff (optional):

Sequence name (optional):

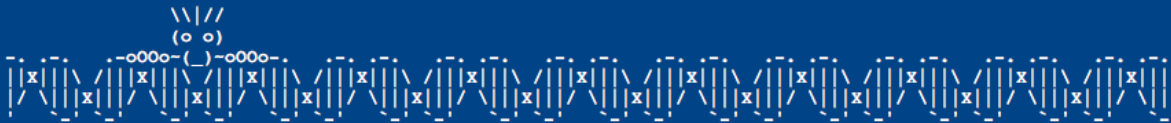
Print options:

Upload your DNA sequence file (one-letter code, upper or lower case, spaces/numbers ignored):  ファイルが選...ていません

Or paste your DNA sequence here (one-letter code, upper or lower case, spaces/numbers ignored):

To have the results mailed to you, enter your email address here (optional):





**For information about Genscan, click here**

This server provides access to the program Genscan for predicting the locations and exon-intron structures of genes in genomic sequences from a variety of organisms.

This server can accept sequences up to 1 million base pairs (1 Mbp) in length. If you have trouble with the web server or if you have a large number of sequences to process, request a local copy of the program (see in bottom of this page) or use the [GENSCAN email server](#). If your browser (e.g., Lynx) does not support file upload or multipart forms, use the [older version](#).

Organism:  optimal exon cutoff (optional):

Sequence name (optional):

Print options:

Upload your DNA sequence file (one-letter code, upper or lower case, spaces/numbers ignored):  ファイルが選...ていません

Or paste your DNA sequence here (one-letter code, upper or lower case, spaces/numbers ignored):

```
atgtgtccatgatgtcagtccttttcaaaaatgaaacatgatagtaacaaattgaagatgacattcagcaatgcctcagttctg
ctgtaaaactgtatcctgaaactaagagtgcaactctgaaacctgaatccaggatggggcgttgctggtcggcgtgccga
ccaaggatgatgttcaagaccaccgatgaactctcaacgagaagacgctcaaggggagcttctcgtaactacaagccg
cgcaccgacctgcccaacgctcgtcgagctgtacatgaagaagtaaaataatgacagaacagaatcagattgcaactgaatg
gcattgttcagaaactatggtcagtgattccaatgagccctgattgatcactatattatgctaagatttcgggtttgatgctgaag
gagctggagctggagaagttcatcacgcaacgctgcccgttctcggagatcaaacacggcgttcgatcctcatgctcaagggggaga
gtctcgtcgtcatgaggatggatgagtagatgatgatctctctctctctctctctctcagtttcagtgatacaagaa
accaaaaatgaaagcaaaaatgagtagaagtttgagttttcatttaagtttaaatgctgctgagacagggcgttggtggag
cagctctgttaactcctcctttcatggttgatctgtttttttgctcagcattgtgcttagcagattttaccagtttaaccgtgaaatag
tacattctgaaactgtaataaactactagcagtgtaggtactattctggaactgaattgctgacagtaaaagggtaatagca
gccagatgaaagttttacaattgaaatcgaacgaacacgaaattcataaattcagctcatcttgctgaagcttacagat
tctatcgattcgatcagctccaataggtcgatggg
```

To have the results mailed to you, enter your email address here (optional):

[Back to the top](#)

```

GENSCAN 1.0      Date run: 8-Nov-108      Time: 00:08:48

Sequence GenomicDNA : 3420 bp : 45.06% C+G : Isochore 1 ( 0 - 100 C+G%)

Parameter matrix: Maize.smat

Predicted genes/exons:

Gn.Ex Type S .Begin ..End .Len Fr Ph I/Ac Do/T CodRg P.... Tscr..
-----
1.00 Prom +   193   232   40
1.01 Init +   337   370   34  0  1   63   73   46 0.998   3.09
1.02 Intr +   469   605  137  2  2   22   86  341 0.999  30.55
1.03 Intr +   718   764   47  0  2   72   75  -17 0.999  -4.97
1.04 Intr +  1373  1698  326  2  2   75  111  572 0.994  57.71
1.05 Intr +  1811  1893   83  0  2   99   87   17 0.936   4.70
1.06 Intr +  1996  2071   76  0  1   20   82   50 0.860  -0.71
1.07 Intr +  2166  2227   62  1  2   97   73   39 0.920   4.27
1.08 Intr +  2326  2421   96  0  0   51  21   98 0.834   2.33
1.09 Intr +  2552  2713  162  1  0   -4   93  446 0.999  39.19
1.10 Term +  2850  2966  117  2  0   27   46  277 0.999  18.91
1.11 PlyA +  3224  3229    6

```

Click [here](#) to view a PDF image of the predicted gene(s)

Click [here](#) for a PostScript image of the predicted gene(s)

Predicted peptide sequence(s):

Predicted coding sequence(s):

>GenomicDNA|GENSCAN predicted peptide 1|379 aa

```

MATAGKVIKCKAAVAWEAGKPLSIEEVEVAPPQAMEVRVKILYTALCHTDVYFWEAKGQT
PVFPRI LGHEAGGIVESVGEVTELAPGDHVLVVFTECKEBCDRCKSEESNMCDLLRINV
DRGVMIGDGKSRFTIKGKPIFHFVGTSTFSEYTVIHVGCLAKINPEAPLDPKVCILSCGFS
TGFATVNVAKPKKGQTVAIFGLGAVGLAAMEGARLSGASRIIGVDLNPAPKFEQAKKFGC
TDFVNPKDHSKPVHEVLIEMTNGGLDRAVECTGNINAMISCFECVHDGCVAVLVGVPTK
DDVFKTHPMNLFNEKTLKGTFFGNYKPRTDLPNVVELYMKKELELEKFI THSVPFSEINT
AFDLMLKGESLRCVMRMDE

```

>GenomicDNA|GENSCAN predicted CDS 1|1140 bp

```

atggcgacagccgggaaggatcaagtgcaaaagcggcggtggcatgggagccgggaag
ccgctgtcgatcgaggagggtggaggtgcccgcgcagccatggaggctccgctcaag
atcctctacaccgcctctgcccacaccgacgtctacttctgggagccaaggggcaaca
cctgttttcctaggatcttgggcatgaagctggaggcattgtggagagtgaggagag
ggtgtgaccgaactcgcggcgaccatgtcctcccgtgttcaccggcgagtgcaag
gagtgatcactgcaaatcggaggagagcaaacatgtgtgacctcctcaggatcaacgtc
gaccggcgctgatcgccgacggcaagtcccgatccacatcaagggaagccatc
ttccacttgtggcacatccacctcagcaggtacaccgtatccatgtcggctgcctc
gcgaagatcaaccggaggcgctctcgacaaggtctgcattctcagctgcggttctct
accggttttggtgcaacagtgaatgtcgcgaaaccgaaaaggccagaccgtggctatt
ttcggctctggagctgttggtcttctgctgctatggaaggtgccaggtgtcctggagcatg
agatcattggtgtgacctgaacctggaatttgaagcaagctagaactttggtcgc

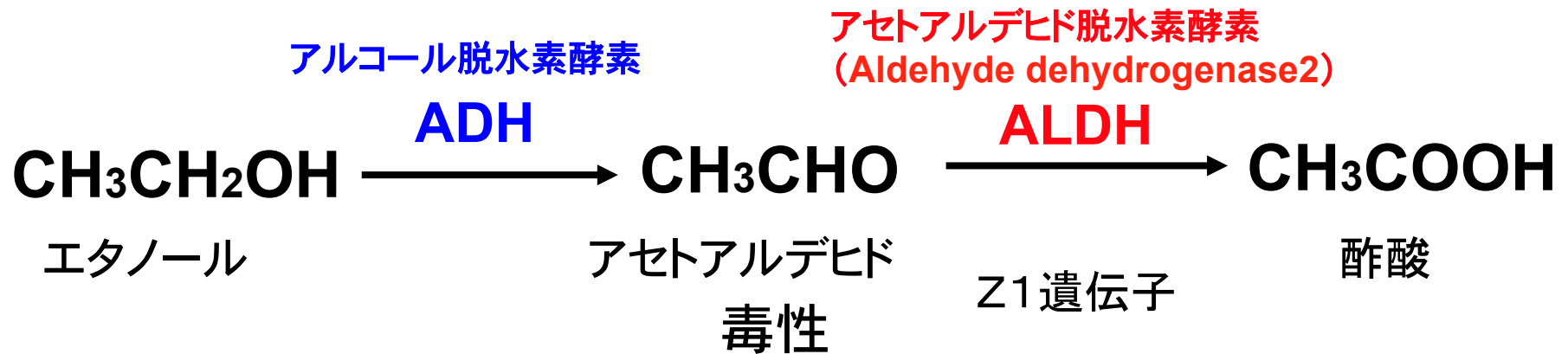
```

予想された遺伝子配列(アミノ酸配列)を  
Blast検索して遺伝子の機能を予測してみよう

# DNAとアミノ酸の アラインメント

DNAの突然変異と形質の違い

## アルコールの代謝



先週の授業でBlast検索した遺伝子

出席と宿題の提出 [提出方法](#)

## リンク集

データベース検索：

1. [PubMed](#): 論文検索
2. [Entrez](#): 総合データベース
3. [Google Scholar](#): 文献データベース
4. [特許データベース](#)

ホモロジー検索：

1. [BLAST](#) [GenomeNET]
2. [FASTA](#) [GenomeNET]

配列解析：

1. [DNA → AA](#) : DNA配列をアミノ酸配列に変換
2. [Spidey](#) : DNAとcDNAのマッチング
3. [BLAST\(NCBI\)](#) : 遺伝子のホモロジー検索
4. [BLAST\(DDBJ\)](#) : 遺伝子のホモロジー検索
5. [Softberry/FGENESH](#) : 遺伝子の予測
6. [GeneMark](#) : 遺伝子の予測
7. [GENESCAN](#) : 遺伝子の予測
8. [clustalW](#) : アライメント
9. [SSPN](#) : スプライシングの予測
10. [ORF Finder](#) : ORFの予測
11. [TFSEARCH](#) : 転写因子結合部位の予測
12. [MOTIF](#) : タンパク質モチーフ検索
13. [pI/Mw](#) : 等電点、分子量の予測

解析用配列

[ある遺伝子AのgenomeDNA配列](#)

Safari ファイル 編集 表示 履歴 ブックマーク ウィンドウ ヘルプ

BLAST: Basic Local Alignment Search Tool

http://blast.ncbi.nlm.nih.gov/Blast.cgi

名古屋大学関係 サイエンス ニュース (321) お役立ち

**BLAST** Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

My NCBI [Sign In] [Regist]

▶ **NCBI/ BLAST Home**

BLAST finds regions of similarity between biological sequences. [more...](#)

**New** Designing or Testing PCR Primers? Try your search in **Primer-BLAST**.

### BLAST Assembled Genomes

Choose a species genome to search, or [list all genomic BLAST databases](#).

- [Human](#)
- [Mouse](#)
- [Rat](#)
- [Arabidopsis thaliana](#)
- [Oryza sativa](#)
- [Bos taurus](#)
- [Danio rerio](#)
- [Drosophila melanogaster](#)
- [Gallus gallus](#)
- [Pan troglodytes](#)
- [Microbes](#)
- [Apis mellifera](#)

### Basic BLAST

Choose a BLAST program to run.

<a href="#">nucleotide blast</a>	Search a <b>nucleotide</b> database using a <b>nucleotide</b> query <i>Algorithms: blastn, megablast, discontinuous megablast</i>
<a href="#">protein blast</a>	Search <b>protein</b> database using a <b>protein</b> query <i>Algorithms: blastp, psi-blast, phi-blast</i>
<a href="#">blastx</a>	Search <b>protein</b> database using a <b>translated nucleotide</b> query
<a href="#">tblastn</a>	Search <b>translated nucleotide</b> database using a <b>protein</b> query
<a href="#">tblastx</a>	Search <b>translated nucleotide</b> database using a <b>translated nucleotide</b> query

### Specialized BLAST

Choose a type of specialized search (or database name in parentheses.)

- Make specific primers with [Primer-BLAST](#)
- Search [trace archives](#)
- Find [conserved domains](#) in your sequence (cds)
- Find sequences with similar [conserved domain architecture](#) (cdart)
- Search sequences that have [gene expression profiles](#) (GEO)
- Search [immunoglobulins](#) (IgBLAST)
- Search for [SNPs](#) (snp)
- Screen sequence for [vector contamination](#) (vecscreen)
- [Align](#) two sequences using BLAST (bl2seq)
- Search [protein](#) or [nucleotide](#) targets in PubChem BioAssay

**News**

[Align Sequences with BLAST](#)

A new BL2seq functionality has been added to the standard BLAST pages that allows you to align a query against a set of subject sequences.  
2008-09-04 12:56:52

[More BLAST news...](#)

**Tip of the Day**

**How to Search Custom Databases in Web-Blast Using Entrez Queries.**

A powerful feature of the BLAST Web interface is the ability to limit BLAST searches to a subset of any database using a standard Entrez query. Skillful use of Entrez queries allows the equivalent of on-the-fly construction of databases of exact composition.

[More tips...](#)

Nucleotide BLAST: Align two or more sequences using BLAST

http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST\_PROGRAMS=megaBlast&PA

名古屋大学関係 サイエンス ニュース (206) 国語辞典 英和辞書 goo辞書 お役立ち

Nucleotide BLAST: Align tw...

**BLAST** Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help My NCBI [Sign In] [Register]

NCBI/BLAST/blastn suite

blastn blastp blastx tblastn tblastx

BLASTN programs search nucleotide subjects using a nucleotide query. more... Reset page Bookmark

Enter Query Sequence

Enter accession number, gi, or FASTA sequence Clear Query subrange

```
GTGATGCAGATCCTGAAGTTCAAGACCATAGAGGAGTTGTTGGGAGAGCCAACAATTCC
ACGTACGGGCTGGCCGACGCTGTCTTCAAAAGGATTTGGACAAGGCCAATTACCTGTCC
CAGGCCCTCCAGGCGGGCACTGTGGGTTCAACTGCTATGATGTGTTTGGAGCCAGTCA
CCCTTTGGTGGCTACAAGATCTCGGGAGTGGCCGGAGTTGGCCGAGTACGGGCTGCAG
GCATACACTGAAGTGAAAAGTCTCACAGTCAAAGTGCCTCAGAAGAAGTCAATAA
```

Z1の予想CDS配列

Or, upload file ファイルを選択 ファイル...いません

Job Title

Enter a descriptive title for your BLAST search

Align two or more sequences

Enter Subject Sequence

Enter accession number, gi, or FASTA sequence Clear Subject subrange

```
ACTGTGTTTGGAGATGTGCAGGATGGCATGACCATCGCCAAGGAGGAGATCTTCGGGCCA
GTGATGCAGATCCTGAAGTTCAAGACCATAGAGGAGTTGTTGGGAGAGCCAACAATTCC
ACGTACGGGCTGGCCGACGCTGTCTTCAAAAGGATTTGGACAAGGCCAATTACCTGTCC
CAGGCCCTCCAGGCGGGCACTGTGGGTTCAACTGCTATGATGTGTTTGGAGCCAGTCA
CCCTTTGGTGGCTACAAGATCTCGGGAGTGGCCGGAGTTGGCCGAGTACGGGCTGCAG
GCATACACTGAAGTGAAAAGTCTCACAGTCAAAGTGCCTCAGAAGAAGTCAATAA
```

Z2の予想CDS配列

Or, upload file ファイルを選択 ファイル...いません

Program Selection

Optimize for

- Highly similar sequences (megablast)
- More dissimilar sequences (discontiguous megablast)
- Somewhat similar sequences (blastn)

Choose a BLAST algorithm

**BLAST** Search nucleotide sequence using Megablast (Optimize for highly similar sequences)

Show results in a new window

Algorithm parameters

Copyright | Disclaimer | Privacy | Accessibility | Contact | Send feedback

NCBI | NLM | NIH | DHHS



NCBI Blast:Nucleotide Sequence (1554 letters)

http://blast.ncbi.nlm.nih.gov/Blast.cgi

名古屋大学関係 サイエンス ニュース (206) 国語辞典 英和辞... - goo辞書 お役立ち

NCBI Blast:Nucleotide Sequ...

**BLAST** Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

My NCBI [Sign In] [Register]

NCBI/BLAST/blastn suite-2sequences/ Formatting Results - H3VR1N2W112

Edit and Resubmit Save Search Strategies Formatting options Download

**Blast 2 sequences**

**Nucleotide Sequence (1554 letters)**

<b>Query ID</b>  cl 17067	<b>Subject ID</b> 17069
<b>Description</b> None	<b>Description</b> None
<b>Molecule type</b> nucleic acid	<b>Molecule type</b> nucleic acid
<b>Query Length</b> 1554	<b>Subject Length</b> 1554
	<b>Program</b> BLASTN 2.2.22+ <a href="#">Citation</a>

Other reports: [Search Summary](#) [Taxonomy reports](#)

**Graphic Summary**

Distribution of 1 Blast Hits on the Query Sequence

Mouse over to see the define, click to show alignments

**Color key for alignment scores**

<40	40-50	50-80	80-200	>=200
-----	-------	-------	--------	-------

Query 0 300 600 900 1200 1500

**Dot Matrix View**

**Descriptions**

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer

Sequences producing significant alignments:  
(Click headers to sort columns)

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
17069		2865	2865	100%	0.0	99%	

**Alignments**  Select All [Get selected sequences](#) **NEW**

```
>|cl|17069
Length=1554
```

NCBI Blast:Nucleotide Sequence (1554 letters)

http://blast.ncbi.nlm.nih.gov/Blast.cgi

名古屋大学関係 サイエンス ニュース (206) 国語辞典 英和辞... - goo辞書 お役立ち

NCBI Blast:Nucleotide Sequ...

Sequences producing significant alignments:  
(Click headers to sort columns)

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
17069		2865	2865	100%	0.0	99%	

**Alignments**  Select All [Get selected sequences](#) NEW

```

>lcl|17069
Length=1554

Score = 2865 bits (1551), Expect = 0.0
Identities = 1553/1554 (99%), Gaps = 0/1554 (0%)
Strand=Plus/Plus

Query 1   ATGTTGcgcgctgccgcccgttcggggcccccgcctgggcccgcctcttgtcagccgcc 60
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
Sbjct 1   ATGTTGCGCGCTGCCGCCGCTTCGGGCCCGCCTGGGCCGCCGCTCTTGTGACGCCGCC 60

Query 61   gccacccagggcgtgcctgcccccaaccagcagcccgaggtcttctgcaaccagattttc 120
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
Sbjct 61   GCCACCCAGGCCGTGCCTGCCCAACCAGCAGCCCGAGGTCTTCTGCAACCAGATTFTTC 120

Query 121  ATAAACAATGAATGGCACGATGCCGTGAGCAGGAAACATCCCACCCTCAATCCGTCC 180
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
Sbjct 121  ATAAACAATGAATGGCACGATGCCGTGAGCAGGAAACATCCCACCCTCAATCCGTCC 180

Query 181  ACTGGAGAGTCACTGTCAGGTAGCTGAAGGGACAAGGAAGATGTGGACAAGGCAGTG 240
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
Sbjct 181  ACTGGAGAGTCACTGTCAGGTAGCTGAAGGGACAAGGAAGATGTGGACAAGGCAGTG 240

Query 241  AAGGCCGCCGGGGCCGCTTCCAGCTGGGCTCACCCTGGGCCGCATGGACGCATCACAC 300
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
Sbjct 241  AAGGCCGCCGGGGCCGCTTCCAGCTGGGCTCACCCTGGGCCGCATGGACGCATCACAC 300

Query 301  AGGGCCGGCTGCTGAACCGCCTGGCCGATCTGATCGAGCGGACCGGACCTACCTGGCG 360
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
Sbjct 301  AGGGCCGGCTGCTGAACCGCCTGGCCGATCTGATCGAGCGGACCGGACCTACCTGGCG 360

Query 361  GCCTTGGAGACCCTGGACAATGGCAAGCCCTATGTCATCTCTACCTGGTGGATTGGAC 420
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
Sbjct 361  GCCTTGGAGACCCTGGACAATGGCAAGCCCTATGTCATCTCTACCTGGTGGATTGGAC 420

Query 421  ATGGTCCCAAATGCTCCGGTATTATGCCGGCTGGGCTGATAAGTACCACGGGAAAC 480
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
Sbjct 421  ATGGTCCCAAATGCTCCGGTATTATGCCGGCTGGGCTGATAAGTACCACGGGAAAC 480

Query 481  ATCCCCATTGACGGAGACTTCTTCAGCTACACAGCCATGAACCTGTGGGGGTGTGGGG 540
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
Sbjct 481  ATCCCCATTGACGGAGACTTCTTCAGCTACACAGCCATGAACCTGTGGGGGTGTGGGG 540

Query 541  CAGATCATTCCGTGGAATTTCCCGCTCCTGATGCAAGCATGGAAGCTGGGCCACGCCCTG 600
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
Sbjct 541  CAGATCATTCCGTGGAATTTCCCGCTCCTGATGCAAGCATGGAAGCTGGGCCACGCCCTG 600

Query 601  GCAACTGGAAACGTGGTTGTGATGAAGGTAGCTGAGCAGACACCCCTCACCGCCCTCTAT 660
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
Sbjct 601  GCAACTGGAAACGTGGTTGTGATGAAGGTAGCTGAGCAGACACCCCTCACCGCCCTCTAT 660

Query 661  GTGGCCAACCTGATCAAGGAGGCTGGCTTTCCCCCTGGTGTGGTCAACATTGTGCCTGGA 720
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
Sbjct 661  GTGGCCAACCTGATCAAGGAGGCTGGCTTTCCCCCTGGTGTGGTCAACATTGTGCCTGGA 720

Query 721  TTTGGCCCACGGCTGGGGCCGCATTCCTCCCATGAGGATGTGGACAAAGTGGCATTTC 780
          |||..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
  
```

NCBI Blast:Nucleotide Sequence (1554 letters)

http://blast.ncbi.nlm.nih.gov/Blast.cgi

名古屋大学関係 サイエンス ニュース (206) 国語辞典 英和辞書 - goo辞書 お役立ち

NCBI Blast:Nucleotide Sequ...

Sbjct	541	CAGATCATTCCGCTGGAATTTCCCGCTCTGATGCAAGCATGGAAGCTGGGCCAGCCTTG	600
Query	601	GCAACTGGAAACGTGGTTGTGATGAAGTAGCTGAGCAGACACCCCTCACGCCCTCTAT	660
Sbjct	601	GCAACTGGAAACGTGGTTGTGATGAAGTAGCTGAGCAGACACCCCTCACGCCCTCTAT	660
Query	661	GTGGCCAACCTGATCAAGGAGGCTGGCTTTCCCCCTGGTGTGGTCAACATTTGTCCTGGA	720
Sbjct	661	GTGGCCAACCTGATCAAGGAGGCTGGCTTTCCCCCTGGTGTGGTCAACATTTGTCCTGGA	720
Query	721	TTTGGCCCCACGGCTGGGGCCGCAATTCCTCCCATGAGGATGTGGACAAAGTGGCATTC	780
Sbjct	721	TTTGGCCCCACGGCTGGGGCCGCAATTCCTCCCATGAGGATGTGGACAAAGTGGCATTC	780
Query	781	ACAGGCTCCACTGAGATTGGCCGCGTAATCCAGGTTGCTGCTGGGAGCAGCAACCTCAAG	840
Sbjct	781	ACAGGCTCCACTGAGATTGGCCGCGTAATCCAGGTTGCTGCTGGGAGCAGCAACCTCAAG	840
Query	841	AGAGTGACCTTGGAGTgggggggAAGAGCCCCAACATCATATGTCAGATGCCGATATG	900
Sbjct	841	AGAGTGACCTTGGAGTGGGGGGGAAGAGCCCCAACATCATATGTCAGATGCCGATATG	900
Query	901	GATTGGCCCGTGAACAGGCCACTTCGCCCTGTTCTTCAACCAGGGCCAGTGTCTGT	960
Sbjct	901	GATTGGCCCGTGAACAGGCCACTTCGCCCTGTTCTTCAACCAGGGCCAGTGTCTGT	960
Query	961	GCCGGCTCCCGACCTTCGTGCAGGAGGACATCTATGATGAGTTTGTGGAGCGGAGCGTT	1020
Sbjct	961	GCCGGCTCCCGACCTTCGTGCAGGAGGACATCTATGATGAGTTTGTGGAGCGGAGCGTT	1020
Query	1021	GCCCGGGCCAAGTCTCGGGTGGTCCGGAAACCCCTTTGATAGCAAGACCGAGCAGGGCCG	1080
Sbjct	1021	GCCCGGGCCAAGTCTCGGGTGGTCCGGAAACCCCTTTGATAGCAAGACCGAGCAGGGCCG	1080
Query	1081	CAGGTGGATGAAACTCAGTTAAGAAGATCCTCGGCTACATCAACACGGGGAAGCAAGAG	1140
Sbjct	1081	CAGGTGGATGAAACTCAGTTAAGAAGATCCTCGGCTACATCAACACGGGGAAGCAAGAG	1140
Query	1141	GGGGCGAAGCTGCTGTGTGGTGGGGGCAATGCTGTGACCGTGGTTACTTCATCCAGCCC	1200
Sbjct	1141	GGGGCGAAGCTGCTGTGTGGTGGGGGCAATGCTGTGACCGTGGTTACTTCATCCAGCCC	1200
Query	1201	ACTGTGTTTGGAGATGTGCAGGATGGCATGACCATCGCCAAGGAGGAGATCTTCGGGCCA	1260
Sbjct	1201	ACTGTGTTTGGAGATGTGCAGGATGGCATGACCATCGCCAAGGAGGAGATCTTCGGGCCA	1260
Query	1261	GTGATGCAGATCCTGAAGTTCAAGACCATAGAGGAGGTTGTTGGGAGAGCCAACAATTC	1320
Sbjct	1261	GTGATGCAGATCCTGAAGTTCAAGACCATAGAGGAGGTTGTTGGGAGAGCCAACAATTC	1320
Query	1321	ACGTACGGGCTGGCCGAGCTGTCTTCAAAAAGGATTTGGACAAGGCCAATTACCTGTCC	1380
Sbjct	1321	ACGTACGGGCTGGCCGAGCTGTCTTCAAAAAGGATTTGGACAAGGCCAATTACCTGTCC	1380
Query	1381	CAGGCCCTCCAGGCGGGCACTGTGGGTCAACTGCTATGATGTTTGGAGCCAGTCA	1440
Sbjct	1381	CAGGCCCTCCAGGCGGGCACTGTGGGTCAACTGCTATGATGTTTGGAGCCAGTCA	1440
Query	1441	CCCTTTGGTGGCTACAAGATGTCGGGGAGTGGCCGGGAGTTGGGCGAGTACGGGCTGCAG	1500
Sbjct	1441	CCCTTTGGTGGCTACAAGATGTCGGGGAGTGGCCGGGAGTTGGGCGAGTACGGGCTGCAG	1500
Query	1501	GCATACACTGAGTGAAGTGTACAGTCAAAGTGCCTCAGAAGAACTATAA 1554	
Sbjct	1501	GCATACACTGAGTGAAGTGTACAGTCAAAGTGCCTCAGAAGAACTATAA 1554	

Select All [Get selected sequences](#) NEW

**Z1CDS**

**GCATACACTGAAGTGAAAACGTGTCACAGTCAAAGTGCCTCAGAAGAACTCATAA**

**Z2CDS**

**GCATACACTAAAGTGAAAACGTGTCACAGTCAAAGTGCCTCAGAAGAACTCATAA**

**\*\*\*\*\* \*\*\*\*\***

Protein BLAST: Align two or more sequences using BLAST

http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&BLAST\_PROGRAMS=blastp&PAGE\_ Google

名古屋大学関係 サイエンス ニュース (206) 国語辞典 英和辞... - goo辞書 お役立ち

Protein BLAST: Align two o...

**BLAST** Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

My NCBI [Sign In] [Register]

NCBI/BLAST/blastp suite

blastp blastx tblastn tblastx

BLASTP programs search protein subjects using a protein query. more... Reset page Bookmark

Enter Query Sequence

Enter accession number, gi, or FASTA sequence Clear Query subrange

MLRAAARFGPRLGRRLLSAAATQAVPAPNQPEVFCNQIFINNEWHDAVSRKTFPTVNPS  
TGEVICQVAEGDKEDVDKAVKAARAFQLGSPWRRMDASHRGRLNRLADLIERDRTYLA  
ALETLDNGKPYVISYLVLDMLKCLRYAGWADKYHGKTIPIIDGDFFSYTRHEPVGCG  
QIIPWNFLLMQAWKLGALATGNVVMKVAEQPLTALYVANLIKEAGFPPGVVNIIVPG  
FGPTAGAAIASHEDVDKVAFTGSTEIGRVIQVAAGSSNLKRVTLLEGGKSPNIIMSDADM

From To Z1のアミノ酸配列

Or, upload file ファイルを選択 ファイル...いません

Job Title

Enter a descriptive title for your BLAST search

Align two or more sequences

Enter Subject Sequence

Enter accession number, gi, or FASTA sequence Clear Subject subrange

MLRAAARFGPRLGRRLLSAAATQAVPAPNQPEVFCNQIFINNEWHDAVSRKTFPTVNPS  
TGEVICQVAEGDKEDVDKAVKAARAFQLGSPWRRMDASHRGRLNRLADLIERDRTYLA  
ALETLDNGKPYVISYLVLDMLKCLRYAGWADKYHGKTIPIIDGDFFSYTRHEPVGCG  
QIIPWNFLLMQAWKLGALATGNVVMKVAEQPLTALYVANLIKEAGFPPGVVNIIVPG  
FGPTAGAAIASHEDVDKVAFTGSTEIGRVIQVAAGSSNLKRVTLLEGGKSPNIIMSDADM

From To Z2のアミノ酸配列

Or, upload file ファイルを選択 ファイル...いません

Program Selection

Algorithm  blastp (protein-protein BLAST)  
Choose a BLAST algorithm

**BLAST** Search protein sequence using Blastp (protein-protein BLAST)  
 Show results in a new window

Algorithm parameters

Copyright | Disclaimer | Privacy | Accessibility | Contact | Send feedback

NCBI | NLM | NIH | DHHS

http://www.ncbi.nlm.nih.gov/ を新規タブで開く

NCBI Blast:Protein Sequence (517 letters)

http://blast.ncbi.nlm.nih.gov/Blast.cgi

名古屋大学関係 サイエンス ニュース (206) 国語辞典 英和辞... - goo辞書 お役立ち

NCBI Blast:Protein Sequenc...

**BLAST** Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

My NCBI [Sign In] [Register]

NCBI/BLAST/blastp suite-2sequences/Formatting Results - H3V346A3112

[Edit and Resubmit](#) [Save Search Strategies](#) [Formatting options](#) [Download](#)

**Blast 2 sequences**

**Protein Sequence (517 letters)**

<b>Query ID</b>	lcl 9111	<b>Subject ID</b>	9113
<b>Description</b>	None	<b>Description</b>	None
<b>Molecule type</b>	amino acid	<b>Molecule type</b>	amino acid
<b>Query Length</b>	517	<b>Subject Length</b>	517
		<b>Program</b>	BLASTP 2.2.22+ <a href="#">Citation</a>

Other reports: [Search Summary](#) [Taxonomy reports](#)

**Graphic Summary**

Distribution of 1 Blast Hits on the Query Sequence

Mouse over to see the define, click to show alignments

**Color key for alignment scores**

<40	40-50	50-80	80-200	>=200
-----	-------	-------	--------	-------

Query 0 100 200 300 400 500

**Dot Matrix View**

**Descriptions**

Sequences producing significant alignments:	Score (Bits)	E Value
lcl 9113 unnamed protein product	<a href="#">1070</a>	0.0

**Alignments**  Select All [Get selected sequences](#) NEW

```

>lcl|9113 unnamed protein product
Length=517

Score = 1070 bits (2767), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 516/517 (99%), Positives = 517/517 (100%), Gaps = 0/517 (0%)

Query 1      MLRAAAREGPRIGRRLLSAAATOAVPAPNOOPEVEFCNOIFINNEWHDAVSRKTEPTVNS 60
  
```

NCBI Blast:Protein Sequence (517 letters)

http://blast.ncbi.nlm.nih.gov/Blast.cgi

名古屋大学関係 サイエンス ニュース (206) 国語辞典 英和辞... - goo辞書 お役立ち

NCBI Blast:Protein Sequenc...

Program BLASTP 2.2.22+ [Citation](#)

Other reports: [Search Summary](#) [Taxonomy reports](#)

### Graphic Summary

Distribution of 1 Blast Hits on the Query Sequence

Mouse over to see the define, click to show alignments

**Color key for alignment scores**

Score Range	Color
<40	Black
40-50	Blue
50-80	Green
80-200	Pink
>=200	Red

Query 0 100 200 300 400 500

### Dot Matrix View

Plot of lcl|9111 vs 9113

### Descriptions

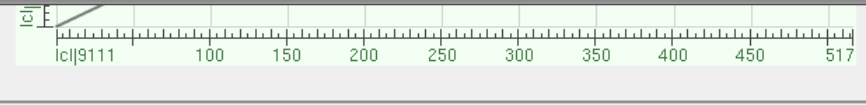
Sequences producing significant alignments:	Score (Bits)	E Value
lcl 9113 unnamed protein product	<a href="#">1070</a>	0.0

NCBI Blast:Protein Sequence (517 letters)

http://blast.ncbi.nlm.nih.gov/Blast.cgi

名古屋大学関係 サイエンス ニュース (206) 国語辞典 英和辞... - goo辞書 お役立ち

NCBI Blast:Protein Sequenc...



**▼ Descriptions**

Sequences producing significant alignments:	Score (Bits)	E Value
lcl 9113 unnamed protein product	<a href="#">1070</a>	0.0

**▼ Alignments**  Select All [Get selected sequences](#) NEW

```

>lcl|9113 unnamed protein product
Length=517

Score = 1070 bits (2767), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 516/517 (99%), Positives = 517/517 (100%), Gaps = 0/517 (0%)

Query 1  MLRAAARFGPRLGRRLLSAAATQAVPAPNQPEVFCNQIFINNEWHDAVSRKTFPTVNPS 60
Sbjct 1  MLRAAARFGPRLGRRLLSAAATQAVPAPNQPEVFCNQIFINNEWHDAVSRKTFPTVNPS 60

Query 61  TGEVICQVAEGDKEDVDKAVKAARAAFQLGSPWRRMDASHRGRLNRLADLIERDRTYLA 120
Sbjct 61  TGEVICQVAEGDKEDVDKAVKAARAAFQLGSPWRRMDASHRGRLNRLADLIERDRTYLA 120

Query 121  ALETLDNGKPYVISYLVLDLDMVLKCLRYAGWADKYHGKTIPIDGDFFSYTRHEPVGVC 180
Sbjct 121  ALETLDNGKPYVISYLVLDLDMVLKCLRYAGWADKYHGKTIPIDGDFFSYTRHEPVGVC 180

Query 181  QIIPWNFPLLMQAWKLGPALATGNVVMKVAEQTPLTALYVANLIKEAGFPVGVNIVPG 240
Sbjct 181  QIIPWNFPLLMQAWKLGPALATGNVVMKVAEQTPLTALYVANLIKEAGFPVGVNIVPG 240

Query 241  FGPTAGAAIASHEDVDKVAFTGSTEIGRVIQVAAGSSNLKRVTLGKSPNIIMSDADM 300
Sbjct 241  FGPTAGAAIASHEDVDKVAFTGSTEIGRVIQVAAGSSNLKRVTLGKSPNIIMSDADM 300

Query 301  DWAVEQAHFALFFNQCCAGSRTFVQEDIYDEFVRSVARAKSRVVGPNPFSKTEQGP 360
Sbjct 301  DWAVEQAHFALFFNQCCAGSRTFVQEDIYDEFVRSVARAKSRVVGPNPFSKTEQGP 360

Query 361  QVDETQFKKILGYINTGKQEGAKLLCGGGIAADRGYFIQPTVFGDVQDGMTIAKEEIFGP 420
Sbjct 361  QVDETQFKKILGYINTGKQEGAKLLCGGGIAADRGYFIQPTVFGDVQDGMTIAKEEIFGP 420

Query 421  VMQILKFKTIEEVVGRANNSTYGLAAAVFTKDLKANYLSQALQAGTVVWNCYDVFGAQS 480
Sbjct 421  VMQILKFKTIEEVVGRANNSTYGLAAAVFTKDLKANYLSQALQAGTVVWNCYDVFGAQS 480

Query 481  PFGGYKMSGSGRELGEYGLQAYTETKTVTVKVPQKNS 517
Sbjct 481  PFGGYKMSGSGRELGEYGLQAYTETKTVTVKVPQKNS 517

```



Z1CDS GCATACACTGAAGTGAAAACGTGCACAGTCAAAGTGCCTCAGAAGAACTCATAA  
Z2CDS GCATACACTAAAGTGAAAACGTGCACAGTCAAAGTGCCTCAGAAGAACTCATAA  
\*\*\*\*\*

グルタミン酸

Z1Amionoacid PFGGYKMSGSGRELGEYGLQAYTEVKTVTVKVPQKNS  
Z2Amionoacid PFGGYKMSGSGRELGEYGLQAYTKVKTVTVKVPQKNS  
\*\*\*\*\*

リジン

Spidey -- mRNA to genomic alignment

http://www.ncbi.nlm.nih.gov/IEB/Research/Ostell/Spidey/

名古屋大学関係 サイエンス ニュース (206) 国語辞典 英和辞書 goo辞書 お役立ち

Spidey -- mRNA to genom...

NCBI Spidey

PubMed Entrez BLAST OMIM Taxonomy Structure

Spidey FAQ  
Spidey documentation  
Spidey executables  
Help/Contact

**Spidey is an mRNA-to-genomic alignment program. For a complete description of how Spidey works, click [here](#). For an example, click [here](#).**

**Genomic sequence (FASTA or GI/Accession):**

Upload file: (ファイルを選択) ファイルが選...ていません

```
CTGAGGCAGGAAAATTGGTTGAACCTGGGAGGTGGAGTTG
CAGTGAGTGGAGATCATGCCACTGCACTCCAGCCTGGGAGT
CAGAGCCGAGACTCCAATCTCAAAAAAAGTAAATAAATGAAAT
AGCCTAAGGAAATGCTGCCAGGCAGTAATGATATGCCACAC
TGGATATGATTTCTGCCCTCCTCTGCTGTGGGTAACAGCT
TCTGTTTCATGCA
```

From: 0 To: 0

**mRNA sequence(s) (One or more FASTA or GI/Accession)**

Upload file: (ファイルを選択) ファイルが選...ていません

```
CAGGCCCTCCAGCGGGCACTGTGTGGTCAACTGCTATGA
TGTGTTTGGAGCCAGTCA
CCCTTTGGTGGCTACAAGATGTCGGGGAGTGGCCGGGAGTT
GGCGGAGTACGGGCTGCAG
GCATACACTAAAGTAAAAGTGCACAGTCAAAGTGCCTCA
GAAGAACTATAA
```

divergent sequences ?  
 Use large intron sizes ?

Align  
Clear

Minimum mRNA-genomic identity ? 0 %  
Minimum length of mRNA covered ? 0 %

**Genomic sequence is: ?**

Vertebrate  
 Drosophila  
 C. elegans  
 Plant

**Output options:**

Text/summary  
 Summary only  
 ASN.1  
 Print multiple alignment

Z1のゲノミックDNA配列

Z2の予想CDS酸配列

Spidey -- mRNA to genomic alignment (808905382)

http://www.ncbi.nlm.nih.gov/IEB/Research/Ostell/Spidey/spideyweb.cgi

名古屋大学関係 ▾ サイエンス ▾ ニュース (206) ▾ 国語辞典 英和辞... - goo辞書 お役立ち ▾

Spidey -- mRNA to genom...

Alignment is on plus strand of genomic sequence and on plus strand of mRNA sequence  
 mRNA coverage: 100%  
 Overall percent identity: 99.9%

	Genomic coordinates	mRNA coordinates	length	identity	mismatches	gaps	Donor site	Acc. site
<a href="#">Exon 1</a>	562-675	1-114	114	100.0%	0	0	d	
<a href="#">Exon 2</a>	15497-15601	115-219	105	100.0%	0	0	d	a
<a href="#">Exon 3</a>	16737-16877	220-360	141	100.0%	0	0	d	a
<a href="#">Exon 4</a>	18846-18925	361-440	80	100.0%	0	0	d	a
<a href="#">Exon 5</a>	23402-23513	441-552	112	100.0%	0	0	d	a
<a href="#">Exon 6</a>	24013-24141	553-681	129	100.0%	0	0	d	a
<a href="#">Exon 7</a>	24885-24998	682-795	114	100.0%	0	0	d	a
<a href="#">Exon 8</a>	25640-25742	796-898	103	100.0%	0	0	d	a
<a href="#">Exon 9</a>	26177-26361	899-1083	185	100.0%	0	0	d	a
<a href="#">Exon 10</a>	31657-31821	1084-1248	165	100.0%	0	0	d	a
<a href="#">Exon 11</a>	33485-33642	1249-1406	158	100.0%	0	0	d	a
<a href="#">Exon 12</a>	37438-37552	1407-1521	115	99.1%	1	0	d	a
<a href="#">Exon 13</a>	43122-43154	1522-1554	33	100.0%	0	0		a

Spidey -- mRNA to genomic alignment (808905382)

http://www.ncbi.nlm.nih.gov/IEB/Research/Ostell/Spidey/spideyweb.cgi

名古屋大学関係 サイエンス ニュース (206) 国語辞典 英和辞... - goo辞書 お役立ち

Spidey -- mRNA to genom...

```

33525 TAGAGGAGGTTGTTGGGAGAGCCAACAATTCACGTTACGGGCTGGCCGCA
      |||
1289 TAGAGGAGGTTGTTGGGAGAGCCAACAATTCACGTTACGGGCTGGCCGCA
      |||
      I E E V V G R A N N S T Y G L A A

33575 GCTGTCTTCACAAAGGATTTGGACAAGGCCAATTACCTGTCCCAGGCCCT
      |||
1339 GCTGTCTTCACAAAGGATTTGGACAAGGCCAATTACCTGTCCCAGGCCCT
      |||
      A V F T K D L D K A N Y L S Q A L

33625 CCAGGCGGGCACTGTGTGGTAAGAGCCT
      |||
1389 CCAGGCGGGCACTGTGTG
      |||
      Q A G T V W

```

[Top](#)

**Exon 12: 37438-37552 (genomic); 1407-1521 (mRNA)**

Protein BLAST: Align two or more sequences using BLAST

```

37438 CAAATTACAGGGTCAACTGCTATGATGTGTTTGGAGCCAGTCACCCCTT
      |||
1407 GGTCAACTGCTATGATGTGTTTGGAGCCAGTCACCCCTT
      |||
      V N C Y D V F G A Q S P F

37478 GGTGGCTACAAGATGTCGGGGAGTGGCCGGGAGTGGGCGAGTACGGGCT
      |||
1447 GGTGGCTACAAGATGTCGGGGAGTGGCCGGGAGTGGGCGAGTACGGGCT
      |||
      G G Y K M S G S C R E L G E Y G L

37528 GCAGGCATACACTCAAGTGAAGTGTGAGTGTGG
      |||
1497 GCAGGCATACACTAAGTGAAGTGTGAGTGTGG
      |||
      Q A Y T K V K T

```

[Top](#)

**Exon 13: 43122-43154 (genomic); 1522-1554 (mRNA)**

```

43122 CCCCTTACAGGTCACAGTCAAAGTGCCTCAGAAGAACTCATAAGAATCATGCA
      |||
1522 GTCACAGTCAAAGTGCCTCAGAAGAACTCATA
      |||
      V T V K V P Q K N S *

```

[Top](#)

Spidey executable    FAQ    Download Source    Privacy statement    Disclaimer

第12exon

第13exon

Z1CDS GCATACACT**GA**AGTGAAAAC**T**GTCACAGTCAAAGTGCCTCAGAAGAACTCATAA  
 Z2CDS GCATACACT**AA**AGTGAAAAC**T**GTCACAGTCAAAGTGCCTCAGAAGAACTCATAA  
 \*\*\*\*\*

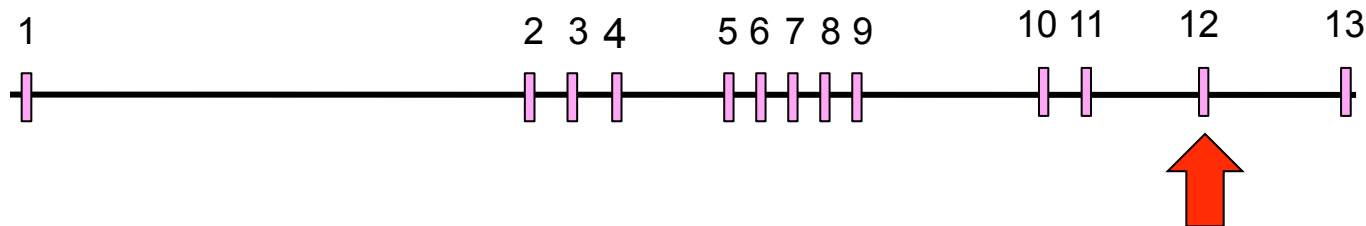
第12exon

グルタミン酸

第13exon

Z1Amionoacid PFGGYKMSGSGRELGEYGLQAYT**E**VKTVTVKVPQKNS  
 Z2Amionoacid PFGGYKMSGSGRELGEYGLQAYT**K**VKTVTVKVPQKNS  
 \*\*\*\*\*

リジン





## アルコールの代謝



### G型(1型)

CAG GCA TAC ACT **GAA** GAG AAA  
--- G A Y T **E** V K --

### A型(2型)

CAG GCA TAC ACT **AAA** GAG AAA  
--- G A Y T **K** V K --

酵素活性がない

## アルコールの代謝



### G型(1型)

CAG GCA TAC ACT GAA GAG AAA  
 --- G A Y T E V K ---  
 酸性アミノ酸

GGtype: G型(1型)をホモにもつお酒に強い

### A型(2型)

CAG GCA TAC ACT AAA GAG AAA  
 --- G A Y T K V K ---  
 塩基性アミノ酸

AGtype: G型、A型をヘテロにもつ  
 お酒にほんの少し強い人(強い人の1/16)  
 (4量体を形成しAタイプが優性に活性を無力化するため)

酵素活性がない

AAtype: A型(2型)をホモにもつお酒に弱い