

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

遺伝子予想ソフト

MEGANTE : <https://megante.dna.affrc.go.jp/>

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

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](#): 何でも検索
4. [Google Scholar](#): 科学に関する検索
5. [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. [Splign](#) : DNAとcDNAのマッピング
2. [Blast align](#) : DNAとcDNAのマッピング
3. [EMBOSS program](#) : DNAとcDNAのマッピング
4. [EMBOSS program](#) : DNA配列をアミノ酸配列に変換
5. [BLAST\(NCBI\)](#) : 遺伝子のホモロジー検索
6. [BLAST\(DDBJ\)](#) : 遺伝子のホモロジー検索
7. [MEGANTE](#) : 遺伝子の予測 
8. [Softberry/FGENESH](#) : 遺伝子の予測
9. [GeneMark](#) : 遺伝子の予測
10. [GENESCAN](#) : 遺伝子の予測
11. [clustalW](#) : アライメント
12. [InterProScan](#) : ドメイン解析
13. [PSORT](#) : 細胞内局在性予測
14. [GEO](#) : 遺伝子発現解析
15. [Gene Ontology](#) : 用語を用いた検索
16. [SSPN](#) : スプライシングの予測
17. [ORF Finder](#) : ORFの予測
18. [TFSEARCH](#) : 転写因子結合部位の予測
19. [MOTIF](#) : タンパク質モチーフ検索
20. [pI/Mw](#) : 等電点、分子量の予測
21. [RAP-DB](#) : イネのゲノムデータベース

INTEGRATED PLANT GENOME ANNOTATION

[What is MEGANTE?](#)[Examples](#)[Publications](#)[Release history](#)[Create an account →](#)

パスワード: 大文字、小文字、数字が必要
Sign in

[Forgot password?](#)

これまでに、沢山の研究であきらかになった情報が集積されていて、
正確な遺伝子の存在場所や機能が簡単に推測できるようになっている。

INTEGRATED PLANT GENOME ANNOTATION



What is MEGANTE?



Examples



Publications



Release history

Create an account →

Sign in

Sign in

[Forgot password?](#)

List of sequences

[Upload sequence](#)

Show 10 entries



<input type="checkbox"/>	Sequence ID	Query name	Uploaded	State	Annotation	Delete
--------------------------	-------------	------------	----------	-------	------------	--------

No uploaded sequences

Showing 0 to 0 of 0 entries



Upload sequences

Number of sequences that can be uploaded: **100**

Query sequence

Enter sequence in FASTA format, or upload FASTA file (maximum length is 10 Mb)

Species

Email notification

 Send email when annotation is finished

Query name (optional)

[← Back to list of sequences page](#)

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

配列解析：

1. [Splign](#) : DNAとcDNAのマッチング
2. [Blast align](#) : DNAとcDNAのマッチング
3. [EMBOSS program"](#) : DNAとcDNAのマッチング
4. [EMBOSS program"](#) : DNA配列をアミノ酸配列に変換
5. [BLAST\(NCBI\)](#) : 遺伝子のホモロジー検索
6. [BLAST\(DDBJ\)](#) : 遺伝子のホモロジー検索
7. [Softberry/FGENESH](#) : 遺伝子の予測
8. [GeneMark](#) : 遺伝子の予測
9. [GENESCAN](#) : 遺伝子の予測
10. [clustalW](#) : アライメント
11. [InterProScan](#) : ドメイン解析
12. [PSORT](#) : 細胞内局在性予測
13. [GEO](#) : 遺伝子発現解析
14. [Gene Ontology](#) : 用語を用いた検索
15. [SSPN](#) : スプライシングの予測
16. [ORF Finder](#) : ORFの予測
17. [TFSEARCH](#) : 転写因子結合部位の予測
18. [MOTIF](#) : タンパク質モチーフ検索
19. [pI/Mw](#) : 等電点、分子量の予測
20. [RAP-DB](#) : イネのゲノムデータベース

解析用配列

[1. ある遺伝子AのgenomeDNA配列](#)



[2. ある遺伝子AのcDNA配列](#)

[3. ある遺伝子XのgenomeDNA配列](#)

[4. ある遺伝子XのcDNA配列](#)

[5. ある遺伝子YのgenomicDNA](#)

[6. ある遺伝子YのcDNA](#)

[7. Z1遺伝子のgenomicDNA](#)

[8. Z1遺伝子のCDS](#)

[9. Z1遺伝子の予想アミノ酸配列](#)

[10. Z2遺伝子のgenomicDNA](#)

[11. Z2遺伝子のCDS](#)

[12. Z2遺伝子のアミノ酸配列](#)

[13. 野生型SD1-genomicDNA配列](#)

[14. 野生型SD1-cDNA配列](#)

[15. カルロス-cDNA](#)

New query was successfully uploaded.

Upload sequence

List of sequences

Show 10 entries



<input type="checkbox"/>	Sequence ID ↑↓	Query name ↑↓	Uploaded ↓↑	State ↑↓	Annotation	Delete
<input type="checkbox"/>	▼ A	Unspecified	2017-10-17 16:01	Waiting	Unavailable	
<input type="checkbox"/>	▼ Y	Unspecified	2017-10-17 15:55	Running	Unavailable	

Showing 1 to 2 of 2 entries

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

配列解析：

1. [Splign](#) : DNAとcDNAのマッチング
2. [Blast align](#) : DNAとcDNAのマッチング
3. [EMBOSS program"](#) : DNAとcDNAのマッチング
4. [EMBOSS program"](#) : DNA配列をアミノ酸配列に変換
5. [BLAST\(NCBI\)](#) : 遺伝子のホモロジー検索
6. [BLAST\(DDBJ\)](#) : 遺伝子のホモロジー検索
7. [Softberry/FGENESH](#) : 遺伝子の予測
8. [GeneMark](#) : 遺伝子の予測
9. [GENESCAN](#) : 遺伝子の予測
10. [clustalW](#) : アライメント
11. [InterProScan](#) : ドメイン解析
12. [PSORT](#) : 細胞内局在性予測
13. [GEO](#) : 遺伝子発現解析
14. [Gene Ontology](#) : 用語を用いた検索
15. [SSPN](#) : スプライシングの予測
16. [ORF Finder](#) : ORFの予測
17. [TFSEARCH](#) : 転写因子結合部位の予測
18. [MOTIF](#) : タンパク質モチーフ検索
19. [pI/Mw](#) : 等電点、分子量の予測
20. [RAP-DB](#) : イネのゲノムデータベース

解析用配列

1. [ある遺伝子AのgenomeDNA配列](#)
2. [ある遺伝子AのcDNA配列](#)
3. [ある遺伝子XのgenomeDNA配列](#)
4. [ある遺伝子XのcDNA配列](#)
5. [ある遺伝子YのgenomicDNA](#) 
6. [ある遺伝子YのcDNA](#)
7. [Z1遺伝子のgenomicDNA](#)
8. [Z1遺伝子のCDS](#)
9. [Z1遺伝子の予想アミノ酸配列](#)
10. [Z2遺伝子のgenomicDNA](#)
11. [Z2遺伝子のCDS](#)
12. [Z2遺伝子のアミノ酸配列](#)
13. [野生型SD1-genomicDNA配列](#)
14. [野生型SD1-cDNA配列](#)
15. [カルロス-cDNA](#)

Upload sequences

Number of sequences that can be uploaded: 100

Query sequence

Enter sequence in FASTA format, or upload FASTA file (maximum length is 10 Mb)

```
>Y
atccgccgaaaccgaaaccaccgaagctcgaaccggcgaggcaagcagcagcagcatgcagcagacgcagagccttctcttcc
aagcaaaaaatattcagatgacgcctcctcctcttccccctcctccacttgcttgccttcaggcgtacttgcccgagactcctaccact
ctctgatccctatatatacaccctcctccctcctcgtctcctcaccagaattccaagcaacgaactgcgagtgattcaagaaaaagaaa
cctgagctttcgatctcttcggagtggttcttgttcttgaaaaagaggggattaatggcgacagccgggaagtgatcaagtgcaaaggt
cogtcttgctgttcttgagcgggttcgtgagaaaattagtttgatgtttgtggggaggagaagctgattagatcgggttttggggttttg
cagcggcggtggcatgggagcgggaagcogctgtcgcagaggtggaggttgcgcgcgcagggcatggaggtccgcgtcaagatcc
tctacaccgccctctgccacaccgacgtctacttctgggagggcaagtaattaatcagcatctgccatggctgagctttctcaccttctcaa
gcttgcaagaatgatcagtgactaactaacttgggtgtgtgtgtgttctgcttggatcttcgcaggggcaaacacctgtttccctaggatc
ttgggcatgaagctggagggatgtgtaactgaaccaacctgtccaattcttcttctactacactgtttagtggttcagctctattct
```


Species

Oryza sativa

Email notification

 Send email when annotation is finished

Query name (optional)

Short description to identify the query

[← Back to list of sequences page](#)

New query was successfully uploaded.

List of sequences

Upload sequence

Show 10 entries

Download Refresh Delete

<input type="checkbox"/>	Sequence ID ↑↓	Query name ↑↓	Uploaded ↓↑	State ↑↓	Annotation	Delete
<input type="checkbox"/>	▼ A	Unspecified	2017-10-17 16:01	Waiting	Unavailable	
<input type="checkbox"/>	▼ Y	Unspecified	2017-10-17 15:55	Running	Unavailable	

Showing 1 to 2 of 2 entries

< 1 >

List of sequences

Upload sequence

Show 10 entries

Download Refresh Delete

<input type="checkbox"/>	Sequence ID ↑↓	Query name ↑↓	Uploaded ↓↑	State ↑↓	Annotation	Delete
<input type="checkbox"/>	▼A	Unspecified	2017-10-17 16:01	Finished	Download Annotation	Delete
<input type="checkbox"/>	▼Y	Unspecified	2017-10-17 15:55	Finished	Download Annotation	Delete

Showing 1 to 2 of 2 entries

< 1 >

A: 2.9 kbp の範囲を query から表示、塩基番号 1 から 2,900

Browser **Select Tracks** Snapshots Custom Tracks Preferences

検索

ランドマークまたは領域:

query:1..2,900

検索

出力 Decorated FASTA File 設定... 実行

Save Snapshot Load Snapshot

データソース

A

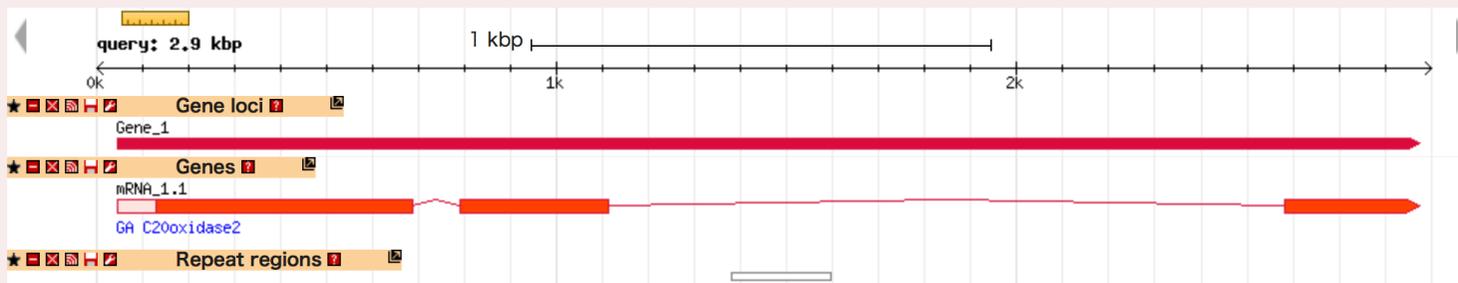
スクロール/ズーム: << < - 表示 2.9 kbp + > >> 反転

オーバービュー



領域

詳細ビュー



Select Tracks 強調表示を解除

mRNA_1.1

Attributes

Type	Gene supported by ESTs or cDNAs
Description	GA C20oxidase2
Gene coordinates	47 - 2873 (+ strand)
Gene length	2827 bp
Protein length	389 aa

Database hits

BLAST against UniProtKB (10 best hits)

Description	Species	ID	E-value	%Identity	%Coverage
SD1	<i>Oryza rufipogon</i>	G8HA36	0	100	100
SD1	<i>Oryza sativa</i>	G8HA06	0	100	100
GA C20oxidase2	<i>Oryza sativa</i> subsp. japonica	B6F2D9	0	100	100
Gibberellin-20 oxidase-2	<i>Oryza sativa</i> subsp. indica	F7J3E1	0	100	100
Gibberellin 20 oxidase 2	<i>Oryza sativa</i> subsp. japonica	Q0JH50	0	100	100
Gibberellin 20 oxidase 2	<i>Oryza sativa</i> subsp. indica	P0C5H5	0	100	100
SD1	<i>Oryza sativa</i> subsp. japonica	G8HA97	0	100	100
DNA-directed RNA polymerase subunit beta	<i>Oryza nivara</i>	A0A0E0FWN5	0	99	83
SD1	<i>Oryza glumipatula</i>	G8HA68	0	99	100
SD1	<i>Oryza sativa</i> subsp. japonica	G8HA95	0	99	100

Functional domains (InterPro)

Description	ID	Match position (aa)
Isopenicillin N synthase-like	IPR027443	30..363
Non-haem dioxygenase N-terminal domain	IPR026992	64..168
Oxoglutarate/iron-dependent dioxygenase	IPR005123	224..324

Gene Ontologies (GO)

Description	ID	Category
oxidoreductase activity	GO:0016491	Molecular function
oxidation-reduction process	GO:0055114	Biological process

Type	Start	End	Length
5'UTR	47	130	84 bp
CDS	131	687	557 bp
CDS	790	1111	322 bp
CDS	2583	2873	291 bp

```
>mRNA_1.1 position=query:47..2873 (+ strand)
CTGCACACAC ACACACACTC ACACTCACAC ACGCTCTCAA CTCAACACAG CGCTCACTTC TCATCTCCAA
TCTCATGGTG GCCGAGCACC CCACGCCACC ACAGCCGCAC CAACCACCGC CCATGGACTC CACCGCCGGC TCTGGCATTG
CGCCCCGGCG GCCGGCGGGC GTGTGCACC TGAGGATGGA GCCCAAGATC CCGGAGCCAT TCGTGTGGCC GAACGGCGAC
GCGAGCGCGG CGTCGGCGGG GGAGCTGGAC ATGCCCGTGG TCGACGTGGG CGTCTCCGCG GACGGCGAGC CCGAGGGGCT
GCGCCGCGCC GCGCGCGCAG TGCCCGCCGC GTGCCCCACC CACGGGTCTT TCCAGGTGTC CGAGCACGGC GTCGACGGCC
CTCTGGCGCG CGCCCGGCTC GACGGCGCCA GCGACTTCTT CCGCCTCCCG CTCGCCGAGA AGCCCGCGCG GCGCGCGCTC
CCGGGCACCC TGTCCGGCTA CACCAGCGCC CACGCCGACC GCTTCGCCTC CAAGCTCCCA TGGAAAGGGA CCCTCTCCTT
CGGCTTCCAC GACCGCGCGC CCGCCCCCGT CGTCGCCGAC TACTTCTCCA GCACCCCTCG CCCCAGCTTC GCGCAATGG
GTAATAA AAAACGATGCGT ACGACATTCG ATTTCAAATT CAAAACAAT TCAAAACACA CCGACCGAGA TTATGCTGAA
TTCAAACGCG TTTGTGCGCG CAGGAGGGTG TACCAGAAGT ACTGCGAGGA GATGAAGGAG CTGTGCGTGA CGATCATGGA
ACTCTGGAGC CTGAGCCTGG GCGTGGAGCG AGGCTACTAC AGGGAGTCTT TCGCGGACAG CAGCTCAATC ATGCGGTGCA
ACTACTACCC GCCATGCCCC GAGCCGGAGC GGACGCTCGG CACGGGCCCC CACTGCGACC CCACCGCCCT CACCATCCTC
CTCCAGGACG ACGTCGGCGG CCTCGAGGTC CTGCTGGAGC GCGAATGGCG CCGCTCCAGC CCGCTCCCGG GCGCCATGGT
CATCAACATC GCGCACACCT TCATGTAAA CCATCTCCTA TTCTCCTCTC CTCTGTTCTC CTCTGCTTCG AAGCAACAGA
ACAAGTAATT CAAGCTTTTT TTTCTCTCTC GCGCGAAATT GACGAGAAA ATAAGATCGT GGTAGGGGCG GGGCTTCAG
CTGAAGCGCG GAAGAAACCG ACCTGACGTG ATTTCTCTGT TCCAAACACA ACAAATGGAA TGCCCCACTC CTCCATGCTG
TATGATTAT CTACATCTT ATAGTTAATA GGAGTAAGTA ACAAGCTACT TTTtcatat tatagttcgt ttgattttt
ttttttaaag tttttttagt tttatccaaa tttattgaaa aacttagcaa cgtttataat accaaattag tctcatttag
tttaattatt tataatattt gataatatat ttatgttata ttaaaaatat tactatattt ttctataaac attatataaa
gccattata atataaaatg gaaggagtaA TTAATATGGA TCTCCCCGA CATGAGAAAT TTTTCCGATG GTGTGACGAC
GCCATGTAAG CTTCGGTGGG CTTGACGGC CAGAGGTGCC AACAGCCAG TCCAAACAAC CCTGGGTCCC CCCCTAACAC
TCCAAACAGT AGTGAAGTAG GTCTCGTCCG GTTTTAGTAT TTGATGACAA ACAAAGTGTG AGTTGAGTGA GCCACCACA
ACTTGCACAC GAGCACATAC ATTTGTGTCC ATTCCTGCCA GTCACCTCCA TCTCTAGTCC TAACCTCTAT CTAGCGATGT
AAGCGGATAA TTTTATCATC CGTATATAAA CCTGTTTGTT ATAGTTAATT TCCTATATAA TACTATAACA GTATACATT
TAAAAGAAA CAAAATTAGG ATAAACAGGC CCTGCTCCTA TCCATCCATG GCACCTGGAA GGACCAGACT CGGTATGTC
ATGCCAAGCC AAGATATGGG TTAATGGAAGA GTAGAGAAGA GGAGAGATGA GAGATAAGCA TCGCTTCTCC TCCTCGTTGG
ATGTGTATT TGGAGGGATT TGTGTAGTAG TAGCAGCGGC CCGCCGGGGA CCGATGCGGA TGGTGGCGCT TCGGTGGCG
TTTTCCCGGG GGGGTTTTGG TTTGGCCCTT GGGGGGATG GCATGCGCGC GCGTGGCGCT GCACGCCACA CACACGGCGC
CGCACCAGC TACGTCGTCG TCGCCCGCGG CCGCAGGTAG CTTAGGTTGG TGTGTTCCGC GCGCGGGCGG GGATTTGTTCC
ATGCCGATCG ATTTGGCGCC ACCCTCGCGC CGGCTCTTGT CCGCTCGTGC GCTCTCTCG CCGGTTTTGT CTTGTGCGG
TTGCTCAGCC GCGCAGCGGG GCACGGACAT TGGCGATGTA GCCCTGCACG TGTGGCCCTC TCCGTTGATG AATGATGAT
TATGATGTA TTTTTTTTGT TCTGAAGGAA TTTGTGGGGA ATTTGTTGTT GTGCAGCGCG TGTGCAACGG GAGGTAAG
AGCTGCCTGC ACAGGGCGGT GGTGAACAG CCGCGGGAGC GCGGTCGCT GGCCTTCTTC CTGTCCCGCC GGGGACAC
GGTGGTGGG CCGCCGCGCA GCGCCGCCAC CGCCGAGCAC TACCCGGACT TCACTGGGCG CGACCTCATG CGCTTCACCG
AGCGCCACTA CCGCGCCGAC ACCCGCACCG TCGACGCGTT CACGGCGTGG CTGCGCGCGC CCGCCCGCGA CCGCGCGCGG
ACGGCGCAGG TCGAGGGCGC CAGCTGA
```

Sequences

ORF sequence

```
>mRNA_1.1 ORF sequence
ATGGTGGCGAGCACCCACGCCACCACAGCCGCACCAACCACCGCCCATGGACTCCACC
GCCGGCTCTGGCATTGCCGCCCCGGCGCGCGCGCGGTTGCGACCTGAGGATGGAGCCC
AAGATCCCAGGACCATTCGTGTGGCCGAACGGCGACGCGAGGCGCGCTCGCGCGCGAG
CTGGACATGCCCTGGTTCGACGTGGCGTGTCTCGCGACGGCGACGCCGAGGGGCTCGC
CGCGCCGGCGCGAGGTGGCCCGCGCTGCGCCACGACGCGGTTCTTCCAGGTGTCGAG
CACGGCGTCGACCGGCTCTGGCGCGCGCGCGCTCGACGGCGCCAGCGACTTCTCCCG
CTCCCGCTCGCCGAGAAGCGCCGCGCGCGCGCTCCGGGACCGTGTCCGGCTACACC
AGCGCCACGCCGAGCGCTTCGCTTCCAAGCTCCCATGGAAGGAGACCCCTCTCTCCGGC
TTCCACGACCGCGCGCGCGCCCCGCTCTGCGCCGACTACTTCTCCAGCACCTCGGCCCC
GACTTCGCGCAATGGGGAGGGTGTACAGAAGTACTCGGAGGAGATGAAGGAGCTGTG
CTGACGATCATGGAACCTCTGGAGCTGAGCCTGGGCGTGGAGCGAGGCTACTACAGGAG
TCTTTCGGGACGACGACTCAATCATCGGTTGCAACTACTACCCGCCATGCCGGAGCCG
GAGCGGACGCTCGGCACGGGCCCGCACTGCGACCCACCGCCCTCACCATCTCTCCAG
GACGACGTCGGCGGCTCGAGTCTCTGTCGACGGCGAATGGCGCCCGTCTAGCCCGCTC
CCGGCGCCATGGTTCATCAACATCGCGCACACTTTCATGGCGCTGTGCAACGGGAGTAT
AAGAGTGCCTGCACAGGGCGGTGGTGAACAGCGCGGGGAGCGCGGCTCGTGGCGTTC
TTCCTGTGCCCGGGAGGACAGGCTGGTGGCGCCCGCGGAGCGCGCCACGCGCCAG
CACTACCCGGACTTCACTGGGCGGACTCATGGCTTACGCGACGCGCCACTACCGCGC
GACACCCGACGCTCGAGCGCTTACGCGCTGGCTCGCGCGCGCGCGCCGACGCGCGC
GCGACGGCGAGGTCGAGGGCGGCACTGA
```

Protein sequence (translated ORF)

```
>mRNA_1.1 Protein sequence
MVAEHPPTPPQHPQPPMDSTAGSGIAAPAAAAVCDLRMEPKIPEPFVWPNGDARPASAAE
LDMPVVDVGLRDGDAEGLRRAAAQVAACATHGFQVSEHGVDAAALARAALDGASDFFR
LPLAEKRARRRVPPTVSGYTAHADRFASKLPWKEITLSFGFHDRRAAPVVDYFSSITLGP
DPAEMCRVYVRYPERMKEITLSTIMLLELSTLGVRCYVDFEFASSIMRNVYDPCDEP
```

5'UTR	47	130	84 bp
CDS	131	687	557 bp
CDS	790	1111	322 bp
CDS	2583	2873	291 bp

```
>mRNA_1.1 position=query:47..2873 (+ strand)
CTGCACAC ACACACACT ACACTCACAC ACGCTCTCAA CTCACTCCCG CTC AACACAG CGCTCACTTC TCATCTCCAA
TCTCATGGTG GCCGAGCACC CCACGCCACC ACAGCCGCAC CAACCACCGC CCAATGGACTC CACCGCCGGC TCTGGCAATG
CCGCCCCCGG GCGCGCCGCG GTGTGCGACC TGAGGATGGA GCCCAAGATC CCGGAGCCAT TCGTGTGGCC GAACGGCGAC
GCGAGGCCGG CGTCCGGCGG GGAGCTGGAC ATGCCCGTGG TCGACTGGG CGTGTCCGCG GACGGCGACG CCGAGGGGCT
GCGCCGCGCG GCGGCGCAGG TGGCGCCGCG GTGCCCCACG CACGGTCTCT TCCAGTGTCT CGAGCACGGC GTCCAGCCCG
CTCTGGCGCG GCGCCGCTC GACGGCGCCA GCGACTTCTT CCGCTCCCG CTCGCCGAGA AGCCCGCGCG GCGCGCGCTC
CCGGGCACC TGTCCGGTA CACCAGCGCC CACGCCGACC GCTTCGCCTC CAAGCTCCCA TGGAAAGAGA CCTCTCCTT
CGGCTCCAC GACCGCCCG CCGCCCCCGT CGTCCCGGAC TACTTCTCCA GCAACCTCGG CCGGCACTTC GCGCAATGG
GGTAATTAAA ACGATGGTGG ACGACATTGC ATTTCAAAT CAAAACAAAT TCAAACACA CCGACCGAGA TTATGCTGAA
TTCAAACCGC TTTGTCCGCG CAGGAGGTG TACCAGAAGT ACTGCGGAGA GATGAAGGAG CTGTCCGTGA CGATCATGGA
ACTCTTGAG CTGAGCCTGG SCGTGGAGCG AGGCTACTAC AGGAGTTCT TCGCGGACAG CAGCTCAATC ATCGCGTGCA
ACTACTACC GCCATGCCCG GAGCCGAGC GGACGCTCG CACGGGCCCG CACTGCGACC CCACCCGCTC CACCATCCTC
CTCCAGGACG ACGTCCGGCG CCTCGAGTCT CTCGTCGACG GCGAATGGCG CCGCTCCCG CCGCTCCCG GCGCATGGT
CATCAACATC GCGCACCTT TCATGGTAAA CCATCTCCTA TTCTCTCTCT CTCTGTCTCT CTCTGTCTCG AAGCAACAGA
ACAAGTAATT CAAGCTTTT TTTCTCTCTC GCGCGAAAT GACGAGAAA ATAAGATCGT GGTAGGGGCG GGGCTTTCAG
CTGAAAGCGG GAAGAAACCG ACCTGACGTG ATTTCTCTGT TCCAATCACA ACAAATGGAA TGCCCACTC CTCCATGTGT
TATGATTTAT CTCACATCTT ATAGTTAATA GGAGTAAGTA ACAAGTACT TTTtcatat atagttcgt ttgattttt
tttttaaaag ttttttagt tttatccaaa tttattgaaa aacttagcaa cgtttataat accaaatag tctcatttag
tttaaatgt tataatattt gataatatat ttatgttata ttaaaaatat taactatatt ttotataaac attattaaaa
gccatttata atataaaatg gaagagtaA TTAATATGGA TCTCCCCGA CATGAGAATA TTTTCCGATG GTGTGACGAC
GCCATGTAAG CTCCTGGTGG CCTGGACGCG CAGAGGTGCC AACAGCCAGC TCCAAACACC CCGTGGTCCC CCCCACAC
TCCAAACAGT AGTGAAGTAGT GTCTGTCTCG GTTTTAGTAT TTGATGACAA ACAAAAGTGT AGTTGAGTTA GCCACCACA
ACTTGACAC GAGCACATAC ATTTGTGTCC ATTTCTCGCA GTCACTTCCA TCTCTAGTCC TAACCTCTAT CTAGCGATGT
AAGCGGATAA TTTTCATCAT CGTATATAAA CCTGTTTGT ATAGTTAAT TCCATATAAA TACTATAACA GTATACATT
TAAAAGAAAA CAAAATAGG ATAAACAGGC CCTGTCCTA TCCATCCATG GCATTTGAA GGACAGACT CGGTATGCC
ATGCCAAGCC AAGATATGGG TTAGGAGA GAAGAGATGA GAGATAAGCA TCGCTTCTC TCCCTGTGG
ATGTGATTT TGGAGGGATT TGTGTAGTAG TAGCAGCGCG CCGCGGGGA CCGATGCGGA TGGTGGCGCT TTCGGTGGCG
TTTTCGCCGG GGGGTTTTG TTTGCCGCTT GGGGGGATG GCATGGCGCG CCGTGGCGCG GCGTGGCGCG CACACGGCGG
CGCACGACG TACGTCGTCG TCGCGCGGG GCGACGCTAG CTTAGGTTGG TGTGTTCCG GCGCGGGCG GGTATGTTCC
ATGCCGATG ATTTGGCGCC ACCCTCGCGG CGGCTTGT CGGCTCGTGC GCTCTCTCG CCGGTTTTGT CCTGTGCGCG
TTGCTCAGC GGGACGGGG GCACGGCAT TGGGATGTA CCCCAGCAG TGTCCGCTC TCGTGTGATG AATGATGATG
TATGTATGTA TTTTTTTTT TCTGAAGGAA TTTGTGGGA ATTTGTGTGT CTGCAGGGC TCTGAAGCG GAGTATAAG
AGCTGCCTGC ACAGGGCGGT GGTGAACCAG CCGCGGAGC GGGGTGCGT GGCCTTCTC CTGTGCCCG GGGAGGACG
GGTGTGCGG CCGCGCCGA GCGCGCCAC GCGCGAGC TACCCGACT TCACTGGGC CGACTCATG CGCTTACGC
AGCGCCACTA CCGCGCGAC ACCCGCAGC TCGACGCTT CACGCGTGG CTGCGCGCG GGGCCCGGA CCGCGCGG
ACGGGCGAG TCGAGCGCG CAGCTGA
```

Sequences

ORF sequence

```
>mRNA_1.1 ORF sequence
ATGGTGGCCGAGCACCCACGCCACCACGCGCCACCAACCACCGCCATGGACTCCACC
GCCGGCTTGGCAATTGCCGCCCGCGCGCGGGTGTGGCACTGAGGATGGAGCCC
AAGATCCCGGAGCCATTCTGTGGCGGAACGCGCAGCGGAGCCCGCTCGCGCGGAG
CTGGACATGCCCGTGGTCGACGTGGCGTGTCTCGCGACGGCGCAGCGAGGGGTGCGC
CGCGCGCGCGGAGGTGGCGCGCGCTGCGCCACGACGCGGTTCTTCCAGGTGCCGAG
CACGGCTCGACGCGCTCTGGCGCGCGCGCTCGACGGCGCAGCGACTTCTCTCGC
CTCCCGCTCGCGAGAACGCGCGCGCGCGCTCCCGGCGACGCTCGCGCTACACC
AGCGCCACGCGACCGCTTCCCTCACAAGTCCCATGGAAGGAGACCTCTCTTCGGC
TTCACGACCGCGCGCGCGCGCGCTCGTCCGCGACTACTTCTCAGCACCTCGGGCCC
GACTTCGGCCAAATGGGAGGTTTACCAGAAGTACTCGGAGGAGATGAAGGAGCTGCG
CTGACGATCATGGAATCTCTGGAGCTGAGCTGGCGTGGAGCGGCTACTACAGGAG
TTCTTCGGGACGAGCTCAATCATCGGCTGCAACTACTACCCGCCATGCCGGAGCGG
GAGCGGACGCTCGGCACGGGCGCGCACTCGGACCCACCGCCCTCACCATCTCTCCAG
GAGCAGCTCGCGCGCTCGAGTCTCTGTGACGGGCAATGGCGCCCGTACGCGCGCTC
CCCGCGCATGGTATCAACATCGCGCACACTTATGGCGCTGCGAACGGGAGGTAT
AAGAGCTGCTGCACAGGGCGGTGGTGAACAGCGCGGGGAGCGCGGTCTGCTGGCTTC
TTCTGTGCGCGGGGAGGACAGGTTGGTGGCGCGCGCGGAGCGCGCCACCGCGAG
CACTACCCGGACTTCACTGGGCGGACCTCATGCGCTTACGACGCGCCACTACCGCGCC
GACACCCGACGCTCGACGCTTACGCGCTGGCTCGCGCGCGCGCGCGCGCGCGCGCC
GCGACGCGCGAGTTCGAGCGGCGGCTGAGCGGCGGCTGAGCGGCGGCGCGCGCGCG
```

Protein sequence (translated ORF)

```
>mRNA_1.1 Protein sequence
MVAEHPTPPQHPQPPMDSTAGSIAAPAAAAVCDLRMEPKIPEPFVWPNGDARPASAAE
LDMPVVDVGLRDGAELGRRAAAQVAACATHGFQVSEHGVDAAALARAALDGSDFFR
LPLAEKRARRRVPGTVSGYTSAHADRFASKLPWKEITLSFGFHRAAAPPVADYFSSLTGP
DFAPMRVYQKYEEMKELSLTIMELLESLGVERGYREFFADSSIMRNCYPPPEP
ERTLGTGFHCDPTALITLLQDDVGLLEVLVDGEWRVPSVPGAMVINIGDTFMALSNRY
KSCLHRAVNVQRERRSLAFLCPREDRVVPPSAATPQHYDPFTWADLMRFQRYRA
DTRTLDAFTRWLAPFAADAATAQVEAAS*
```

List of sequences

Upload sequence

Show 10 entries



<input type="checkbox"/>	Sequence ID	Query name	Uploaded	State	Annotation	Delete
<input type="checkbox"/>	▼ A	Unspecified	2017-10-17 16:01	Finished		
<input type="checkbox"/>	▼ Y	Unspecified	2017-10-17 15:55	Finished		

Showing 1 to 2 of 2 entries

Y: 3.42 kbp の範囲を query から表示、塩基番号 1 から 3,420

Browser [Select Tracks](#) [Snapshots](#) [Custom Tracks](#) [Preferences](#)

検索

ランドマークまたは領域:

query:1..3,420

検索

出力 Decorated FASTA File 設定... 実行

Save Snapshot Load Snapshot

データソース

Y

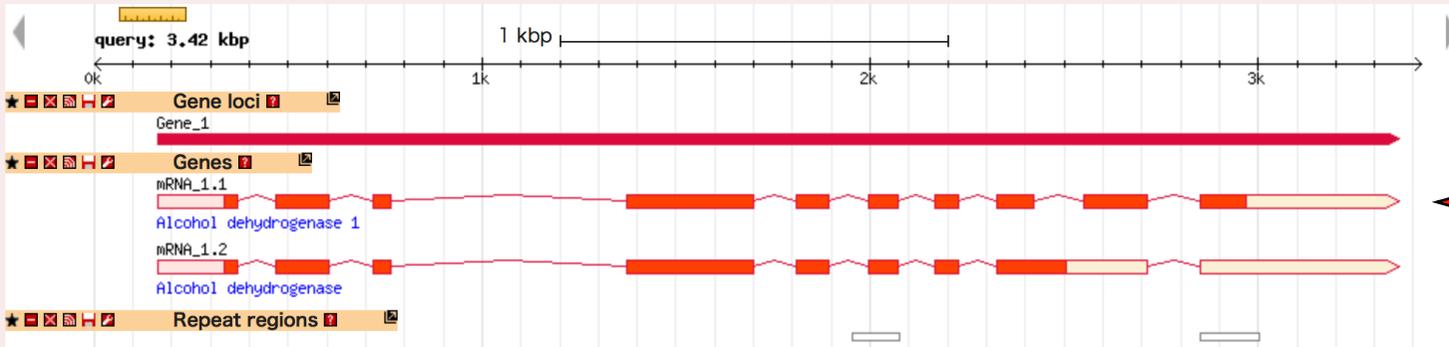
スクロール/ズーム: <<< << < 表示 3.42 kbp > >> >>> 反転

オーバービュー



領域

詳細ビュー



Select Tracks 強調表示を解除

Gene structure

Type	Start	End	Length
5'UTR	166	336	171 bp
CDS	337	370	34 bp
CDS	469	605	137 bp
CDS	718	764	47 bp
CDS	1373	1698	326 bp
CDS	1811	1893	83 bp
CDS	1996	2071	76 bp
CDS	2166	2227	62 bp
CDS	2326	2421	96 bp
CDS	2552	2713	162 bp
CDS	2850	2966	117 bp
3'UTR	2967	3362	396 bp

>mRNA 1.1 position=query:166..3362 (+ strand)

```

GCCCGCAGAG CTCTACCAC TCTCTGATCC CTATATATAC ACCCCCTCCT CCCCTCTCGT CTCTCACCA GAATTCGAAG
CAACGAACGT CGAGTGATTC AAGAAAAAAG AAAACCTGAG CTTTCGATCT CTTCCGAGTG GTTTCCTTGT CTTTGAAAAA
GAGGGGATTC AATGGCGACA GCCGGGAAGG TGATCAAGTG CAAAGGTCCG TTCTTGCTGT TCTTGGAGCG GTTCGTGAGA
AAATAGTTT GATGTTTGT GGGGAGGAGA AGCTGATAG ATCGGTTTTT GGGGTTTTTG CAGCGGGCGT GGCATGGGAG
GCCGGGAAGC CGCTGTCCGAT CGAGGAGGTG GAGGTGCGC CGCCGAGGC CATGGAGTTC CGCTCAAGA TCCTCTACAC
CGCCCTCTGC CACACCGAGC TCTACTTCTG GGAGCCAAAG GTAATTAATC AGCATCTGCC ATGGCTGAGC TTTCTCACCT
TCTCAAGCTT GCAAGAATGA TCACTGACTA ACTAACTTGG TGTGTGTGTG GTTTCGTCTT TGATCTTCGC AGGGCAAAAC
ACCTGTTTTT CCTAGGATCT TGGCCATGA AGCTGGAGGG TATGTGTAAT CTGAACCAAC CCTGTCCAAT TCTGTCTCTT
CTACTACACT GTTTAGTGGT TCACTCTATT CTGTCTTGT ATTTGTTTTT TTTTGTTTTA AAATGGCTTT TGGTGTTTAG
TGGAATTTAG TGTGTACTTG TACGTCTATG GTTGTTTGGT TTATGGTATT CGGTACTTAA CTATAGTTAG TGTGCACGTT
GGGAGTCCCT GATTAGTCAG TGGCATTTTC ATTAAGTCA GTAAAGAAAT CTACCAAGTG TCCAATTTGT CATGATATTT
TACTGGTTTG GTATAAAATC ATATGAAGAA TCAAAGAACA GCGTCATTGA TGACATATAC AACGACAAT GGATCCCTAG
GATAATATAC TTTTCTTCAT GTGCTCCCTC ATGCTAGTGC TTTATTACTT TCTCTCTTTT GATGTTTTGA GGTACTCAA
GGAAAAAGAC ACACCTGAATC ATTATCCATC TTATGAACTG TTTAAACTTG ACTTACTTAT CACTTCTCTG CTCAAATGGA
CCCTTTTTTT TTTTCCATTT CCATGGAAGA TATATGTGTA ATCTAACAA TCAAATTGAT CTAACAAAAC AATTTTTTTG
CTCAGCAT TGTGGAGAGT GTGGGAGAGG GTGTGACCGA ACTCGCGCCG GCGGACCATG TCCTCCCGGT GTTCACCGCG
GAGTGCAAGG AGTGTGATCA CTGCAAATCG GAGGAGGACA ACATGTGTGA CCTCCTCAGG ATCAACGCTG ACCGGCGCT
CATGATCGGC GACGGCAAGT CCGGATTCAC CATCAAGGGG AAGCCATCT TCCACTTTGT TGGCACATCC ACCTTCAGCG
AGTACACCGT CATCCATGTC GGCCTGCTCG GAAAGATCAA CCGGAGGCG CCTCTCGACA AGGTCTGCAT TCTCAGCTGC
GGTTTCTCTA CCGGTAAGAT GATGGATGGC TTCATGATTA TGATAGTAGT AGTTTTTACA GTGAAATTTG AAAGTCTTTT
TTTTTTTTGA TGCATTAAGT TGATGGGATT TTGCTGTCT TGCAGTTTTT GGTCAACAG TGAATGTCG GAAACCGAAA
AAGGGCCAGA CCGTGGCTAT TTTCCGCTCT GGAGCTGTTG GTCTTGCTGT AAGTGACAG CACAAGATA TATCATTTCC
ATTTGATTTG CAGATTGTTG AGATCTTCAA actcaaatg atgaatgttt gtaaaattac atgattatag gctatggaag
gtgccaggct gtctggagca tggaggatca ttgggtggtg cctgaacct gogaattcg aacaaggtac AATCTTTTAT
CATCTTTGAG TGAATAATC ATGTTTGGTA TATGTTCTGG ATTGAACACT GAACCTGCA ACTGTTTTT ATCTCTTCAG
CTAAGAAGTT TGGTGCAC TACTCTGTAA ACCCCAAGGA CCACAGCAAG CCACTGCAGG AGTACATCA CATTGCTCA
CTAATTTCTC TTATATCAC ACACCTTCA TGACATGGAG TGTACATACT TAACATCAA CTTGCAATG GTGATCTCAG
GTGCTGATG AGATGACCAA CGGTGGACTC GACAGGGCCG TTGAATGCAC GGGCAACATC AACCCATGA TATCTGCTT
CGAATGTGTC CATGATGTC GTCTCTTCT TACAACATG AAACATGATA TGAACAAAT GAAGATGATC ATTTGAGCAA
TGCTCCAGT TCTGCTGTAA ACTGTATCCT GAAACTAAG AGTGGCAATC TGAACCTGA ATCCAGGAT GGGCGTGG
GGTCTGGTC GCGTGGCGA CCAAGGATGA TGTGTTCAAG ACCCACCGA TGAATTTCT CAACGAGAAG ACCTCAAGG
GGAGTCTCT CGCTAACTAC AAGCCGCGA CCGACCTGCC CAACGCTCT GAGCTGTACA TGAAGAAGT AAAATAAAT
GCAGAACAGA ATCAGATTTG CAACCTGAAT GTGCATTTGT CAGAAATCTA TGGTCACTGA TTCCAATGAG CCACTGATTTG
ATCACTATTT TATGCTAATG ATTTCTGGTG TTTGATGCT GAAGGAgctg gagctggaga agtcaatcaac gcaacagctg
ccgttctcgg agatcaaacac ggcgttggat ctcactgca aggggggagag tctcctctgc gtcactgaga tggatgagta
gagatgatga tgatctatct ctctctctct ctctctTGCA GTTTCAGTGC ATACAAGAAA CCAAAAATGA AAGCAAAAA
TGGATGAGAA GTTTGAGGTT TTCCATTTAA GTTTAATAAT GGTCTGAGA CAGGGCTTGT GTGGAGTCAG CTCTGTTTAC
TCCCTCCTTT CATGGCTTGA TCTTGTTTAT TGTCAGCATT GTGGCTTAGC AGATTTTAC CAGTTTTTACC CTGTAATAA
TGACCATCT GAAATCGTAA TAAACTACTA CGAGTGTGAG GTACTATTGC TGAACCTGAA TTGATTTGTC AGTAAAGAGG
TGAATAGCAG CCAATATGAA GGTTTTACAA TTTGAACATC GAAACGAAAT CCACACGAAA TTCATAAAT TCATCGC

```

Sequences

ORF sequence

>mRNA 1.1 ORF sequence

```

ATGGCGACAGCCGGGAAGGTGATCAAGTGCAAAGCGCGGTGGCATGGGAGGCCGGGAAG
CCGCTGTGATCGAGGAGGTGGAGTTGCGCCCGCCGAGGCCATGGAGGTTCCGCTCAAG
ATCCTTACACCGCCCTCTGCCACACCGACGCTACTTCTGGGAGGCAAGGGCAAAACA
CCTGTTTTCCCTAGGATCTTGGCCATGAAGCTGGAGGCATTTGGGAGGTGTTGGAGAG
GGTGTGACCGAACTCGCGCCGGGACCATGCTCCCGGTTTCAACCGGAGTGCAG
GACTGTGATCACTCAAATCGGAGGAGCAACATGTTGACCTCTCAGGATCAACCTC
GACCCGGGCTCATGATCGCGAGCGCAACTCCGATTACCATCAAGGGGAAGCCATC
TTCCACTTTGTTGGACATCCACTTCCAGCGAGTACACCGTCACTCATGTCGGCTGCCTC
GCGAAGATCAACCCGGAGGCCCTCTCGACAAGTCTGCACTTCTCAGCTGCGGTTCTCT
ACCGTTTTGGTGCACAGTGAATGTCGCGAAACCGAAAAAGGGCCAGACCGTGGCTATT
TTCGCTCTGGAGCTGTTGCTTTGCTgctatggaaggtgccaggctgtctggagcatcg
agcatgattgtctggagctcaagcctgcaaatgcaaacagCTAAGTTCCTGCTGCTGCTG

```

CDS	2532	2713	182 bp
CDS	2850	2966	117 bp
3'UTR	2967	3362	396 bp

```
>mRNA_1.1 position=query:166..3362 (+ strand)
GCCCGGAGA CTCTACAC TCTCTGATCC CTATATATAC ACCCCCTCCT CCCCTCTCGT CTCCTCACCA GAATCCAAG
CAACGAAGT CGAGTGATTC AAGAAAAAG AAAACCTGAG CTTTCGATCT CTTCCGGATG GTTTCGTGTT CTTTGAAGAA
GAGGGGGATT AATGGCGACA GCCGGGAAGG TGATCAAGTG CAAAGTCCCG TTCCTGCTGT TCTTGGAGCG GTTCGTGAGA
AAATTAGTTT GATGTTTTGT GGGGAGGAGA AGCTGATTAG ATCGGGTTT GGGGTTTTG CAGCGGGGT GGCATGGGAG
GCCGGGAAAG CGCTGTGAT CGAGGAGGT GAGGTTGCGC CGCCGAGCG CATGGAGTC CGCTCAAGA TCCCTACAC
CGCCCTCTG CACACCCGAC TCTACTTCTG GGAGCCCAAG GTAATTAATC AGCATCTGCC ATGGCTGAGC TTTCTCACCT
TCTCAAGCTT GCAAGAATGA TCACTGACTA ACTAAGTTGG TGTGTGTGTG GTTCTGCTT TGATCTTGC AGGGCAAAAC
ACCTGTTTTT CCTAGGATCT TGGGCCATGA AGCTGGAGGG TATGTCTAAT CTGAACCAAC CCTCTCCAAT TCTGTTCTCT
CTACTACACT GTTTAGTGGT TCACTCTATT CTGTTCTTCT ATTTGTTTTT TTTTCTTTTA AAATGGCTTT TGGTCTTAG
TGAATTTAG TGTGTACTTG TACGTACATG GTTGTGTTGG TTATGTTATT CGGTTACTAA CTATAGTTAG TGTGCACGTT
GGGAGTCTT GATTAGTCAG TGGCATTTTC ATTAAGTTCA GTAAAGAAAT CTACCAAGTG TCCAATTTGT CATGATATTT
TACTGGTTTG GTATAAAATC ATATGAAGAA TCAAGAACA GCGTCATTGA TGACATATAC AACGACAATT GGATCTCTAG
GATAATATAC TTTTCTTCAT GTGCTCCTCT ATGCTAGTGC TTTATTAATT TCTCTCTTTT GATGTTTTGA GGTACTCAAA
GGAAAAAGAC ACACGAATC ATTAATCCATC TTATGAAC TGTTAAACTTG ACTTACTTAT CACTTCTCTG CTCAAATGGA
CCCTTTTTTT TTTTCCATTT CCATGGAAGA TATATGTGTA ATCTAACAA TCAAATTGAT TCAACAAAAC AATTTTTTTG
CTCACAGCAT TGTGGAGAGT GTGGGAGAGG GTGTGACCGA ACTCGCGCCG GCGGACCATG TCCCTCCGGT GTTCACCGGC
GAGTGCRAAG AGTGTGATCA CTGCAATTCG GAGGAGGCA ACATGTGTGA CCTCCTCAGG ATCAACCTCG ACCGCGCGT
CATGATCGGC GACGGCAAGT CCCGATTCAC CATCAAGGGG AAGCCCATCT TCCACTTTGT TGGCACATCC ACCTCAGCG
AGTACACCGT CATCACTGTC GGCTGCTTCG CGAAGATCAA CCCGGAGGCG CCTCTCGACA AGGCTCGCAT TCTCAGCTGC
GGTTCCTCA CCGTAAGAT GATGGATGGC TTCATGATTA TGTAGTAGT AGTTTTCACA GTGAATTTG AAAGCTTTT
TTTTTTTTGA TGCATTAAGT TGAATGGATT TTGCTGTTCT TGCAGTTTTT GGTGCAACAG TGAATGTCGC GAAACGGAAA
AAGGGCCAGA CCGTGGCTAT TTCGGTCTT GGAGCTGTTG GTCCTGCTGT AAGTGACACG CACAAGATA TATCATTTCC
ATTTGATTG CAGATTGTTG AGATCTTGAA actcaaaact atgaatgttt gtaaaattac atgattatag gctatggaag
gtgcaaggt gtctgagaca tggagatca ttggtgtgga cctgaacect gogaattog aacaagtaac AATCTTTTAT
CATCTTTGAG TGAAAATTAC ATGTTTGTA TATGTTCTGG ATTGAACACT GAACCTGCA ACTTGTTTTC ATCTTCTCAG
CTAAGAGTTT TGGCTGCAT GACTTCGTAA ACCCAAGGA CCACAGCAAG CCGATGCAG AGGTACATCA CATTCCTCA
CTAATTTCT TATATACAC ACATGTTCA TGACATGGAG TGTACATACT TAAACATCAA CTCGCAATG GTGATCTCAG
GTGCTGATG AGATGACCAA CGGTGGACTC GACAGGGCCG TTGAATGCAC GGGCAACATC AACGCCATGA TATCTGCTT
CGAATGTGC CATGATGTA GTCTCTTTCT TACAACATG AACATGATA GTAACAAAT GAAGATGATC ATTTGAGCAA
TGCTCCAGT TCTGCTGTA ACTGTATCCT GAAACTAAG AGTGGCAATC TGAACCTGA ATCCAGGAT GGGCGTTCG
GGTGTGGTC GGGGTGCCA CCAAGGATGA TGTGTTCAAG ACCCAGGCA TGAACCTCT CAACGAGAG ACCTCAAGG
GGAGTCTT CCGTAACAT AAGCCGCGCA CCGACTGCC CAACCTGCT GAGCTGTACA TGAAGAGGT AAAAATAAT
GCAGAACAGA ATCAGATTG CAACCTGAAT GTGCATGTT CAGAATCTA TGTCACTGA TTCCAATGAG CCACTGATTG
ATCACTATTT TATGCTAATG ATTTCTGGTG TTTGATGTT CAAGGagctg gagctggaga agtctcaac gcaagagctg
cgttctcgg agatcaaac ggggttggat ctcctgctca agggggagag tctcggctgc gtcctgagga tggatgagta
gagatgatga tgatctatct ctctctctct ctctctTGCA GTTTCAGTGC ATACAAGAAA CCAAAAATGA AAGCAAAAA
TGGATGAGAA GTTTGAGGTT TTCCATTTAA GTTTAATAAT GGTCTGAGA CAGGGCTGT GTGGAGTCAG CTCTGTTTAC
TCCCTCCTTT CATGGCTTGA TCTTTTTFAT TGTACGCATT GTGGCTTAG AGATTTTAC CAGTTTTTAC CTGTAATAA
TGACCATTT GAAATCGTAA TAAACTACTA GCAGTGTGAG TGAATTTGC TGAACGTAA TTAGTTGTGC AGTAAAGAGG
TGAATAGCAG CCAGTATGAA GGTTTTACAA TTGAACATC GAAACGAAAT CCACACGAAA TTCATAAATT TCATCCG
```

Sequences

ORF sequence

```
>mRNA_1.1 ORF sequence
ATGGCGACAGCCGGGAAGGTGATCAAGTCAAAAGCGCGGTGGCATGGGAGCCGGGAAG
CCGCTCTCGATCGAGGAGGTGGAGTTGCGCCCGCAGGCGCATGAGGTTCCCGTCAAG
ATCCTCTACACCGCCCTCTGCCACCGACGCTACTTCTGGGAGGCAAGGGGCAAAACA
CCTGTTTTCCCTAGGATCTTGGGCCATGAAGTGGAGCATTTGAGAGAGTGTGGAGAG
GGTGTGACCGAATCGCGCCGGGACCATGTCCTCCGGGTTCACCGCGAGTGAAG
GAGTGTGATCACTGCAAAATCGGAGGAGGCAACATGTGTGACCTCCTCAGGATCAACGTC
GACCGGGCGTCAATGATCGGCGACGGCAAGTCCGATTCACCATCAAGGGGAAGCCATC
TTCACATTTGTTGGACATCCACCTTCAGCGAGTACACCGTCACTCATGTCGGCTGCCTC
GCCAAGATCAACCCGAGGCGCCTCTCGCAAGSTCTGCAATTCAGCTCGGTTTCTCT
ACCGGTTTTGGTGCAACAGTGAATGTCGGGAAACGAAAAAGGGCCAGACCCGTGGCTATT
TTCCGCTTGGAGCTGTGGTCTTGTGctatggaaggtgcccaggtgctctggagatcg
aggatcattggtgtggacctgaacctgcaaaatcgaaacaagCTAAGAAGTTTGGCTGC
ACTGACTTCGTAACCCCAAGGACCAAGCAAGCCAGTGCACAGGCTGTGATTGAGATG
ACCAACGTTGGACTCGACAGGGCCGTTGAATGCACGGGCAACATCAACGCCATGATATCC
TGCTTCGAATGTGTCATGATGATGGGCGTTCGGGTGCTGTTGCGGCTGCCGCAACAG
GATGATGTGTTCAAGACCCACCCGATGAATTCCTCAACGAGAAAGACGCTCAAGGGGACG
TTCTTCGGTAACTACAAGCCGCGCACCGACTGCCAACAGTCTGAGCTGTACATGAAG
AAGGAgctggagctggagaagttcatcacgacagcgtgcccgttctcggagatcaaacag
gcgtctgatctcatgctcaacggggagagctctccgctgctgctgagatgagatgagtag
```

Protein sequence (translated ORF)

```
>mRNA_1.1 Protein sequence
MATAGKVIKCKAAVAWEAGKFLSIEEVEVAPQAMEVRVKILYALCHTDVYFWEAKGQT
PVFPRILGHEAGGIVESVGEVTEAPGDHVLVPTGCEKCDHCKSESNMCDLLRINV
DRGVMIGDKSRFTIKKPIFHFVGTSTFSEYTVIHVGLAKINPEAPLDKVCILSCGFS
TGFATVNVAKPKKQTVAFIPLGAVGLAAMEGARLSGASRIIGVDLNPAPKFEAKKFGC
TDFVNPKDHSPVHEVLIEMTNGLDRAVECTGNINAMISCFECVHDGWAVLVGVPTK
DDVFKTHPMNLFNEKILKGTFFGNYPKPTDLNPNVVELYMKKELEKFIHSHVPSFSEINT
AFDMLKGESLRCVMRMD*
```

DNA配列の中から遺伝子を予測する

遺伝子予想ソフト

MEGANTE : <https://megante.dna.affrc.go.jp/>

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

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

[連鎖解析](#)

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](#): 何でも検索
4. [Google Scholar](#): 科学に関する検索
5. [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. [Splign](#) : DNAとcDNAのマッチング
2. [Blast align](#) : DNAとcDNAのマッチング
3. [EMBOSS program](#) : DNAとcDNAのマッチング
4. [EMBOSS program](#) : DNA配列をアミノ酸配列に変換
5. [BLAST\(NCBI\)](#) : 遺伝子のホモロジー検索
6. [BLAST\(DDBJ\)](#) : 遺伝子のホモロジー検索
7. [MEGANTE](#) : 遺伝子の予測
8. [Softberry/FGENESH](#) : 遺伝子の予測 
9. [GeneMark](#) : 遺伝子の予測
10. [GENESCAN](#) : 遺伝子の予測
11. [clustalW](#) : アライメント
12. [InterProScan](#) : ドメイン解析
13. [PSORT](#) : 細胞内局在性予測
14. [GEO](#) : 遺伝子発現解析
15. [Gene Ontology](#) : 用語を用いた検索
16. [SSPN](#) : スプライシングの予測
17. [ORF Finder](#) : ORFの予測
18. [TFSEARCH](#) : 転写因子結合部位の予測
19. [MOTIF](#) : タンパク質モチーフ検索
20. [pI/Mw](#) : 等電点、分子量の予測
21. [RAP-DB](#) : イネのゲノムデータベース

TEST ON LINE

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

FGENESH

Reference: Solovyev V, Kosarev P, Seledsov I, Vorobyev D. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol.* 2006,7, Suppl 1: P. 10.1-10.12.

HMM-based gene structure prediction (multiple genes, both chains). The Fgenesh gene-finder was selected as the most accurate program for plant gene identification. *Plant Molecular Biology* (2005), 57, 3, 445-460: "Five ab initio programs (FGENESH, GeneMark.hmm, GENSCAN, GlimmerR and Grail) were evaluated for their accuracy in predicting maize genes. FGENESH yielded the most accurate and GeneMark.hmm the second most accurate predictions" (FGENESH identified 11% more correct gene models than GeneMark on a set of 1353 test genes).

Paste nucleotide sequence here:

```
agagggtgaatagcagccagatgaaggtttacaattgaacatcgaacgaaatccacacgaa  
atcataaattcatcgcatctgctgaagcttaccagattctatcgattcgatctagctccaatagtc  
gatggg
```

ある遺伝子YのGenomicDNA

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

Local file name:

Select organism specific gene-finding parameters :

[Total 353 genome-specific parameters are available for genefinders of FGENESH suite](#)

Oryza sativa Indica Group(long-grained rice)

[\[Help\]](#) [\[Show advanced options\]](#)

[\[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 created for our academic customers, including Broad Institute/MIT, Washington University, University of Minnesota and The Institute for Genomic Research (TIGR).

Your use of Softberry programs signifies that you accept [Terms of Use](#)

Last modification date: 12 Dec 2013

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

配列解析：

1. [Splign](#) : DNAとcDNAのマッチング
2. [Blast align](#) : DNAとcDNAのマッチング
3. [EMBOSS program"](#) : DNAとcDNAのマッチング
4. [EMBOSS program"](#) : DNA配列をアミノ酸配列に変換
5. [BLAST\(NCBI\)](#) : 遺伝子のホモロジー検索
6. [BLAST\(DDBJ\)](#) : 遺伝子のホモロジー検索
7. [Softberry/FGENESH](#) : 遺伝子の予測
8. [GeneMark](#) : 遺伝子の予測
9. [GENESCAN](#) : 遺伝子の予測
10. [clustalW](#) : アライメント
11. [InterProScan](#) : ドメイン解析
12. [PSORT](#) : 細胞内局在性予測
13. [GEO](#) : 遺伝子発現解析
14. [Gene Ontology](#) : 用語を用いた検索
15. [SSPN](#) : スプライシングの予測
16. [ORF Finder](#) : ORFの予測
17. [TFSEARCH](#) : 転写因子結合部位の予測
18. [MOTIF](#) : タンパク質モチーフ検索
19. [pI/Mw](#) : 等電点、分子量の予測
20. [RAP-DB](#) : イネのゲノムデータベース

解析用配列

1. [ある遺伝子AのgenomeDNA配列](#)
2. [ある遺伝子AのcDNA配列](#)
3. [ある遺伝子XのgenomeDNA配列](#)
4. [ある遺伝子XのcDNA配列](#)
5. [ある遺伝子YのgenomicDNA](#) 
6. [ある遺伝子YのcDNA](#)
7. [Z1遺伝子のgenomicDNA](#)
8. [Z1遺伝子のCDS](#)
9. [Z1遺伝子の予想アミノ酸配列](#)
10. [Z2遺伝子のgenomicDNA](#)
11. [Z2遺伝子のCDS](#)
12. [Z2遺伝子のアミノ酸配列](#)
13. [野生型SD1-genomicDNA配列](#)
14. [野生型SD1-cDNA配列](#)
15. [カルロス-cDNA](#)

TEST ON LINE

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

FGENESH

Reference: Solovyev V, Kosarev P, Seledsov I, Vorobyev D. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol.* 2006,7, Suppl 1: P. 10.1-10.12.

HMM-based gene structure prediction (multiple genes, both chains). The Fgenesh gene-finder was selected as the most accurate program for plant gene identification. *Plant Molecular Biology* (2005), 57, 3, 445-460: "Five ab initio programs (FGENESH, GeneMark.hmm, GENSCAN, GlimmerR and Grail) were evaluated for their accuracy in predicting maize genes. FGENESH yielded the most accurate and GeneMark.hmm the second most accurate predictions" (FGENESH identified 11% more correct gene models than GeneMark on a set of 1353 test genes).

Paste nucleotide sequence here:

```
agagggtgaatagcagccagatgaagggtttacaattgaacatcgaacgaaatccacacgaa  
atcataaattcatcgcatctgctgaagcttaccagattctatcgattcgatctagctccaatagtc  
gatggg
```

ある遺伝子YのGenomicDNA

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

Local file name:

Select organism specific gene-finding parameters :

[Total 353 genome-specific parameters are available for genefinders of FGENESH suite](#)

Oryza sativa Indica Group(long-grained rice)

[\[Help\]](#) [\[Show advanced options\]](#)

[\[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 created for our academic customers, including Broad Institute/MIT, Washington University, University of Minnesota and The Institute for Genomic Research (TIGR).

Your use of Softberry programs signifies that you accept [Terms of Use](#)

Last modification date: 12 Dec 2013

[Show picture of predicted genes in PDF file](#)

FGENESH 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):

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

```

ATGGCCAGACCCGGGAAAGGTGATCAAGTGCAAAGCCGGCGGTGGCATGGGAGGCCGGGAAG
CCCGTGTGCGATCGAGGAGGTGGAGGTTGCCCGCCCGCAGGCCATGGAGGTCCGCGTCAAG
ATCCTCTACACCCGCTCTGCCACACCCGACCGTCTACTTCTGGGAGGCCAAGGGGCAAAACA
CCTGTTTTCCCTAGGATCTTGGGCCATGAAGCTGGAGGCATTGTGGAGAGTGTGGGAGAG
GGTGTGACCGAACTCGCGCCGGCGGACCATGTCTCCCGGTGTTACCAGGCGAGTGCAAG
GAGTGTGATCACTGCAAAATCCGAGGAGAGCAACATGTGTGACCTCCTCAGGATCAACGTC
GACCCGCGCGTCAATGATGATCGCGGACGGCAAGTCCCGATTCAACCATCAAGGGGAAGCCCATC
TTCCACTTTGTTGGCACATCCACCTTCAGCGAGTACACCGTCATCCATGTCGGGTGCCTC
GCGAAGATCAACCCGGAGGCGCTCTCGACAAGGTCTGCATTCTCAGCTGCGGTTTCTCT
ACCGGTTTTGGTGCAACAGTGAATGTCCGGAACCGAAAAAGGGCCAGACCGTGGCTATT
TTCGGTCTTGGAGCTGTTGGTCTTGCCTATGGAAGGTGCCAGGCTGTCTGGAGCATCG
AGGATCATTGGTGTGGACCTGAACCCCTGCGAAATTCGAACAAGCTAAGAAGTTTGGCTGC
ACTGACTTCGTAACCCCAAGGACCACAGCAAGCCAGTGCACGAGGTGCTGATTGAGATG
ACCAACGGTGGACTCGACAGGCGCGTGAATGCACGGGCAACATCAACGCCATGATATCC
TGCTTCGAATGTGTCATGATGGATGGGGCGTTGGCGTGGTGGCGGTGCCGACCAAG
GATGATGTGTTCAAGACCCACCCGATGAACCTTCTCAACGAGAAGACGCTCAAGGGGACG
TTCTTCGGTAACCTACAAGCCGCGCACCGACCTGCCAACGTCGTCGAGCTGTACATGAAG
AAGGAGCTGGAGCTGGAGAAGTTCATCACGCACAGCGTGCCTTCTCGGAGATCAACACG
CGCTTCGATCTCATGCTCAAGGGGGAGAGTCTCCGCTGCGTCAATGAGGATGGATGAGTAG

```

>FGENESH: 1 10 exon (s) 337 - 2966 379 aa, chain +

```

MATAGKVIKCKAAAVANEAGKPLSIEEVEVAPPQAMEVVRVKILYTALCHTDVYFWEAKGQT
PVFPRILGHEAGGIVESVGEVTELPAGDHVLPVFTGCKEKCDCHEESNMCDLLRINV
DRGVMIGDGKSRFTIKGKPIHFVGTSTFSEYTVIHVGLAKINPEAPLDDKVCILSCGFS
TGFGATVNVAKPKKQTVVAFGLGAVGLAAMEGARLSGASRIIGVDLNPAPKFEQAKKFK
TDFVNPKDHSPVHEVLIEMTNGGLDRAVECTGNINAMISCFECVHDGWSVAVLVGVPTK
DDVFKTHPMNFLNEKTLKGTFFPNYKPRDLPNVVELYMKKELELEKFIHSHVFPSEINT
AFDMLKGESLRCVMRME

```

Oryza sativaのモデルで解析

予想されたATG-TAG(CDS)

予想されたアミノ酸

Wordに保存

Organism; *Oryza sativa*

MATAGKVIKCKAAVAWEAGKPLSIEEVEVAPPQAMEVRVKILYALCHTDVYFWEAKGQT
PVFPRILGHEAGGIVESVGEVTELAGDHVLPVFTGECKECDHCKSEESNMCDLLRINV
DRGVMIGDGKSRFTIKGKPIFHFVGTSTFSEYTVIHVGCLAKINPEAPLDKVCILSCGFS
TGFGATVNVAKPKKGQTVAFGLGAVGLAAMEGARLSGASRIIGVDLNPAAKFEQAKKFGC
TDFVNPDKHSPVHEVLIEMTNGGLDRAVECTGNINAMISCFECVHDGWGVAVLVGVPTK
DDVFKTHPMNFLNEKTLKGTFFGNYPRTDLPNVVELYMKKELELEKFI THSVPFSEINT
AFDLMLKGESLRCVMMRDE

TEST ON LINE

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

FGENESH

Reference: Solovyev V, Kosarev P, Seledsov I, Vorobyev D. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol.* 2006,7, Suppl 1: P. 10.1-10.12.

HMM-based gene structure prediction (multiple genes, both chains). The Fgenesh gene-finder was selected as the most accurate program for plant gene identification. *Plant Molecular Biology* (2005), 57, 3, 445-460: "Five ab initio programs (FGENESH, GeneMark.hmm, GENSCAN, GlimmerR and Grail) were evaluated for their accuracy in predicting maize genes. FGENESH yielded the most accurate and GeneMark.hmm the second most accurate predictions" (FGENESH identified 11% more correct gene models than GeneMark on a set of 1353 test genes).

Paste nucleotide sequence here:

```
tgagggtgaatagcagccagtatgaagggtttacaattgaacatcgaaacgaaatccacacgaa  
ttcataaattcatcgcatctgctgaagcctaccagattctatcgatctagctccaataggtc  
atggg
```

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

Local file name: ファイル未選択

Select organism specific gene-finding parameters :

Total 353 genome-specific parameters are available for genefinders of FGENESH suite

Human (Homo sapiens)

[[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 created for our academic customers, including Broad Institute/MIT, Washington University, University of Minnesota and The Institute for Genomic Research (TIGR).

Your use of Softberry programs signifies that you accept [Terms of Use](#)

Last modification date: 12 Dec 2013

ある遺伝子YのGenomicDNA

[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
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 CDS1     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 +

```
ATGGAGGTCGCCGTC AAGATCCTCTACACCGCCCTCTGCCACACCGACGCTACTTCTGG
GAGGCCAAGGGGCAAACACTGTTTCCCTAGGATCTTGGGCCATGAAGCTGGAGGCATT
GTGGAGAGTGTGGGAGAGGGGTGTGACCGAAGTCCGGCCGGGGCACCATGTCTCCCGGTG
TTCACCGCGAGTGC AAGGAGTGTGATCACTGCAAAATCGGAGGAGCAACATGTGTGAC
CTCCTCAGGATCAACGTCGACCGCGCGTCA TGATCGGCGACGGCAAGTCCCGATTACC
ATCAAGGGGAAGCCATCTTCCACTTTGTTGGCACATCCACCTTCAGCGAGTACACCGTC
ATCCATGTCCGGCTGCCTCGCGAAGATCAACCCGGAGGCGCCTCTCGACAAGGTCTGCATT
CTCAGCTGCCGTTTCTTACC GGTTTGGTGC AAGTGAATGTCCGGAACCGGAAAAAG
GGCCAGACCGTGGCTATTTTCGGTCTTGGAGCTGTTGGTCTTGGCTGCTATGGAAGGTGCC
AGGCTGTCTGGAGCATCGAGGATCATTTGGTGTGGACCTGAACCCCTGCGAAATTCGAACAA
GCTAAGAAGTTTGGCTGCACTGACTTCGTAACCCCAAGGACCACAGCAAGCCAGTGCAC
GAGGTGCTGATTGAGATGACCAACGGTGGACTCGACAGGGCCGTTGAATGCACGGGGCAAC
ATCAACGCCATGATATCCTGCTTCCGAATGTGCCATGATGGATGGGGCGTTGCCGTGCTG
GTCGGCGTGCCGACCAAGGATGATGTGTTCAAGACCCACCCGATGAACCTTCCCTCAACGAG
AAGACGCTCAAGGGGACGTTCTTCGGTA ACTACAAGCCGCGCACCGACCTGCCCAACGTC
GTCCAGCTGTACATGAAGAAGGAGCTGGAGCTGGAGAAGTTCATCACGCACAGCGTGCCG
TTCTCGGAGATCAACACGGCGTTCGATCTCATGCTCAAGGGGGAGAGTCTCCGCTGCGTC
ATGACCGATGGATGAGTAG
```

>FGENESH: 1 9 exon (s) 537 - 2966 345 aa, chain +

```
MEVRVKILY TALCHTDVYFWEAKGQTPVFPRI LGHEAGGIVESVGEVTE LAPGDHVL PV
FTGECKECDHCKSESNMCDLLRINVD RGVMI GDGKSRFTIKGKPI FHFVGTSTFSEYTV
IHVGC LAKINPEAPLDKVCILSCGFSTGFGATVNVAKPKKGQTVAI FGLGAVGLAAMEGA
RLSGASRIIGVDLNP AKFEQAKKFGCTDFVNP KDHSKPVEVLIEMTNGGLDRAVECTGN
INAMISCFECVHDG WGVAVLVGVP TKDDVFKTHPMNFLNEKTLKGTF FGNYPRTDLPNV
VELYMKKELELEKFI THSVFPFSEINTAFDLM LKGESLRCVMRME
```

Humanのモデルで解析

予想されたATG-TAG

予想されたアミノ酸

Modelによって予想されるORFは多少異なる

Organism; 植物 (Oryza sativa)

MATAGKVIKCKAAVAWEAGKPLSIEEVEVAPPQAMEVRVKILY TALCHTDVYFW EAKGQT
PVFPRILGHEAGGIVESVGEGVTELAPGDHVL PVFTGECKECDHCKSEESNMCDLLRINV
DRGVMIGDGKSRFTIKGKPIHFVGTSTFSEYTVIHVGCLAKINPEAPLDKVCILSCGFS
TGFGATVNVAKPKKGQTVAIFGLGAVGLAAMEGARLSGASRIIGVDL NPAKFEQAKKFGC
TDFVNP KDHSKPVHEVLIEMTNGGLDRAVECTGNINAMISCFECVHDGWGVAVLVGVPTK
DDVFKTHPMNFLNEKTLKGTFFGNYKPRTDLPNVVELYMKKELELEK FITHSVPFSEINT
AFDLMLKGESLRCVMRMDE

Organism; 人間

MEVRVKILY TALCHTDVYFW EAKGQTPVFPRILGHEAGGIVESVGEGVTELAPGDHVL PV
FTGECKECDHCKSEESNMCDLLRINVDRGVMIGDGKSRFTIKGKPIHFVGTSTFSEYTV
IHVGCLAKINPEAPLDKVCILSCGFSTGFGATVNVAKPKKGQTVAIFGLGAVGLAAMEGA
RLSGASRIIGVDL NPAKFEQAKKFGCTDFVNP KDHSKPVHEVLIEMTNGGLDRAVECTGN
INAMISCFECVHDGWGVAVLVGVPTKDDVFKTHPMNFLNEKTLKGTFFGNYKPRTDLPNV
VELYMKKELELEK FITHSVPFSEINTAFDLMLKGESLRCVMRMDE

Organism; ショウジョウバエ

MFCGEEKLIRSGFGVF AA VAWEAGKPLSIEEVEVAPPQAMEVRVKILY TALCHTDVYFW
EAKGQTPVFPRILGHEAGGIVESVGEGVTELAPGDHVL PVFTGECKECDHCKSEESNMCD
LLRINVDRGVMIGDGKSRFTIKGKPIHFVGTSTFSEYTVIHVGCLAKINPEAPLDKVC I
LSCGFSTGFGATVNVAKPKKGQTVAIFGLGAVGLAAMEGARLSGASRIIGVDL NPAKFEQ
AKKFGCTDFVNP KDHSKPVHEVLIEMTNGGLDRAVECTGNINAMISCFECVHDGWGVAVL
VGVPTKDDVFKTHPMNFLNEKTLKGTFFGNYKPRTDLPNVVELYMKKELELEK FITHSVP
FSEINTAFDLMLKGESLRCVMRMDE

生物情報工学II (担当 芦苺) 遺伝子の配列解析

第6回目-1 (11/1) [遺伝子の予測と機能解析](#)

-2 (11/1) [演習](#)

第7回目-1 (11/8) [DNA配列から遺伝子を探す](#)

-2 (11/8) [DNA、アミノ酸のアラインメント、系統樹解析](#)

-3 (11/8) [演習](#)

第8回目-1 (11/15) [Genetyxを用いた解析](#)

-2 (11/15) [演習](#)

第9回目-1 (11/22) [KEGGを用いた解析](#)

[連鎖解析](#)

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](#): 何でも検索
4. [Google Scholar](#): 科学に関する検索
5. [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. [Splign](#): DNAとcDNAのマッチング
2. [Blast align](#): DNAとcDNAのマッチング
3. [EMBOSS program](#): DNAとcDNAのマッチング
4. [EMBOSS program](#): DNA配列をアミノ酸配列に変換
5. [BLAST\(NCBI\)](#): 遺伝子のホモロジー検索 
6. [BLAST\(DDBJ\)](#): 遺伝子のホモロジー検索
7. [Softberry/FGENESH](#): 遺伝子の予測
8. [GeneMark](#): 遺伝子の予測
9. [GENESCAN](#): 遺伝子の予測
10. [clustalW](#): アライメント
11. [InterProScan](#): ドメイン解析
12. [PSORT](#): 細胞内局在性予測
13. [GEO](#): 遺伝子発現解析
14. [Gene Ontology](#): 用語を用いた検索
15. [SSPN](#): スプライシングの予測

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

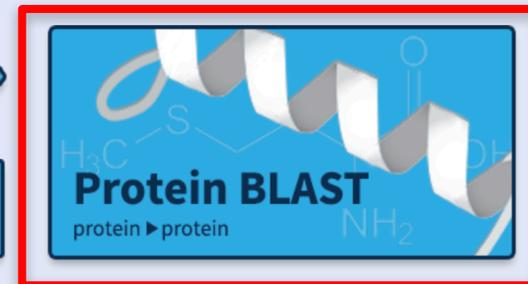
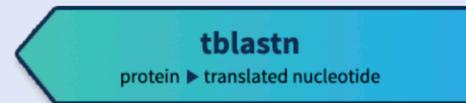
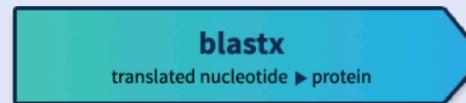
QuickBLASTP

Try **QuickBLASTP** for a fast protein search of nr.

Tue, 23 May 2017 13:00:00 EST

[More BLAST news...](#)

Web BLAST



BLAST Genomes

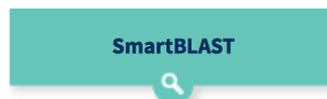
Enter organism common name, scientific name, or tax id.

[Human](#) [Mouse](#) [Rat](#) [Microbes](#)

Standalone and API BLAST



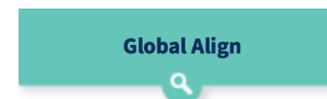
Specialized searches



Find proteins highly similar to your query



Design primers specific to your PCR template



Compare two sequences across their entire span (Needleman, Muzny)



Find conserved domains in your sequence

Standard Protein BLAST

blastn blastp blastx tblastn tblastx

BLASTP programs search protein databases using a protein query. [more...](#)

[Reset page](#) [Bookmark](#)

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#)

IHVGCLAKINPEAPLDKVCILSCGFSTGFGATVNVAKPKKGQTVAIFGLGAVGLAAMEGA
RLSGASRIIGVDLNPAAKFEQAKKFGCTDFVNPDKDHSKPVHEVLIEMTNGGLDRAVECTGN
INAMISCFECVHDGWGVAVLVGVPTKDDVFKTHPMNFLNEKTLKGTFFGNYKPRDLPNV
VELYMKKELELEKFITHSVPFSEINTAFDMLKGESLRCVVMRME

Query subrange [Clear](#)

From

To

Or, upload file

[ファイルを選択](#) [ファイル未選択](#) [Clear](#)

Job Title

Enter a descriptive title for your BLAST search [Clear](#)

Align two or more sequences [Clear](#)

Choose Search Set

Database

Non-redundant protein sequences (nr) [Clear](#)

Organism
Optional

Homo sapiens (taxid:9606) Exclude [+](#)

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown. [Clear](#)

Exclude
Optional

Models (XM/XP) Uncultured/environmental sample sequences

Entrez Query
Optional

[YouTube](#) [Create custom database](#)

Enter an Entrez query to limit search [Clear](#)

Program Selection

Algorithm

- Quick BLASTP (Accelerated protein-protein BLAST) **New**
- blastp (protein-protein BLAST)
- PSI-BLAST (Position-Specific Iterated BLAST)
- PHI-BLAST (Pattern Hit Initiated BLAST)
- DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)

Choose a BLAST algorithm [Clear](#)

BLAST

Search database **Non-redundant protein sequences (nr)** using **Blastp (protein-protein BLAST)**

Show results in a new window

[+ Algorithm parameters](#)

BLAST is a registered trademark of the National Library of Medicine

[Support center](#) [Mailing list](#) [YouTube](#)

NCBI

National Center for Biotechnology Information, U.S. National Library of Medicine
8600 Rockville Pike, Bethesda MD, 20894 USA

[Policies and Guidelines](#) | [Contact](#)



BLAST Results

Your search is limited to records that include: Homo sapiens (taxid:9606) » Full Entrez Query

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

[YouTube](#) [How to read this page](#) [Blast report description](#)

Job title: Y-Human

RID YB2VJZ32014 (Expires on 10-18 12:47 pm)

Query ID Icl|Query_39248
 Description Y-Human
 Molecule type amino acid
 Query Length 345

Database Name nr
 Description All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples from WGS projects
 Program BLASTP 2.7.0+ » Citation

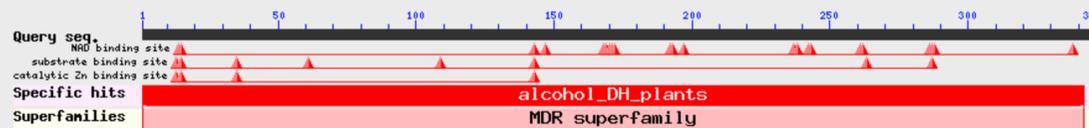
Other reports: [Search Summary](#) [Taxonomy reports](#) [Distance tree of results](#) [Multiple alignment](#) [MSA viewer](#)

New Analyze your query with [SmartBLAST](#)

Graphic Summary

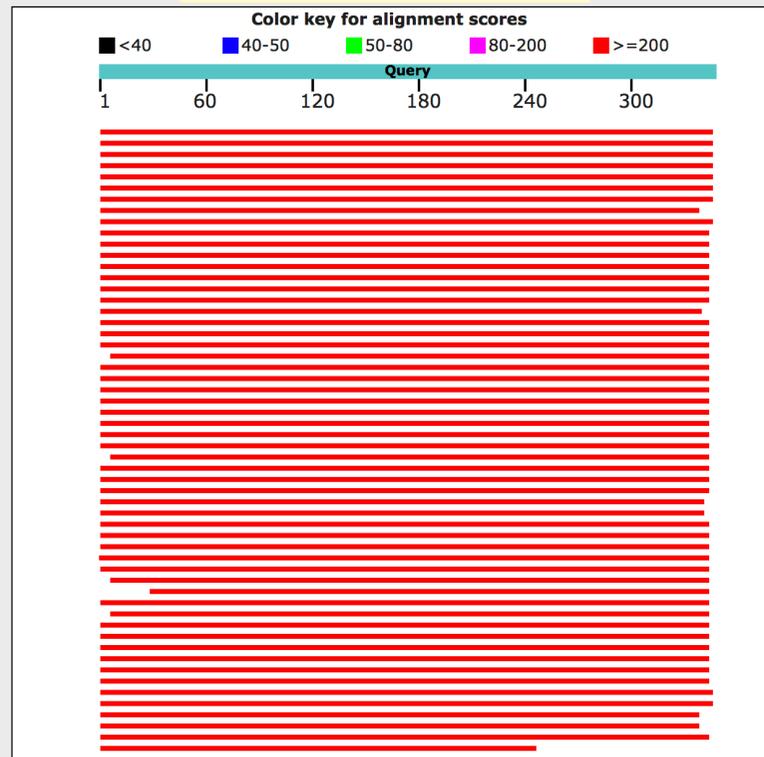
Show Conserved Domains

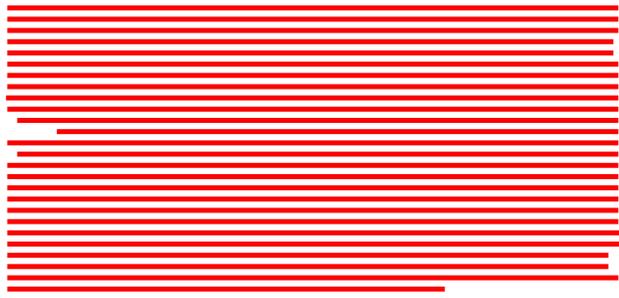
Putative conserved domains have been detected, click on the image below for detailed results.



Distribution of the top 73 Blast Hits on 72 subject sequences

Mouse over to see the title, click to show alignments





Descriptions

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> Chain A, S-nitrosoglutathione Reductase (gsnor) In Complex With N6022	353	353	99%	5e-120	51%	3QJ5_A
<input type="checkbox"/> Chain A, Structure Of Human Liver Chichi Alcohol Dehydrogenase (A Glutathione- Dependent Formaldehyde Dehydrogenase)	353	353	99%	5e-120	51%	1TEH_A
<input type="checkbox"/> alcohol dehydrogenase class-3 [Homo sapiens]	353	353	99%	6e-120	51%	NP_000662.3
<input type="checkbox"/> alcohol dehydrogenase 5 (class III), chi polypeptide [Homo sapiens]	352	352	99%	8e-120	51%	AAV38636.1
<input type="checkbox"/> ADH5 [Homo sapiens]	350	350	99%	5e-119	51%	CAG38730.1
<input type="checkbox"/> Chain A, Structure Of The Binary Complex Of The E67I Mutant Of Human Glutathione-Dependent Formaldehyde Dehydrogenase With Nad(H)	350	350	99%	8e-119	51%	2FZW_A
<input type="checkbox"/> alcohol dehydrogenase class III [Homo sapiens]	349	349	99%	3e-118	51%	AAA51597.1
<input type="checkbox"/> alcohol dehydrogenase 5 (class III), chi polypeptide, isoform CRA_b [Homo sapiens]	345	345	96%	1e-116	52%	EAX06085.1
<input type="checkbox"/> hCG17981 [Homo sapiens]	336	336	99%	2e-113	50%	EAW88504.1
<input type="checkbox"/> Chain A, Three-Dimensional Structures Of Three Human Alcohol Dehydrogenase Variants: Correlations With Their Functional Differences	335	335	98%	1e-112	48%	1HDY_A
<input type="checkbox"/> alcohol dehydrogenase 1B isoform 1 [Homo sapiens]	335	335	98%	1e-112	48%	NP_000659.2
<input type="checkbox"/> Chain A, Three-Dimensional Structures Of Three Human Alcohol Dehydrogenase Variants: Correlations With Their Functional Differences	332	332	98%	1e-111	48%	1HDX_A
<input type="checkbox"/> RecName: Full=Alcohol dehydrogenase 1B; AltName: Full=Alcohol dehydrogenase subunit beta	332	332	98%	1e-111	48%	P00325.2
<input type="checkbox"/> Chain A, Three-Dimensional Structures Of Three Human Alcohol Dehydrogenase Variants: Correlations With Their Functional Differences	331	331	98%	3e-111	48%	1HDZ_A
<input type="checkbox"/> alcohol dehydrogenase [Homo sapiens]	330	330	98%	8e-111	48%	CAA33487.1
<input type="checkbox"/> alcohol dehydrogenase beta-1 subunit (EC 1.1.1.1) [Homo sapiens]	329	329	98%	1e-110	48%	AAA51592.1
<input type="checkbox"/> rcADH5 [Homo sapiens]	328	328	97%	2e-110	50%	ABB92441.1
<input type="checkbox"/> alcohol dehydrogenase beta subunit [Homo sapiens]	328	328	98%	4e-110	47%	AAA51884.1
<input type="checkbox"/> Chain A, Crystallization Of Human Beta3 Alcohol Dehydrogenase (10 MgML) IN 100 MM SODIUM PHOSPHATE (PH 7.5), 7.5 MM NAD+ And 1 Mm 4-Iodopyrazole	328	328	98%	4e-110	47%	1HTB_A
<input type="checkbox"/> alcohol dehydrogenase beta-3 subunit [Homo sapiens]	328	328	98%	5e-110	47%	AAB48003.1
<input type="checkbox"/> alcohol dehydrogenase 1B isoform 2 [Homo sapiens]	325	325	97%	1e-109	47%	NP_001273579.1
<input type="checkbox"/> Chain A, Human Gamma-2 Alcohol Dehydrogenase	325	325	98%	7e-109	47%	1HT0_A
<input type="checkbox"/> Chain A, Crystal Structure Of Human Class Ii Alcohol Dehydrogenase (Adh4) In Complex With Nad And Zn	325	325	98%	8e-109	45%	3COS_A
<input type="checkbox"/> Alcohol dehydrogenase 4 (class II), pi polypeptide [Homo sapiens]	325	325	98%	8e-109	45%	AAH22319.1
<input type="checkbox"/> unnamed protein product [Homo sapiens]	325	325	98%	9e-109	47%	CAA27842.1
<input type="checkbox"/> alcohol dehydrogenase 4 (class II), pi polypeptide, isoform CRA_a [Homo sapiens]	324	324	98%	1e-108	45%	EAX06088.1
<input type="checkbox"/> alcohol dehydrogenase [Homo sapiens]	324	324	98%	1e-108	45%	CAA39813.1
<input type="checkbox"/> alcohol dehydrogenase 1C [Homo sapiens]	324	324	98%	1e-108	47%	NP_000660.1
<input type="checkbox"/> alcohol dehydrogenase 4 isoform 2 [Homo sapiens]	324	324	98%	1e-108	45%	NP_000661.2
<input type="checkbox"/> alcohol dehydrogenase IB (class I), beta polypeptide, isoform CRA_b [Homo sapiens]	323	323	97%	1e-108	47%	EAX06097.1

DNA配列の中から遺伝子を予測する

遺伝子予想ソフト

MEGANTE : <https://megante.dna.affrc.go.jp/>

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

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

[連鎖解析](#)

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](#): 何でも検索
4. [Google Scholar](#): 科学に関する検索
5. [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. [Splign](#) : DNAとcDNAのマッチング
2. [Blast align](#) : DNAとcDNAのマッチング
3. [EMBOSS program](#) : DNAとcDNAのマッチング
4. [EMBOSS program](#) : DNA配列をアミノ酸配列に変換
5. [BLAST\(NCBI\)](#) : 遺伝子のホモロジー検索
6. [BLAST\(DDBJ\)](#) : 遺伝子のホモロジー検索
7. [MEGANTE](#) : 遺伝子の予測
8. [Softberry/FGENESH](#) : 遺伝子の予測
9. [GeneMark](#) : 遺伝子の予測
10. [GENESCAN](#) : 遺伝子の予測 
11. [clustalW](#) : アライメント
12. [InterProScan](#) : ドメイン解析
13. [PSORT](#) : 細胞内局在性予測
14. [GEO](#) : 遺伝子発現解析
15. [Gene Ontology](#) : 用語を用いた検索
16. [SSPN](#) : スプライシングの予測
17. [ORF Finder](#) : ORFの予測
18. [TFSEARCH](#) : 転写因子結合部位の予測
19. [MOTIF](#) : タンパク質モチーフ検索
20. [pI/Mw](#) : 等電点、分子量の予測
21. [RAP-DB](#) : イネのゲノムデータベース

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):

The GENSCAN Web Server at MIT

Identification of complete gene structures in genomic DNA



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

Server update, November, 2009: We've been recently upgrading the GENSCAN webserver hardware, which resulted in some problems in the output of GENSCAN. We apologize for the inconvenience. These output errors were resolved.

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 instructions at the bottom of this page).

Organism: Sub exon cutoff (optional):

Sequence name (optional):

Print options:

Upload your DNA sequence file (upper or lower case, spaces/numbers ignored):

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

```
atcaacgccatgatatccttgccttgcgaatgtgtccatgatgtcagtccttcttacaacatgaaacatgatagtaacaaattgaagat
gatcatttcagcaatgcctccagttctgctgtaaacctgatcctgaaaactaagagtggaacatgaaacctgaatccagggatggg
gcgttgcggtgctggtcggcgtgccgaccaaggatgatgttcaagaccaccgatgaactcctcaacgagaagacgctca
aggggacgttctcggtaactcaagccgcgaccgacctgcccaacgtcgtcagctgacatgaagaaggtaaaataattgc
agaacagaatcagatttgaacctgaatgtgcattgttcagaatctatgtcagtgattccaatgagccactgattgatactattt
atgctaattgattctgggtttgatgctgaaggagctggagctggagaagttcatcacgcacagcgtgccgttctcggagatcaac
acggcgttcgatctcatgctcaaggggagagctccgctgctcatgaggatggatgagtagagatgatgatcctctctc
tctctctcttgagttcagtgcatacaagaacaaaaatgaaagcaaaaaatggatgagaagtttgaggtttccatttaagttt
aataatggtcgtgagacagggctgtgtggagtcagctctgttactcctccttcatgcttattgttattgtagcattgtggc
ttagcagatttcaccagtttaccctgtaatttagtaccattctgaaatcgtaataaacactagcagtgtaggtagtactatgtcggaa
ctgaaattgattgtgcagtaaaagaggtgaatagcagccagatgaaagttttacaattgaaacatcgaaacgaaatccacagaaatt
cataaatttcacgcatcttgctgaagcttaccagattctatcgattcgatctagctccaataggtcgatggg
```

ある遺伝子YのGenomicDNA

[Back to the top](#)

Predicted coding sequence(s):

>/tmp/10_17_17-04:13:55.fasta|GENSCAN_predicted_peptide_1|379_aa

```
MATAGKVIKCKAAVAWEAGKPLSIEEVEVAPPQAMEVRVKILYALCHTDVYFWEAKGQT  
PVFPRILGHEAGGIVESVGEVTE LAPGDHVL PVFTGECKECDHCKSESNMCDLLRINV  
DRGVMIGDGKSRFTIKGKPIFHFVGTSTFSEYTVIHVGLAKINPEAPLDKVCILSCGFS  
TGFGATVNVAKPKKQTVAIFGLGAVGLAAMEGARLSGASRIIGVDLNPAPKEQAKKFGC  
TDFVNPKDHSKPVHEVLIEMTNGGLDRAVECTGNINAMISCFECVHDGWGVAVLVGVPTK  
DDVFKTHPMNFLNEKTLKGTFFGNYKPRDLPNVVELYMKKELELEKFI THSVPFSEINT  
AFDLMLKGESLRCVVRMDE
```

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する。

>/tmp/10_17_17-04:13:55.fasta|GENSCAN_predicted_CDS_1|1140_bp

```
atggcgacagccgggaaggtgatcaagtgc aaagcggcggatgggagccgggaag  
ccgctgtcgatcgaggaggtggaggttgcgccgccgagccatggaggtccgcgtcaag  
atcctctacaccgcctctgcccacaccgacgtctacttctgggagccaagggcaaca  
cctgtttccctagatcttggccatgaagctggaggcattgtggagagtgtggagag  
ggtgtgaccgaactcgcgccggcgaccatgtcctcccgggtgtcaccggcgagtcaag  
gagtgatgactgcaaatcggaggagagcaacatgtgtgacctcctcaggatcaacgtc  
gaccgcgcgctcatgatcggcgacggcaagtcccgattcaccatcaaggggaagccatc  
ttccactttgttggcacatccacctcagcgagtacaccgtcatccatgtcggctgctc  
gcgaagatcaaccggaggcctctcgacaaggtctgcattctcagctgoggtttctct  
accggttttggtgcaacagtgaatgtcgcgaaccgaaaaggccagaccgtggctatt  
ttcggcttggagctgttggcttctgctgctatggaaggtgccaggctgtctggagcatcg  
aggatcattggtgtggacctgaacctgcgaaattcgaacaagctaagaagtttggtgc  
actgacttcgtaaacccaaggaccacagcaagccagtgacagaggtgctgattgagatg  
accaacggtggactcgacagggccgttgaatgcacgggcaacatcaacgccatgatatcc  
tgcttcgaatgtgtccatgatggatggggcgttgcgggtgctggcggcgtgccgaccaag  
gatgatgtgttcaagaccaccgatgaacttctcaacgagaagacgctcaaggggacg  
ttcttcggtaactacaagcccgccaccgacctgcccaacgtcgtcgagctgtacatgaag  
aaggagctggagctggagaagttcatcacgcacagcgtgccgttctcggagatcaacacg  
gcgttcgatctcatgctcaagggggagagctcctcgtcgtcatgaggatggatgagtag
```

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

生物情報工学II (担当 芦苺) 遺伝子の配列解析

第6回目-1 (11/1) [遺伝子の予測と機能解析](#)

-2 (11/1) [演習](#)

第7回目-1 (11/8) [DNA配列から遺伝子を探す](#)

-2 (11/8) [DNA、アミノ酸のアラインメント、系統樹解析](#)

-3 (11/8) [演習](#)

第8回目-1 (11/15) [Genetyxを用いた解析](#)

-2 (11/15) [演習](#)

第9回目-1 (11/22) [KEGGを用いた解析](#)

[連鎖解析](#)

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](#): 何でも検索
4. [Google Scholar](#): 科学に関する検索
5. [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. [Splign](#) : DNAとcDNAのマッチング
2. [Blast align](#) : DNAとcDNAのマッチング
3. [EMBOSS program](#) : DNAとcDNAのマッチング
4. [EMBOSS program](#) : DNA配列をアミノ酸配列に変換
5. [BLAST\(NCBI\)](#) : 遺伝子のホモロジー検索 
6. [BLAST\(DDBJ\)](#) : 遺伝子のホモロジー検索
7. [Softberry/FGENESH](#) : 遺伝子の予測
8. [GeneMark](#) : 遺伝子の予測
9. [GENESCAN](#) : 遺伝子の予測
10. [clustalW](#) : アライメント
11. [InterProScan](#) : ドメイン解析
12. [PSORT](#) : 細胞内局在性予測
13. [GEO](#) : 遺伝子発現解析
14. [Gene Ontology](#) : 用語を用いた検索
15. [SSPN](#) : スプライシングの予測

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

QuickBLASTP

Try **QuickBLASTP** for a fast protein search of nr.

Tue, 23 May 2017 13:00:00 EST

[More BLAST news...](#)

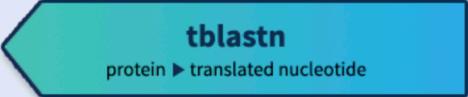
Web BLAST



Nucleotide BLAST
nucleotide ▶ nucleotide



blastx
translated nucleotide ▶ protein



tblastn
protein ▶ translated nucleotide



Protein BLAST
protein ▶ protein

BLAST Genomes

Enter organism common name, scientific name, or tax id.

[Human](#) [Mouse](#) [Rat](#) [Microbes](#)

Standalone and API BLAST



Download BLAST
Get BLAST databases and executables



Use BLAST API
Call BLAST from your application



Use BLAST in the cloud
Start an instance at a cloud provider

Specialized searches



SmartBLAST

Find proteins highly similar to your query



Primer-BLAST

Design primers specific to your PCR template



Global Align

Compare two sequences across their entire span (Needleman, Winkler)



CD-search

Find conserved domains in your sequence

Standard Nucleotide BLAST

blastn blastp blastx tblastn tblastx

BLASTN programs search nucleotide databases using a nucleotide query. [more...](#)[Reset page](#) [Bookmark](#)

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#)[Clear](#)Query subrange [?](#)From To

Or, upload file

ファイルを選択 | ファイル未選択 [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#) Align two or more sequences [?](#)

Choose Search Set

Database

 Human genomic + transcript Mouse genomic + transcript Others (nr etc.):Nucleotide collection (nr/nt) [?](#)

Organism

Optional

 Exclude [+](#)Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

Exclude

Optional

 Models (XM/XP) Uncultured/environmental sample sequences

Limit to

Optional

 Sequences from type material

Entrez Query

Optional

 [YouTube](#) [Create custom database](#)Enter an Entrez query to limit search [?](#)

Program Selection

Optimize for

 Highly similar sequences (megablast) More dissimilar sequences (discontiguous megablast) Somewhat similar sequences (blastn)Choose a BLAST algorithm [?](#)

BLAST

Search database Nucleotide collection (nr/nt) using Megablast (Optimize for highly similar sequences)

 Show results in a new window[+ Algorithm parameters](#)



Align Sequences Nucleotide BLAST

[blastn](#) [blastp](#) [blastx](#) [tblastn](#) [tblastx](#)

BLASTN programs search nucleotide subjects using a nucleotide query. [more...](#)

[Reset page](#) [Bookmark](#)

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#)

Query subrange [Clear](#)

From

To

Or, upload file

[Clear](#)

Job Title

Enter a descriptive title for your BLAST search [Clear](#)

Align two or more sequences [Clear](#)

Enter Subject Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#)

Subject subrange [Clear](#)

From

To

Or, upload file

[Clear](#)

Program Selection

Optimize for

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

Choose a BLAST algorithm [Clear](#)

BLAST

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

Show results in a new window

[+ Algorithm parameters](#)

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...

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)

Query ID Id 17067	Subject ID 17069
Description None	Description None
Molecule type nucleic acid	Molecule type nucleic acid
Query Length 1554	Subject Length 1554
	Program BLASTN 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

<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**

```
>lcl|17069
Length=1554
```


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   ATGTTGcgcgctgccgcccgttcggggccccgctgggcccgcgcctcttgcagccgcc 60
Sbjct 1   ATGTTGCGCGCTGCCGCCGCTTCGGGCCCGCCTGGGCCCGCCTCTTGTTCAGCCGCC 60

Query 61   gccacccaggccgTGCCCTGCCCCCAACCAGCAGCCCGAGGCTTCTGCAACCAGATTTTC 120
Sbjct 61   GCCACCCAGGCCGTGCTGCCCCCAACCAGCAGCCCGAGGCTTCTGCAACCAGATTTTC 120

Query 121  ATAACAATGAATGGCAGATGCCGTCAGCAGGAAAACATCCCCACCGTCAATCCGTCC 180
Sbjct 121  ATAACAATGAATGGCAGATGCCGTCAGCAGGAAAACATCCCCACCGTCAATCCGTCC 180

Query 181  ACTGGAGAGGTATCTGTCTCAGGTAGCTGAAGGGGACAAGGAAGATGTGGACAAGGCAGTG 240
Sbjct 181  ACTGGAGAGGTATCTGTCTCAGGTAGCTGAAGGGGACAAGGAAGATGTGGACAAGGCAGTG 240

Query 241  AAGGCCGCCCGGGCCGCTTCCAGCTGGGCTCACCTTGGCGCCGATGGACGCATCACAC 300
Sbjct 241  AAGGCCGCCCGGGCCGCTTCCAGCTGGGCTCACCTTGGCGCCGATGGACGCATCACAC 300

Query 301  AGGGCCGGCTGTGAACCGCCTGGCCGATCTGATCGAGCGGACCGGACCTACCTGGCG 360
Sbjct 301  AGGGCCGGCTGTGAACCGCCTGGCCGATCTGATCGAGCGGACCGGACCTACCTGGCG 360

Query 361  GCCTTGGAGACCTGGACAATGGCAAGCCCTATGTCTCTCTACCTGGTGGATTTGGAC 420
Sbjct 361  GCCTTGGAGACCTGGACAATGGCAAGCCCTATGTCTCTCTACCTGGTGGATTTGGAC 420

Query 421  ATGGTCTCAAATGTCTCCGGTATTATGCCGGCTGGGCTGATAAGTACCACGGGAAAACC 480
Sbjct 421  ATGGTCTCAAATGTCTCCGGTATTATGCCGGCTGGGCTGATAAGTACCACGGGAAAACC 480

Query 481  ATCCCCATTGACGGAGACTTCTTCAGCTACACGCCATGAACCTGTGGGGTGTGCGGG 540
Sbjct 481  ATCCCCATTGACGGAGACTTCTTCAGCTACACGCCATGAACCTGTGGGGTGTGCGGG 540

Query 541  CAGATCATCCGTGGAATTTCCCGCTCCTGATGCAAGCATGGAAGCTGGGCCAGCCTTG 600
Sbjct 541  CAGATCATCCGTGGAATTTCCCGCTCCTGATGCAAGCATGGAAGCTGGGCCAGCCTTG 600

Query 601  GCAACTGGAACGTGGTTGTGATGAAGGTAGCTGAGCAGACACCCCTCACCCGCTCTAT 660
Sbjct 601  GCAACTGGAACGTGGTTGTGATGAAGGTAGCTGAGCAGACACCCCTCACCCGCTCTAT 660

Query 661  GTGGCCAACCTGATCAAGGAGGCTGGCTTTCCCCCTGGTGTGGTCAACATTTGTCCTGGA 720
Sbjct 661  GTGGCCAACCTGATCAAGGAGGCTGGCTTTCCCCCTGGTGTGGTCAACATTTGTCCTGGA 720

Query 721  TTTGGCCCCACGGCTGGGGCCGCCATTGCCTCCCATGAGGATGTGGACAAGTGGCATT 780
```

NCBI Blast:Nucleotide Sequence (1554 letters)

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

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

NCBI Blast:Nucleotide Sequ...

Sbjct	541	CAGATCATTCCGTGGAATTTCCCGCTCCTGATGCAAGCATGGAAGCTGGGCCAGCCTTG	600
Query	601	GCAACTGGAAACGTGGTTGTGATGAAGGTAGCTGAGCAGACACCCCTCACCGCCCTCTAT	660
Sbjct	601	GCAACTGGAAACGTGGTTGTGATGAAGGTAGCTGAGCAGACACCCCTCACCGCCCTCTAT	660
Query	661	GTGGCCAACCTGATCAAGGAGGCTGGCTTTCCCCCTGGTGGTCAACATTGTGCCTGGA	720
Sbjct	661	GTGGCCAACCTGATCAAGGAGGCTGGCTTTCCCCCTGGTGGTCAACATTGTGCCTGGA	720
Query	721	TTTGGCCCCACGGCTGGGGCCGCATTGCCCTCCCATGAGGATGTGGACAAAGTGGCATT	780
Sbjct	721	TTTGGCCCCACGGCTGGGGCCGCATTGCCCTCCCATGAGGATGTGGACAAAGTGGCATT	780
Query	781	ACAGGCTCCACTGAGATTGGCCCGTAATCCAGGTTGCTGCTGGGAGCAGCAACCTCAAG	840
Sbjct	781	ACAGGCTCCACTGAGATTGGCCCGTAATCCAGGTTGCTGCTGGGAGCAGCAACCTCAAG	840
Query	841	AGAGTGACCTTGGAGTgggggggAAGAGCCCCAACATCATATGTCAGATGCCGATATG	900
Sbjct	841	AGAGTGACCTTGGAGTGGGGGGGAAGAGCCCCAACATCATATGTCAGATGCCGATATG	900
Query	901	GATTGGGCCGTGGAACAGGCCCACTTCGCCCTGTTCTTCAACCAGGGCCAGTGTGCTGT	960
Sbjct	901	GATTGGGCCGTGGAACAGGCCCACTTCGCCCTGTTCTTCAACCAGGGCCAGTGTGCTGT	960
Query	961	GCCGGCTCCCGACCTTCGTGCAGGAGGACATCTATGATGAGTTGTGGAGCGGAGCGTT	1020
Sbjct	961	GCCGGCTCCCGACCTTCGTGCAGGAGGACATCTATGATGAGTTGTGGAGCGGAGCGTT	1020
Query	1021	GCCCCGGCCAAGTCTCGGGTGGTCGGGAACCCCTTTGATAGCAAGACCAGCAGGGGCCG	1080
Sbjct	1021	GCCCCGGCCAAGTCTCGGGTGGTCGGGAACCCCTTTGATAGCAAGACCAGCAGGGGCCG	1080
Query	1081	CAGGTGGATGAAACTCAGTTTAAGAAGATCCTCGGCTACATCAACACGGGAAGCAAGAG	1140
Sbjct	1081	CAGGTGGATGAAACTCAGTTTAAGAAGATCCTCGGCTACATCAACACGGGAAGCAAGAG	1140
Query	1141	GGGGCGAAGCTGCTGTGTGGTGGGGCATTGCTGCTGACCGTGGTTACTTCATCCAGCCC	1200
Sbjct	1141	GGGGCGAAGCTGCTGTGTGGTGGGGCATTGCTGCTGACCGTGGTTACTTCATCCAGCCC	1200
Query	1201	ACTGTGTTTGGAGATGTGCAGGATGGCATGACCATCGCCAAGGAGGAGATCTTCGGGCCA	1260
Sbjct	1201	ACTGTGTTTGGAGATGTGCAGGATGGCATGACCATCGCCAAGGAGGAGATCTTCGGGCCA	1260
Query	1261	GTGATGCAGATCCTGAAGTTCAAGACCATAGAGGAGGTTGTTGGGAGAGCCAACAATTC	1320
Sbjct	1261	GTGATGCAGATCCTGAAGTTCAAGACCATAGAGGAGGTTGTTGGGAGAGCCAACAATTC	1320
Query	1321	ACGTACGGGCTGGCCCGAGCTGTCTTCACAAAGGATTTGGACAAGGCCAATTACCTGTCC	1380
Sbjct	1321	ACGTACGGGCTGGCCCGAGCTGTCTTCACAAAGGATTTGGACAAGGCCAATTACCTGTCC	1380
Query	1381	CAGGCCCTCCAGGGGGCACTGTGTGGTCAACTGCTATGATGTGTTTGGAGCCCAGTCA	1440
Sbjct	1381	CAGGCCCTCCAGGGGGCACTGTGTGGTCAACTGCTATGATGTGTTTGGAGCCCAGTCA	1440
Query	1441	CCCTTTGGTGGCTACAAGATGTCCGGGAGTGGCCGGGAGTTGGCCGAGTACGGGCTGCAG	1500
Sbjct	1441	CCCTTTGGTGGCTACAAGATGTCCGGGAGTGGCCGGGAGTTGGCCGAGTACGGGCTGCAG	1500
Query	1501	GCATACACTGAGTGAAGTGTCAAGTCAAGTGCCTCAGAAGAACTCATAA	1554
Sbjct	1501	GCATACACTGAGTGAAGTGTCAAGTCAAGTGCCTCAGAAGAACTCATAA	1554

Select All [Get selected sequences](#) NEW

Z1CDS

GCATACACTGAAGTGAAA**ACTGTCACAGTCAAAGTGCCTCAGAAGAACTCATAA**

Z2CDS

GCATACACTAAGTGAAA**ACTGTCACAGTCAAAGTGCCTCAGAAGAACTCATAA**

******* *******

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 英和辞書 - 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

MLRAAARFGPRLGRLLSAAATQAVPAPNQQPEVFCNQIFINNEWHDAVSRKTFPTVNPS
TGEVICQVAEGDKEDVDKAVKAARAAFQLGSPWRRMDASHRGRLNRLADLIERDRTYLA
ALETLDNGKPYVISYLVLDLDMVLKCLRYAGWADKYHGKTIPIGDFFSYTRHEPVGCG
QIIPWNFPLLMQAWKLGALATGNVVMKVAEQPLTALYVANLIKEAGFPFGVNIIVPG
FGPTAGAAIASHEDVDKVAFTGSTEIGRVIQVAAGSSNLKRVTLLEGGKSPNIIMSDADM

From 10 To 100 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

MLRAAARFGPRLGRLLSAAATQAVPAPNQQPEVFCNQIFINNEWHDAVSRKTFPTVNPS
TGEVICQVAEGDKEDVDKAVKAARAAFQLGSPWRRMDASHRGRLNRLADLIERDRTYLA
ALETLDNGKPYVISYLVLDLDMVLKCLRYAGWADKYHGKTIPIGDFFSYTRHEPVGCG
QIIPWNFPLLMQAWKLGALATGNVVMKVAEQPLTALYVANLIKEAGFPFGVNIIVPG
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)

Query ID	lc1 9111	Subject ID	9113
Description	None	Description	None
Molecule type	amino acid	Molecule type	amino acid
Query Length	517	Subject Length	517
		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

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

0 100 200 300 400 500

Dot Matrix View

Descriptions

Sequences producing significant alignments:	Score	E
	(Bits)	Value
lc1 9113 unnamed protein product	<u>1070</u>	0.0

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

```

>lc1|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      MLRAAAREGPRIGRRLLSAAATOAVPAPNOOPEVFCNOLEFNNEWHDAVSRKTEPTVNPS 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 英和辞書 - 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	Purple
>=200	Red

Query

▼ Dot Matrix View

Plot of lcl|9111 vs 9113

▼ Descriptions

Sequences producing significant alignments:	Score (Bits)	E Value
lcl 9113 unnamed protein product	<u>1070</u>	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	<u>1070</u>	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 TGEVICQVAEGDKEDVDKAVKAARAAFQLGSPWRRMDASHRGLLNRLADLIERDRTYLA 120
Sbjct 61 TGEVICQVAEGDKEDVDKAVKAARAAFQLGSPWRRMDASHRGLLNRLADLIERDRTYLA 120

Query 121 ALETLDNGKPYVISYLVLDLDMVLKCLRYAGWADKYHGKTIPIDGFFSYTRHEPVGVC 180
Sbjct 121 ALETLDNGKPYVISYLVLDLDMVLKCLRYAGWADKYHGKTIPIDGFFSYTRHEPVGVC 180

Query 181 QIIPWNFPLLMQAWKLGPALATGNVVMKVAEQTPLTALYVANLIKEAGFPVGVNIVPG 240
Sbjct 181 QIIPWNFPLLMQAWKLGPALATGNVVMKVAEQTPLTALYVANLIKEAGFPVGVNIVPG 240

Query 241 FGPTAGAAIASHEDVDKVAFTGSTEIGRVIQVAAGSSNLKRVTLLELGGKSPNIIMSDADM 300
Sbjct 241 FGPTAGAAIASHEDVDKVAFTGSTEIGRVIQVAAGSSNLKRVTLLELGGKSPNIIMSDADM 300

Query 301 DWAVEQAHFALFFNQGCCAGSRTFVQEDIYDEFVRSVARAKSRVVGPNPDSKTEQGP 360
Sbjct 301 DWAVEQAHFALFFNQGCCAGSRTFVQEDIYDEFVRSVARAKSRVVGPNPDSKTEQGP 360

Query 361 QVDETQFKKILGYINTGKQEGAKLLCGGGIAADRGYFIQPTVFGDVQDGMTIAKEEIFGP 420
Sbjct 361 QVDETQFKKILGYINTGKQEGAKLLCGGGIAADRGYFIQPTVFGDVQDGMTIAKEEIFGP 420

Query 421 VMQILKFKTIEEVVGRANNSTYGLAAAVFTKDLKANYLSQALQAGTVVWNCYDVFGAQS 480
Sbjct 421 VMQILKFKTIEEVVGRANNSTYGLAAAVFTKDLKANYLSQALQAGTVVWNCYDVFGAQS 480

Query 481 PFGGYKMSGSGRELGEYGLQAYTEKTVTVKVPQKNS 517
Sbjct 481 PFGGYKMSGSGRELGEYGLQAYTEKTVTVKVPQKNS 517

```

Z1CDS GCATACACT**GAA**GTGAAAAGTGTACAGTCAAAGTGCCTCAGAAGAACTCATAA
Z2CDS GCATACACT**AA**GTGAAAAGTGTACAGTCAAAGTGCCTCAGAAGAACTCATAA

グルタミン酸

Z1Aminoacid PFGGYKMSGSGRELGEYGLQAYT**E**VKTVTVPKVPQKNS
Z2Aminoacid PFGGYKMSGSGRELGEYGLQAYT**K**VKTVTVPKVPQKNS

リジン

Please specify input sequences by GI/Accession or in FASTA format.

Examples (click to select):

- [NM_214647 / NW_732498 \(one model\)](#)
- [AF238306 / NT_033777 \(one model, frameshifts\)](#)
- [NM_020978 / NG_004750 \(multiple models\)](#)

cDNA: Z2の予想CDS酸配列

```
>Z2 CDS
ATGTTGCGCGCTGCCGCCCGCTTCGGGCCCCGCCTGGGCC
GCCGCCTCTTGTGAGCCGCC
GCCACCCAGGCCGTGCCTGCCCCCAACCAGCAGCCCGAGG
TCTTCTGCAACCAGATTTTC
ATAACAATGAATGGCACGATGCCGTCAGCAGGAAAACATTC
```

- Lower quality query sequence (e.g. EST)
- Reverse and complement the query
- More partial alignments
- Use discontinuous megablast (e.g. for cross-species)

Align

Genomic: Z1のゲノミックDNA配列

```
>Z1g
TGCCAGGTGGTCTCATCTCCTGGCCTTTGCCCTTGCTGTTCC
CTGTCATCATTGAGGTCTCACTTGTGATTTCTGACCATGGTA
CTTATAAAAGCAGTGCCGTCTGCCCATCCATGTCACCTCGT
TCATCTCCTTACCTCCGAAATGATCTCGCTTTTGGGTTAC
GGCCGGTCTTTCACCTGGAGCATCAGCCGGGAAGGTCAGG
```

ファイルを選択 ファイル未選択

Whole genome:

Help: for questions, comments, or bug reporting, please visit [NCBI Support Center](#)

#	Query	Subject	Span(bp)	Coverage(%)	Overall(%)	Exon(%)	CDS(%)	In-frame(%)
1	Z2(+)	Z1g(+)	562-43154	100.00	99.94	99.94	0.00	0.00
2	Z2(+)	Z1g(+)	44562-87154	100.00	99.94	99.94	0.00	0.00

[Graphics](#) | [Text](#)

Model	Coverage	CDS	Mismatches and indels
Model 1	100.00% Overall 99.94% Exon 99.94%	0.00% In-frame 0.00% Primary transcript 1554 bp	1 Exons (min/max/ave), bp 33 / 185 / 119 Introns (min/max/ave), bp 435 / 14822 / 3420

Z2 (+) CDS

Z1g (+)

562 43154

Segments Alignment

Segments	Alignment
1 2 3 4 5 6	V N C Y D V F G A Q S P F G G Y K M S G S
7 8 9 10 11 12	1407 GGTCAACTGCTATGATGTTGTTGGAGCCAGTCACCCTTTGGTGGCTACAAGATGTCGGGGAGTG
13	37433 TACAGGGTCAACTGCTATGATGTTGTTGGAGCCAGTCACCCTTTGGTGGCTACAAGATGTCGGGGAGTG
	G R E L G E Y G L Q A Y T K V K T
1472	GCCGGGAGTTGGGCGAGTACGGGCTGCAGGCATACACTAAAGTGAAAAC T
37503	GCCGGGAGTTGGGCGAGTACGGGCTGCAGGCATACACTGAAGTGAAAAC TGTGAG

Model	Coverage	CDS	Mismatches and indels
Model 2	100.00% Overall 99.94% Exon 99.94%	0.00% In-frame 0.00% Primary transcript 1554 bp	1 Exons (min/max/ave), bp 33 / 185 / 119 Introns (min/max/ave), bp 435 / 14822 / 3420

Z2 (+) CDS

Z1g (+)

44562 87154

Segments Alignment

Segments	Alignment
1 2 3 4 5 6	M L R A A A R F G P R L G R R L L S A A A T Q
7 8 9 10 11 12	1 ATGTTGCGCGCTGCCGCCGCTTCGGGCCCGCCTGGGCCGCGCCTCTTGTGTCAGCCGCCGCCACCCAGG
13	44562 ATGTTGCGCGCTGCCGCCGCTTCGGGCCCGCCTGGGCCGCGCCTCTTGTGTCAGCCGCCGCCACCCAGG
	A V P A P N Q Q P E V F C N Q
71	CCGTGCCTGCCCCCAACCAAGCAGCCCGAGGTCTTCTGCAACCAAG
44632	CCGTGCCTGCCCCCAACCAAGCAGCCCGAGGTCTTCTGCAACCAAGGTGAG

第12exon

第13exon

Z1CDS

GCATACACT**GAA**GTGAAA**ACT**GTCACAGTCAAAGTGCCTCAGAAGAACTCATAA

Z2CDS

GCATACACT**AAAG**TGAAA**ACT**GTCACAGTCAAAGTGCCTCAGAAGAACTCATAA

第12exon

グルタミン酸

第13exon

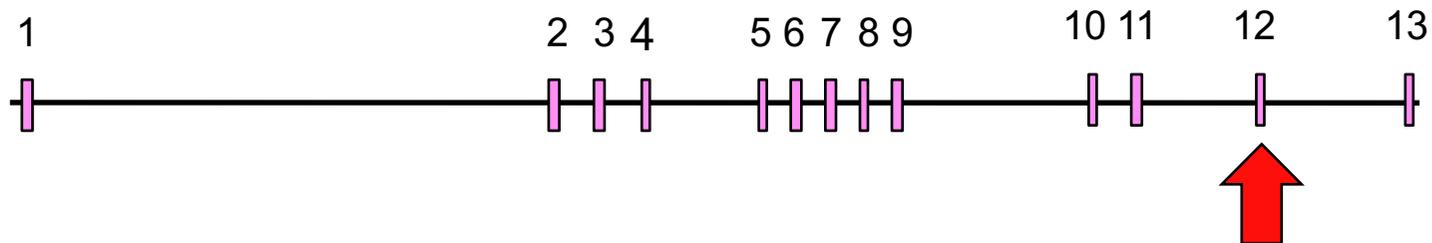
Z1Amionoacid

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

