Bioinformática: QBQ-5722

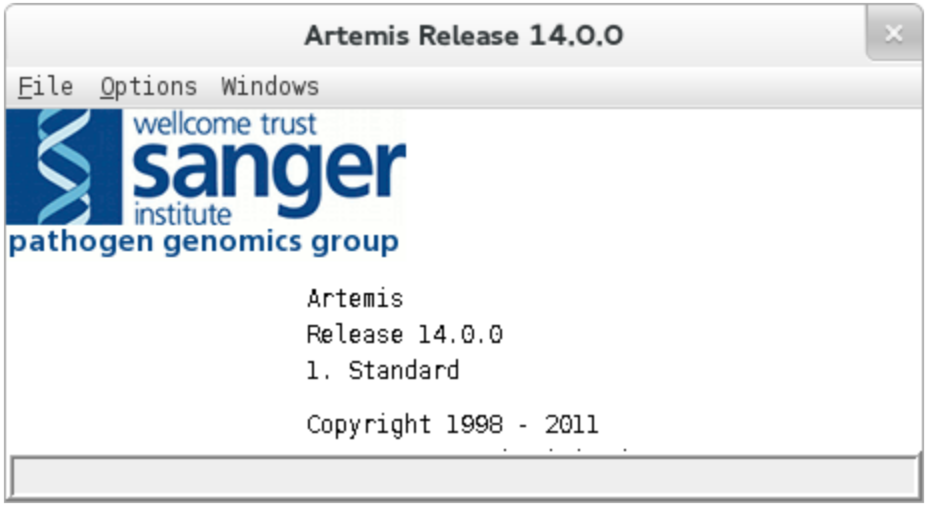
Anotação Artemis: Passo-a-passo

Prof. Dr. João Carlos Setubal

Anotação de genes no Artemis

Na área de trabalho, dentro da pasta **QBQ2507**, dê um duplo clique no ícone **Artemis**;

O Artemis será carregado. Clique no menu superior esquerdo `File` e selecione a opção `Open`;



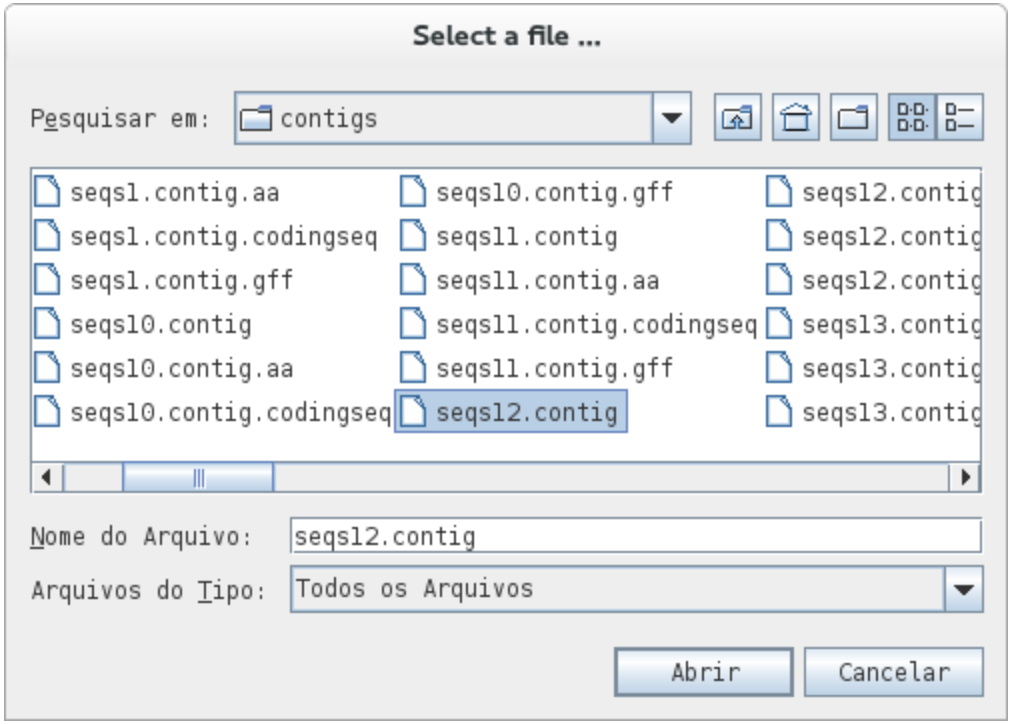
```
Artemis
Release 14.0.0
1. Standard

Copyright 1998 - 2011
```

Anotação de genes no Artemis

No diálogo que aparecerá, altere a caixa de seleção `Arquivos do tipo` de `Sequence files` para `Todos os arquivos`

Navegue até a pasta **QBQ2507** que está no **Desktop** e selecione o arquivo **seqs*.contig** presente. Clique em `Abrir`.

Anotação de genes no Artemis



Select a file ...

Pesquisar em: contigs

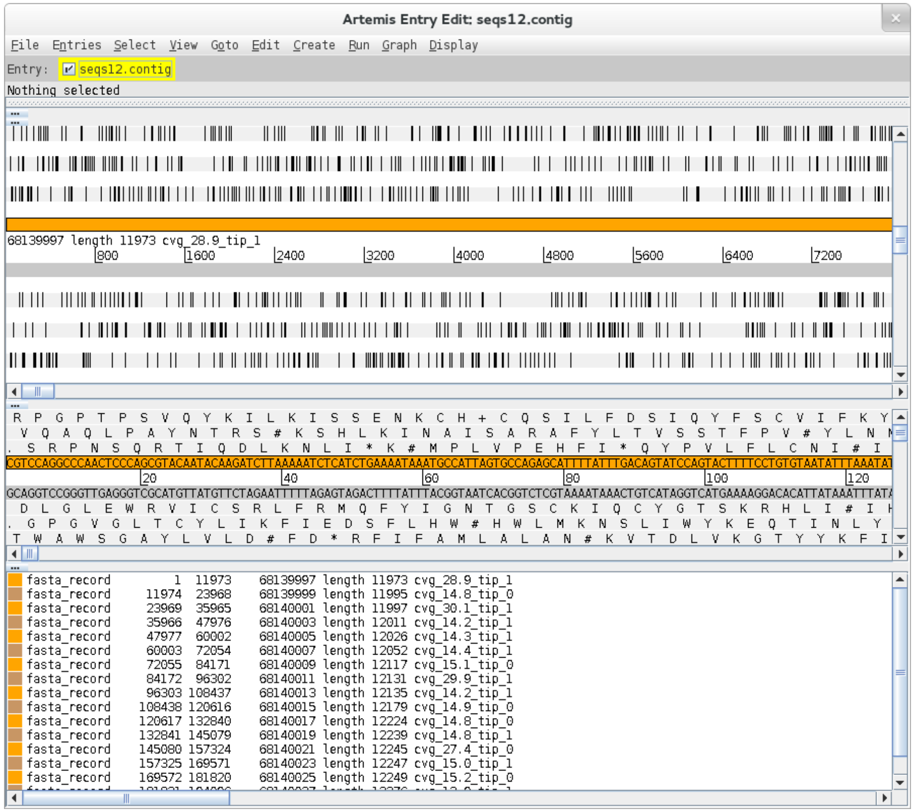
seqs1.contig.aa
seqs1.contig.codingseq
seqs1.contig.gff
seqs10.contig
seqs10.contig.aa
seqs10.contig.codingseq
seqs10.contig.gff
seqs11.contig
seqs11.contig.aa
seqs11.contig.codingseq
seqs11.contig.gff
seqs12.contig
seqs12.contig
seqs12.contig
seqs12.contig
seqs13.contig
seqs13.contig
seqs13.contig

Nome do Arquivo: seqs12.contig

Arquivos do Tipo: Todos os Arquivos

Anotação de genes no Artemis

Uma nova janela se abrirá com somente os cabeçalhos dos contigs identificados



```
R P G P T P S V Q Y K I L K I S S E N K C H + C Q S I L F D S I Q Y F S C V I F K Y
 V Q A Q L P A Y N T R S # K S H L K I N A I S A R A F Y L T V S S T F P V # Y L N
. S R P N S Q R T I Q D L K N L I * K # M P L V P E H F I * Q Y P V L F L C N I # I
CGTCCAGGCCCAACTCCCAGCGTACAATACAAGATCTTAAAAATCTCATCTGAAAATAAATGCCATTAGTGCCAGAGCATTTTATTTGACAGTATCCAGTACTTTTCCTGTGTAATATTTAAATAT
GCAGGTCCGGGTTGAGGGTCGCATGTTATGTTCTAGAATTTTTAGAGTAGACTTTTATTTACGGTAATCACGGTCTCGTAAAATAAACTGTCATAGGTCATGAAAAGGACACATTATAAATTTATA
 D L G L E W R V I C S R L F R M Q F Y I G N T G S C K I Q C Y G T S K R H L I # I
. G P G V G L T C Y L I K F I E D S F L H W # H W L M K N S L I W Y K E Q T I N L Y
T W A W S G A Y L V L D # F D * R F I F A M L A L A N # K V T D L V K G T Y Y K F I
```

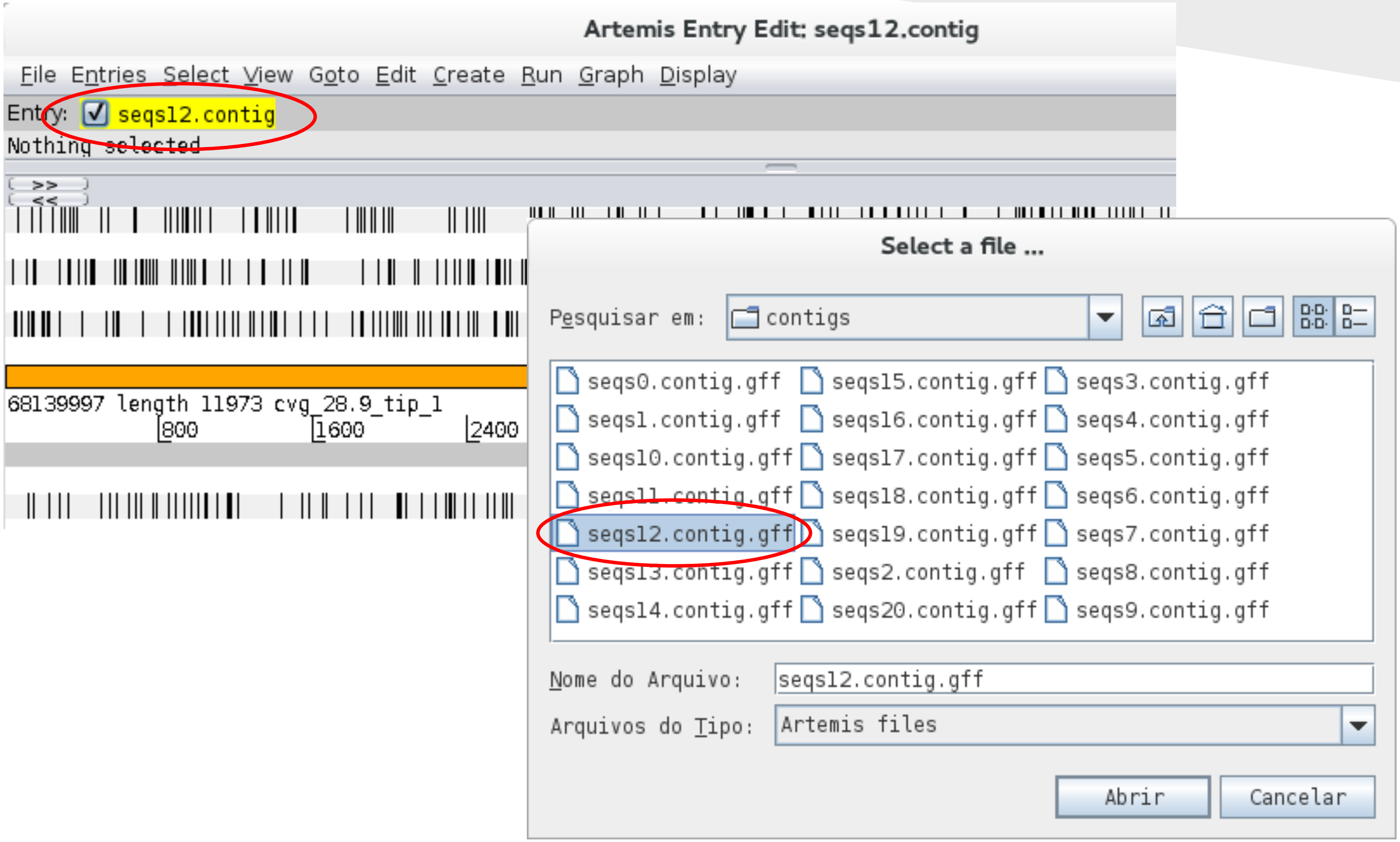
```
fasta_record       1  11973     68139997 length 11973 cvg_28.9_tip_1
fasta_record   11974  23968     68139999 length 11995 cvg_14.8_tip_0
fasta_record   23969  35965     68140001 length 11997 cvg_30.1_tip_1
fasta_record   35966  47976     68140003 length 12011 cvg_14.2_tip_1
fasta_record   47977  60002     68140005 length 12026 cvg_14.3_tip_1
fasta_record   60003  72054     68140007 length 12052 cvg_14.4_tip_1
fasta_record   72055  84171     68140009 length 12117 cvg_15.1_tip_0
fasta_record   84172  96302     68140011 length 12131 cvg_29.9_tip_1
fasta_record   96303 108437     68140013 length 12135 cvg_14.2_tip_1
fasta_record  108438 120616     68140015 length 12179 cvg_14.9_tip_0
fasta_record  120617 132840     68140017 length 12224 cvg_14.8_tip_0
fasta_record  132841 145079     68140019 length 12239 cvg_14.8_tip_1
fasta_record  145080 157324     68140021 length 12245 cvg_27.4_tip_0
fasta_record  157325 169571     68140023 length 12247 cvg_15.0_tip_1
fasta_record  169572 181820     68140025 length 12249 cvg_15.2_tip_0
```

Anotação de genes no Artemis

Na nova janela, navegue até o menu superior esquerdo `File` e selecione a opção `Read an entry...`

No diálogo que aparecerá, navegue até a pasta **QBQ2507** e selecione o **GFF** correspondente ao contig aberta visualizado. Clique em `Abrir`

Anotação de genes no Artemis



Entry: ☑ seqs12.contig

Nothing selected

68139997 length 11973 cvg_28.9_tip_1

Pesquisar em: contigs

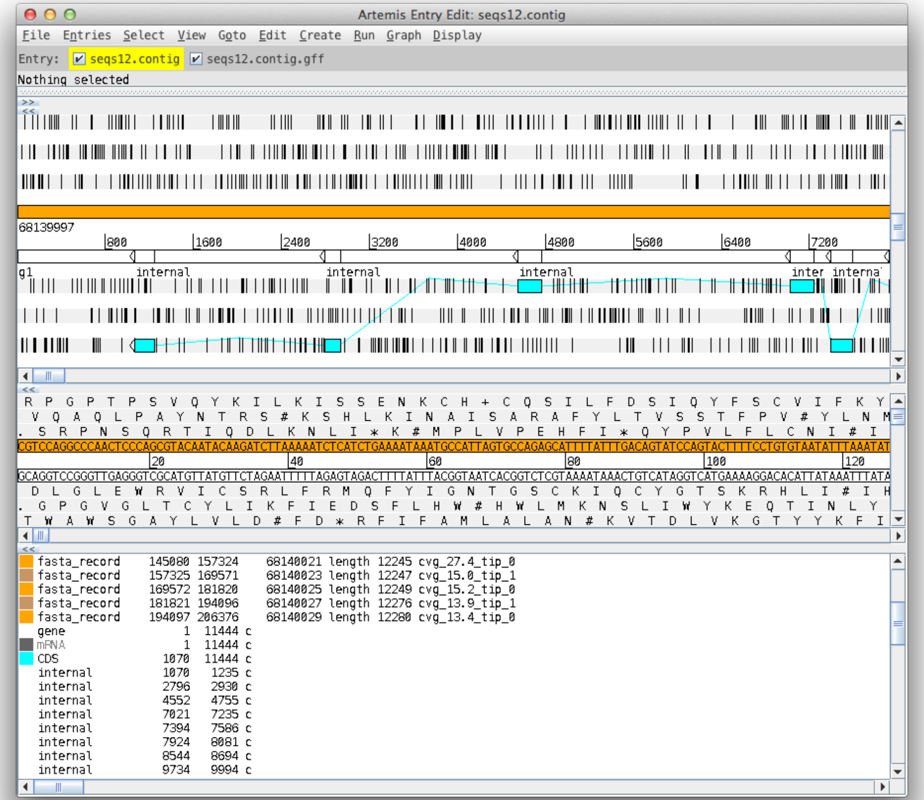
seqs0.contig.gff
seqs1.contig.gff
seqs10.contig.gff
seqs11.contig.gff
seqs12.contig.gff
seqs13.contig.gff
seqs14.contig.gff
seqs15.contig.gff
seqs16.contig.gff
seqs17.contig.gff
seqs18.contig.gff
seqs19.contig.gff
seqs2.contig.gff
seqs20.contig.gff
seqs3.contig.gff
seqs4.contig.gff
seqs5.contig.gff
seqs6.contig.gff
seqs7.contig.gff
seqs8.contig.gff
seqs9.contig.gff

Nome do Arquivo: seqs12.contig.gff

Arquivos do Tipo: Artemis files

Anotação de genes no Artemis

Os genes e CDS's preditos no arquivo GFF aparecerão na lista inferior



Nothing selected

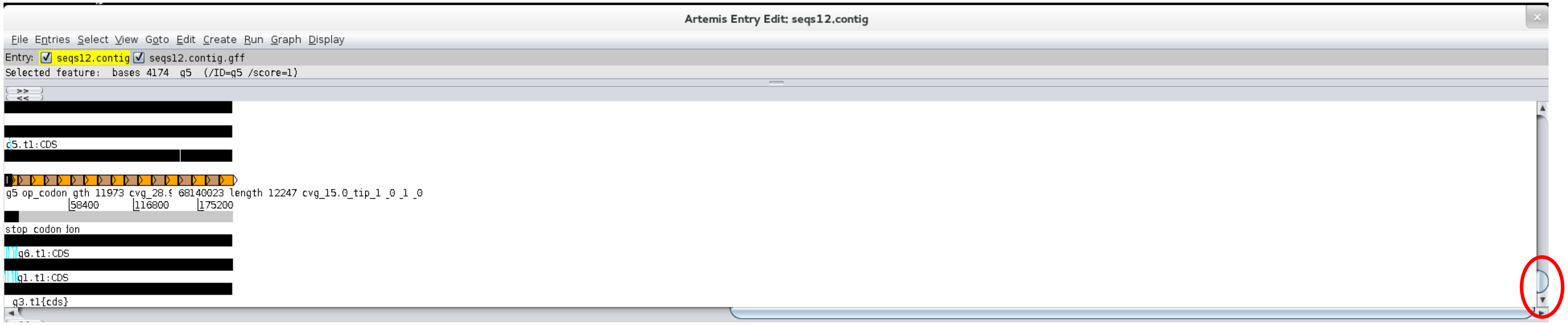
```
fasta_record     145080 157324     68140021 length 12245 cvg_27.4_tip_0
fasta_record     157325 169571     68140023 length 12247 cvg_15.0_tip_1
fasta_record     169572 181820     68140025 length 12249 cvg_15.2_tip_0
fasta_record     181821 194096     68140027 length 12276 cvg_13.9_tip_1
fasta_record     194097 206376     68140029 length 12280 cvg_13.4_tip_0
gene                  1  11444 c
mRNA                  1  11444 c
CDS                1070  11444 c
internal           1070   1235 c
internal           2796   2930 c
internal           4552   4755 c
internal           7021   7235 c
internal           7394   7586 c
internal           7924   8081 c
internal           8544   8694 c
internal           9734   9994 c
```

Anotação de genes no Artemis

O Artemis é dividido em 3 frames principais. Em cada frame, as barras de rolagem **horizontais** são utilizadas para navegar ao longo das sequências, ao passo de que as barras verticais são utilizadas para controlar o zoom de visualização das regiões.

Anotação de genes no Artemis

Frame com zoom mínimo (visualização de todas as sequências carregadas - barra vertical de rolagem posicionada no rodapé.)



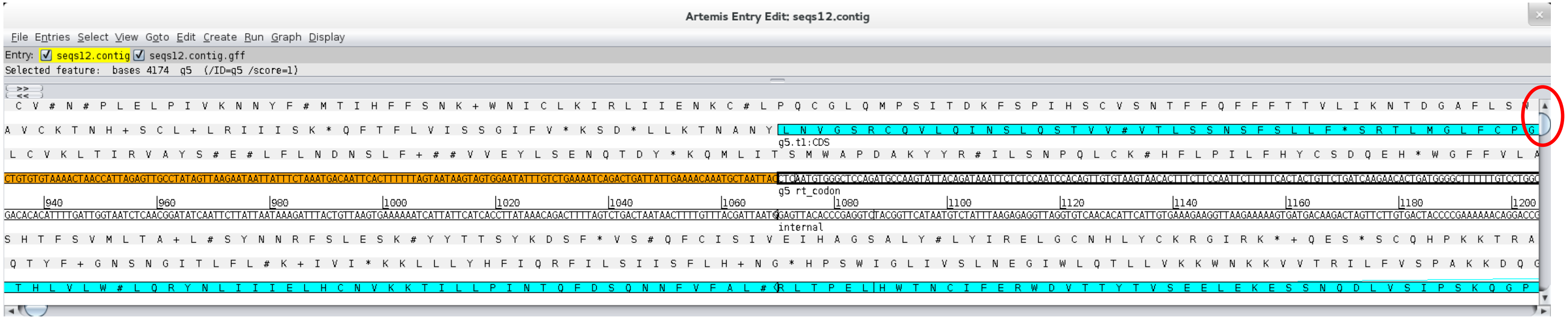
Artemis Entry Edit: seqs12.contig

Entry: ☑ seqs12.contig ☑ seqs12.contig.gff

Selected feature: bases 4174 g5 (/ID=g5 /score=1)

Anotação de genes no Artemis

Frame com zoom máximo (visualização dos nucleotídeos e aminoácidos das sequências)

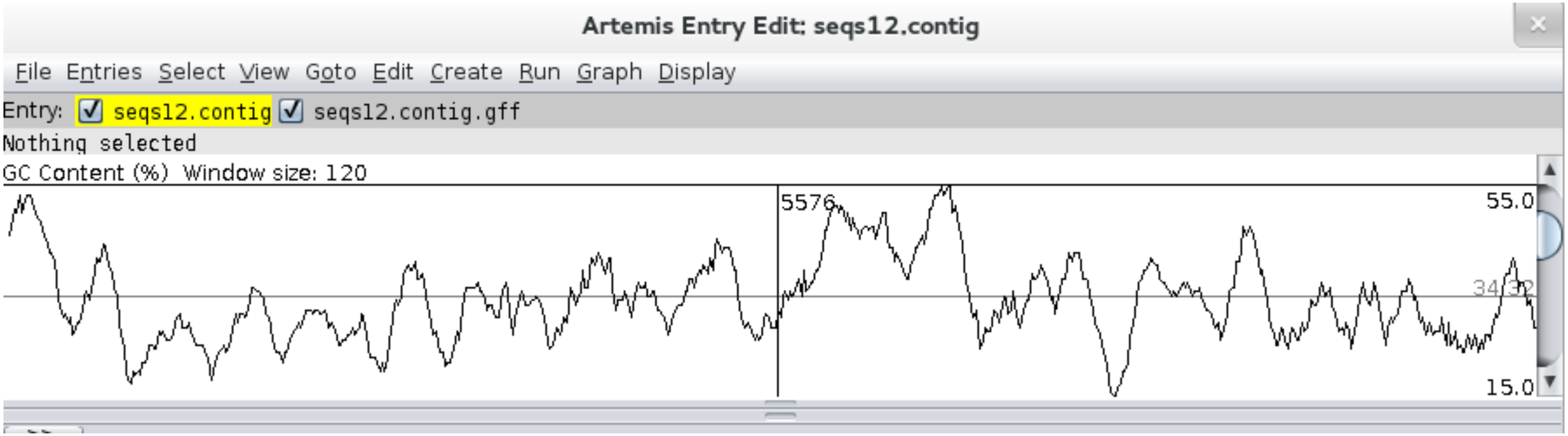


Entry: ☑ seqs12.contig ☑ seqs12.contig.gff

```
Selected feature:  bases 4174  g5  (/ID=g5 /score=1)

 C V # N # P L E L P I V K N N Y F # M T I H F F S N K + W N I C L K I R L I I E N K C # L P Q C G L Q M P S I T D K F S P I H S C V S N T F F Q F F F T T V L I K N T D G A F L S N
A V C K T N H + S C L + L R I I I S K * Q F T F L V I S S G I F V * K S D * L L K T N A N Y L N V G S R C Q V L Q I N S L Q S T V V # V T L S S N S F S L L F * S R T L M G L F C P G
                                                                                          g5.t1:CDS
 L C V K L T I R V A Y S # E # L F L N D N S L F + # # V V E Y L S E N Q T D Y * K Q M L I T S M W A P D A K Y Y R # I L S N P Q L C K # H F L P I L F H Y C S D Q E H * W G F F V L A

CTGTGTGTAAAACTAACCATTAGAGTTGCCTATAGTTAAGAATAATTATTTCTAAATGACAATTCACTTTTTTAGTAATAAGTAGTGGAATATTTGTCTGAAAATCAGACTGATTATTGAAAACAAATGCTAATTACCTCAATGTGGGCTCCAGATGCCAAGTATTACAGATAAATTCTCTCCAATCCACAGTTGTGTAAGTAACACTTTCTTCCAATTCTTTTTCACTACTGTTCTGATCAAGAACACTGATGGGGCTTTTTTGTCCTGGC
                                                                                          g5 rt_codon
    |940               |960               |980               |1000              |1020              |1040              |1060              |1080              |1100              |1120              |1140              |1160              |1180              |1200
GACACACATTTTGATTGGTAATCTCAACGGATATCAATTCTTATTAATAAAGATTTACTGTTAAGTGAAAAAATCATTATTCATCACCTTATAAACAGACTTTTAGTCTGACTAATAACTTTTGTTTACGATTAATGGAGTTACACCCGAGGTCTACGGTTCATAATGTCTATTTAAGAGAGGTTAGGTGTCAACACATTCATTGTGAAAGAAGGTTAAGAAAAAGTGATGACAAGACTAGTTCTTGTGACTACCCCGAAAAAACAGGACCG
                                                                                          internal
S H T F S V M L T A + L # S Y N N R F S L E S K # Y Y T T S Y K D S F * V S # Q F C I S I V E I H A G S A L Y # L Y I R E L G C N H L Y C K R G I R K * + Q E S * S C Q H P K K T R A
 Q T Y F + G N S N G I T L F L # K + I V I * K K L L L Y H F I Q R F I L S I I S F L H + N G * H P S W I G L I V S L N E G I W L Q T L L V K K W N K K V V T R I L F V S P A K K D Q G
 T H L V L W # L Q R Y N L I I I E L H C N V K K T I L L P I N T Q F D S Q N N F V F A L # R L T P E L H W T N C I F E R W D V T T Y T V S E E L E K E S S N Q D L V S I P S K Q G P
```

Anotação de genes no Artemis

Para exibir a porcentagem de conteúdo GC, clique em `Graph` no menu superior e depois selecione a opção `GC Content (%)`. Um gráfico com a quantidade de GC's por região aparecerá na parte superior da janela.



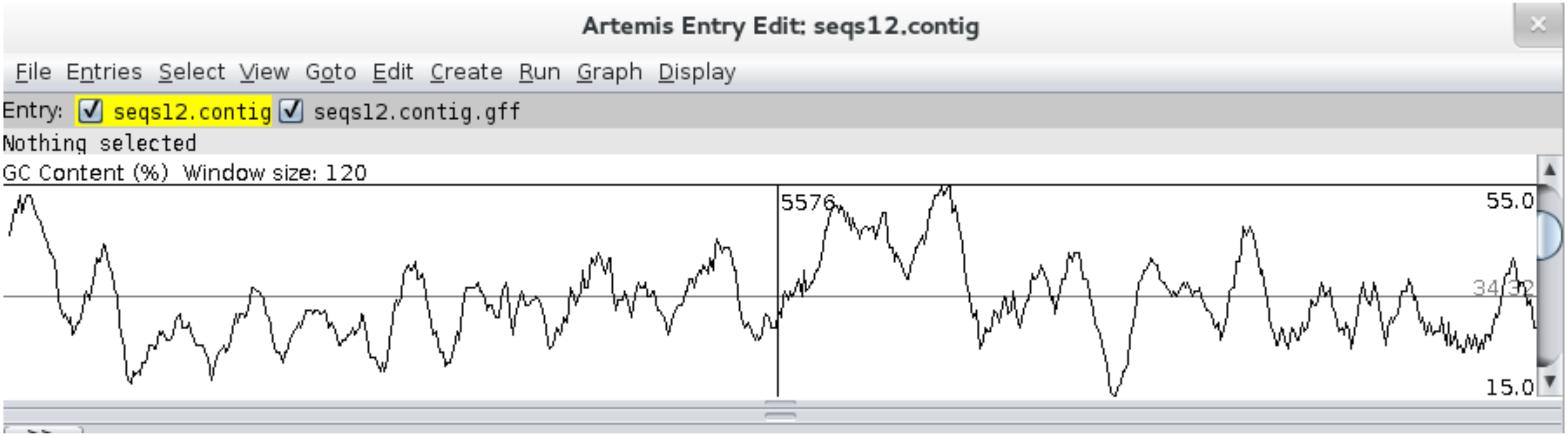
Entry: ☑ seqs12.contig ☑ seqs12.contig.gff

Nothing selected

GC Content (%) Window size: 120

Anotação de genes no Artemis

Para ocultar o gráfico, repita os passos anteriores, ou clique no menu `Graph` e selecione a opção `Hide all graphs`



Artemis Entry Edit: seqs12.contig

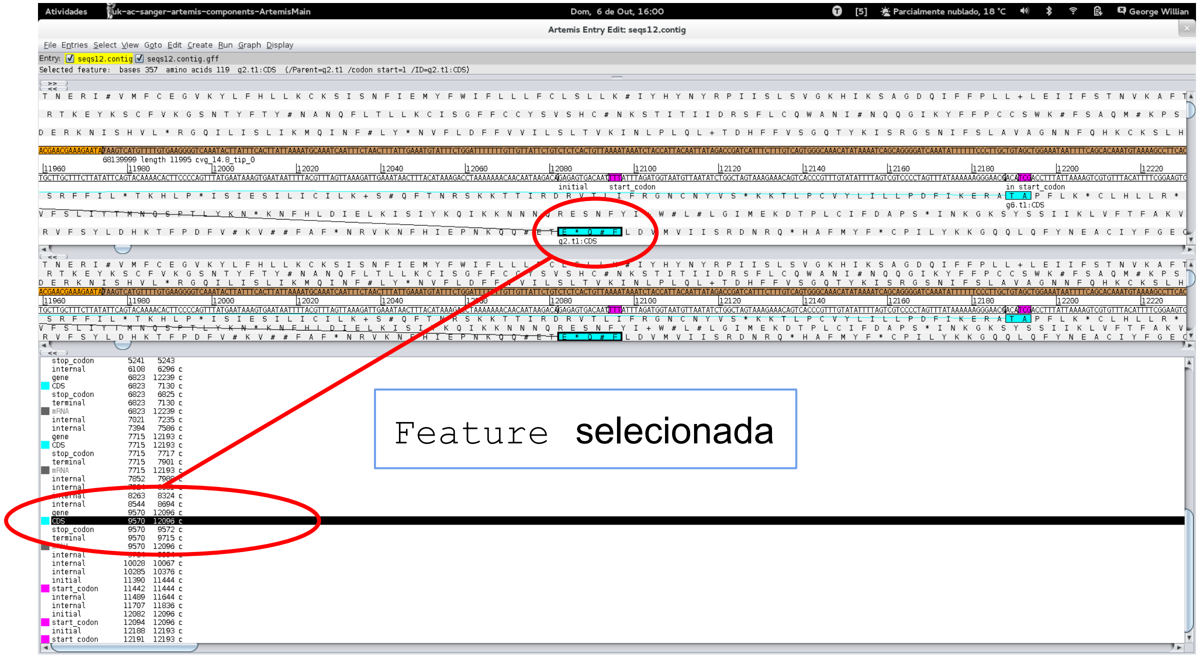
Entry: ☑ seqs12.contig ☑ seqs12.contig.gff

Nothing selected

GC Content (%) Window size: 120

Anotação de genes no Artemis

No frame inferior, dê um duplo clique no conjunto desejado (**gene** ou **CDS**) para que o visualizador seja posicionado nas sequências correspondentes ao conjunto.



Entry: ☑ seqs12.contig ☑ seqs12.contig.gff

Selected feature: bases 357 amino acids 119 q2.t1:CDS (/Parent=q2.t1 /codon start=1 /ID=q2.t1:CDS)

```
   stop_codon   5241   5243
   internal     6108   6296 c
   gene         6823  12239 c
 ▇ CDS          6823   7130 c
   stop_codon   6823   6825 c
   terminal     6823   7130 c
 ▇ mRNA         6823  12239 c
   internal     7021   7235 c
   internal     7394   7586 c
   gene         7715  12193 c
 ▇ CDS          7715  12193 c
   stop_codon   7715   7717 c
   terminal     7715   7901 c
 ▇ mRNA         7715  12193 c
   internal     7852   7988 c
   internal     [illegible]
   internal     8263   8324 c
   internal     8544   8694 c
   gene         9570  12096 c
 ▇ CDS          9570  12096 c
   stop_codon   9570   9572 c
   terminal     9570   9715 c
 ▇ [illegible]  9570  12096 c
   internal     [illegible]
   internal    10028  10067 c
   internal    10285  10376 c
   initial     11390  11444 c
 ▇ start_codon 11442  11444 c
   internal    11489  11644 c
   internal    11707  11836 c
   initial     12082  12096 c
 ▇ start_codon 12094  12096 c
   initial     12188  12193 c
 ▇ start_codon 12191  12193 c
```

Feature selecionada

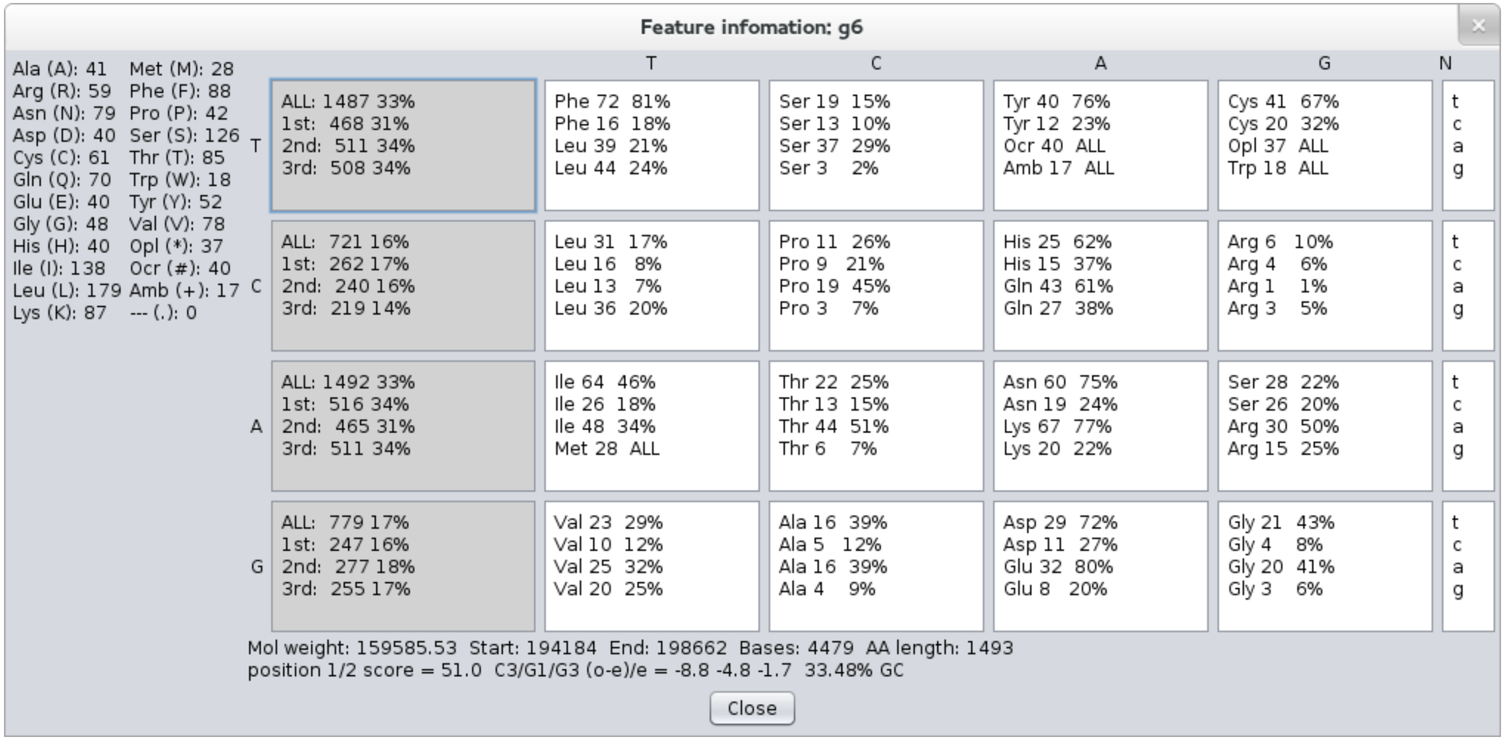
Anotação de genes no Artemis

Diversos parâmetros podem ser modificados:

1. base de dados utilizada;
2. filtros (baixas complexidades, repetições, etc.);
3. limiar do e-value;
4. gap open;
5. ...

Anotação de genes no Artemis

É possível visualizar as estatísticas de uma feature selecionando-a, clicando no menu `View` e selecionando a opção `Feature Statistics`



Feature infomation: g6

Ala (A): 41
Arg (R): 59
Asn (N): 79
Asp (D): 40
Cys (C): 61
Gln (Q): 70
Glu (E): 40
Gly (G): 48
His (H): 40
Ile (I): 138
Leu (L): 179
Lys (K): 87

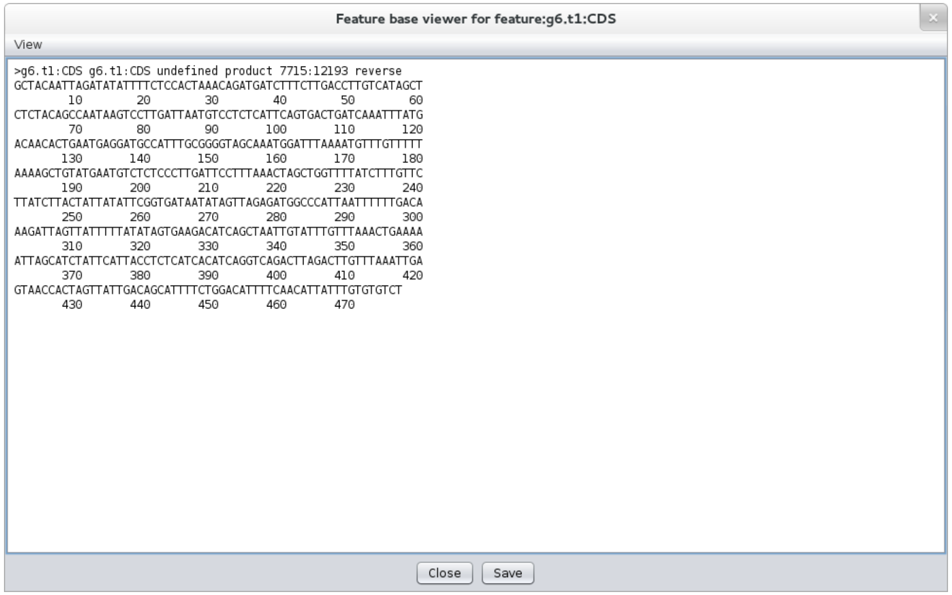
Met (M): 28
Phe (F): 88
Pro (P): 42
Ser (S): 126
Thr (T): 85
Trp (W): 18
Tyr (Y): 52
Val (V): 78
Opl (*): 37
Ocr (#): 40
Amb (+): 17
--- (.): 0

		T	C	A	G	N
T	ALL: 1487 33% 1st: 468 31% 2nd: 511 34% 3rd: 508 34%	Phe 72 81% Phe 16 18% Leu 39 21% Leu 44 24%	Ser 19 15% Ser 13 10% Ser 37 29% Ser 3 2%	Tyr 40 76% Tyr 12 23% Ocr 40 ALL Amb 17 ALL	Cys 41 67% Cys 20 32% Opl 37 ALL Trp 18 ALL	t c a g
C	ALL: 721 16% 1st: 262 17% 2nd: 240 16% 3rd: 219 14%	Leu 31 17% Leu 16 8% Leu 13 7% Leu 36 20%	Pro 11 26% Pro 9 21% Pro 19 45% Pro 3 7%	His 25 62% His 15 37% Gln 43 61% Gln 27 38%	Arg 6 10% Arg 4 6% Arg 1 1% Arg 3 5%	t c a g
A	ALL: 1492 33% 1st: 516 34% 2nd: 465 31% 3rd: 511 34%	Ile 64 46% Ile 26 18% Ile 48 34% Met 28 ALL	Thr 22 25% Thr 13 15% Thr 44 51% Thr 6 7%	Asn 60 75% Asn 19 24% Lys 67 77% Lys 20 22%	Ser 28 22% Ser 26 20% Arg 30 50% Arg 15 25%	t c a g
G	ALL: 779 17% 1st: 247 16% 2nd: 277 18% 3rd: 255 17%	Val 23 29% Val 10 12% Val 25 32% Val 20 25%	Ala 16 39% Ala 5 12% Ala 16 39% Ala 4 9%	Asp 29 72% Asp 11 27% Glu 32 80% Glu 8 20%	Gly 21 43% Gly 4 8% Gly 20 41% Gly 3 6%	t c a g

Mol weight: 159585.53 Start: 194184 End: 198662 Bases: 4479 AA length: 1493
position 1/2 score = 51.0 C3/G1/G3 (o-e)/e = -8.8 -4.8 -1.7 33.48% GC

Anotação de genes no Artemis

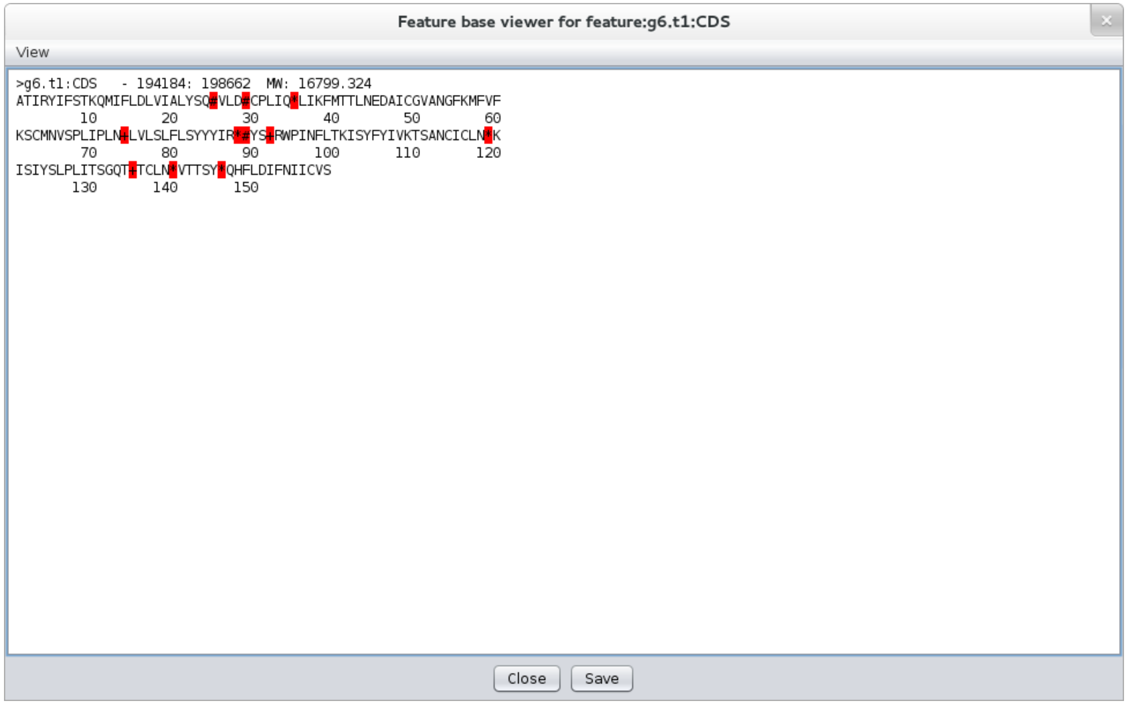
Para visualizar as bases de uma determinada feature, basta selecioná-la, clicar no menu `View` e selecionar a opção `Bases of Selection`, ou ainda, `Bases of Selecion As Fasta`, caso queira visualizar as bases em formato **FASTA**.



```
>g6.t1:CDS g6.t1:CDS undefined product 7715:12193 reverse
GCTACAATTAGATATATTTTCTCCACTAAACAGATGATCTTTCTTGACCTTGTCATAGCT
        10        20        30        40        50        60
CTCTACAGCCAATAAGTCCTTGATTAATGTCCTCTCATTCAGTGACTGATCAAATTTATG
        70        80        90       100       110       120
ACAACACTGAATGAGGATGCCATTTGCGGGGTAGCAAATGGATTTAAAATGTTTGTTTTT
       130       140       150       160       170       180
AAAAGCTGTATGAATGTCTCTCCCTTGATTCCTTTAAACTAGCTGGTTTTATCTTTGTTC
       190       200       210       220       230       240
TTATCTTACTATTATATTCGGTGATAATATAGTTAGAGATGGCCCATTAATTTTTTGACA
       250       260       270       280       290       300
AAGATTAGTTATTTTTATATAGTGAAGACATCAGCTAATTGTATTTGTTTAAACTGAAAA
       310       320       330       340       350       360
ATTAGCATCTATTCATTACCTCTCATCACATCAGGTCAGACTTAGACTTGTTTAAATTGA
       370       380       390       400       410       420
GTAACCACTAGTTATTGACAGCATTTTCTGGACATTTTCAACATTATTTGTGTGTCT
       430       440       450       460       470
```

Anotação de genes no Artemis

Da mesma maneira, é posível visualizar os aminoácidos de um CDS. Selecione o CDS, clique no menu `View` e selecione a opção `Aminoacids of Selection`, ou ainda, `Aminoacids of Selecion As Fasta`.



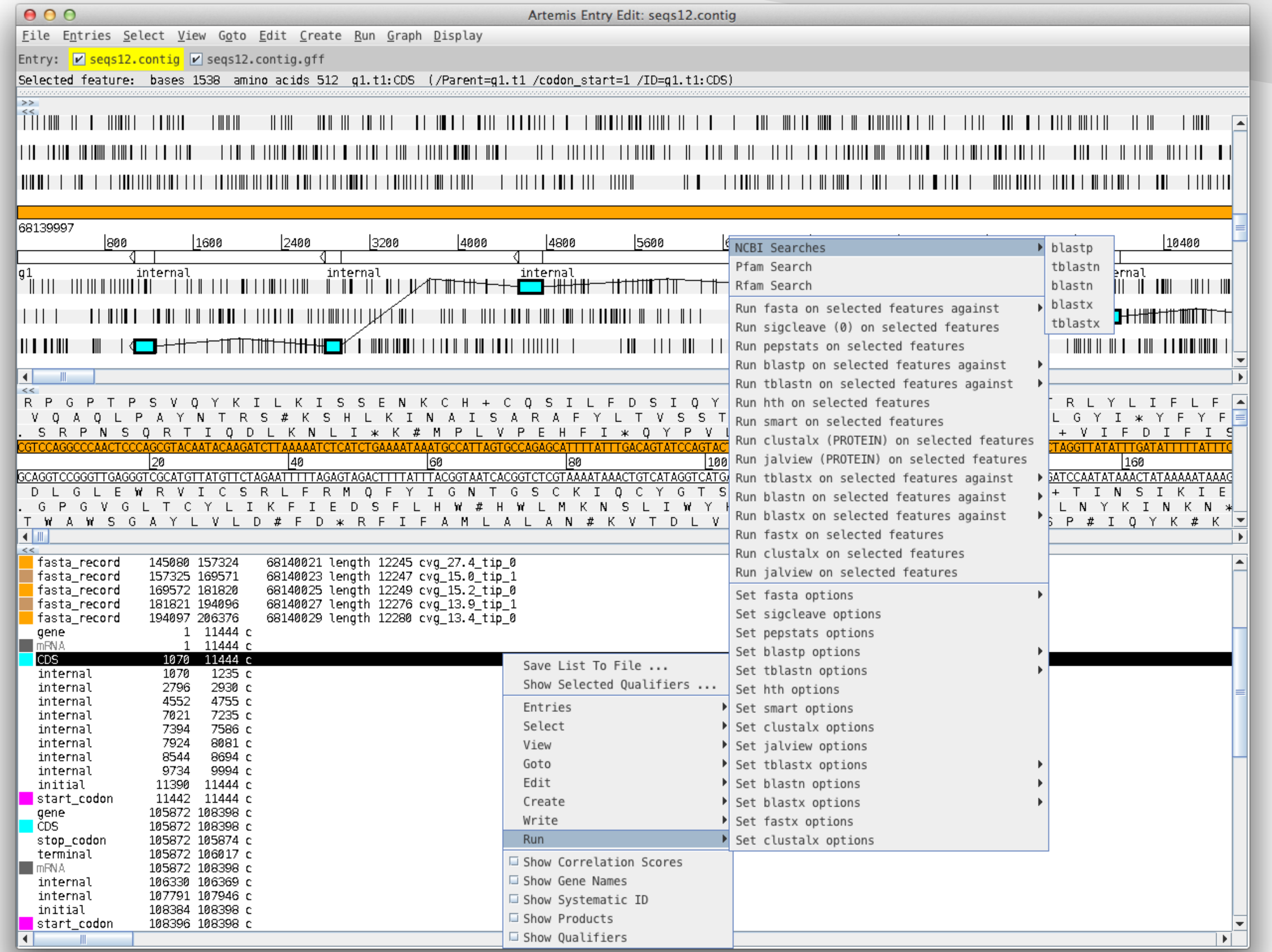
```
>g6.t1:CDS   - 194184: 198662  MW: 16799.324
ATIRYIFSTKQMIFLDLVIALYSQ#VLD#CPLIQ*LIKFMTTLNEDAICGVANGFKMFVF
         10        20        30        40        50        60
KSCMNVSPLIPLN+LVLSLFLSYYYIR*#YS+RWPINFLTKISYFYIVKTSANCICLN*K
         70        80        90       100       110       120
ISIYSLPLITSGQT+TCLN*VTTSY*QHFLDIFNIICVS
        130       140       150
```

Anotação de genes no Artemis

O Artemis permite a execução de 5 tipos diferentes de BLAST utilizando a interface web do NCBI;

Selecione a `Feature` que deseja "blastar", clique com o botão direito e navegue pela opção `Run` até `NCBI Searches`;

Anotação de genes no Artemis

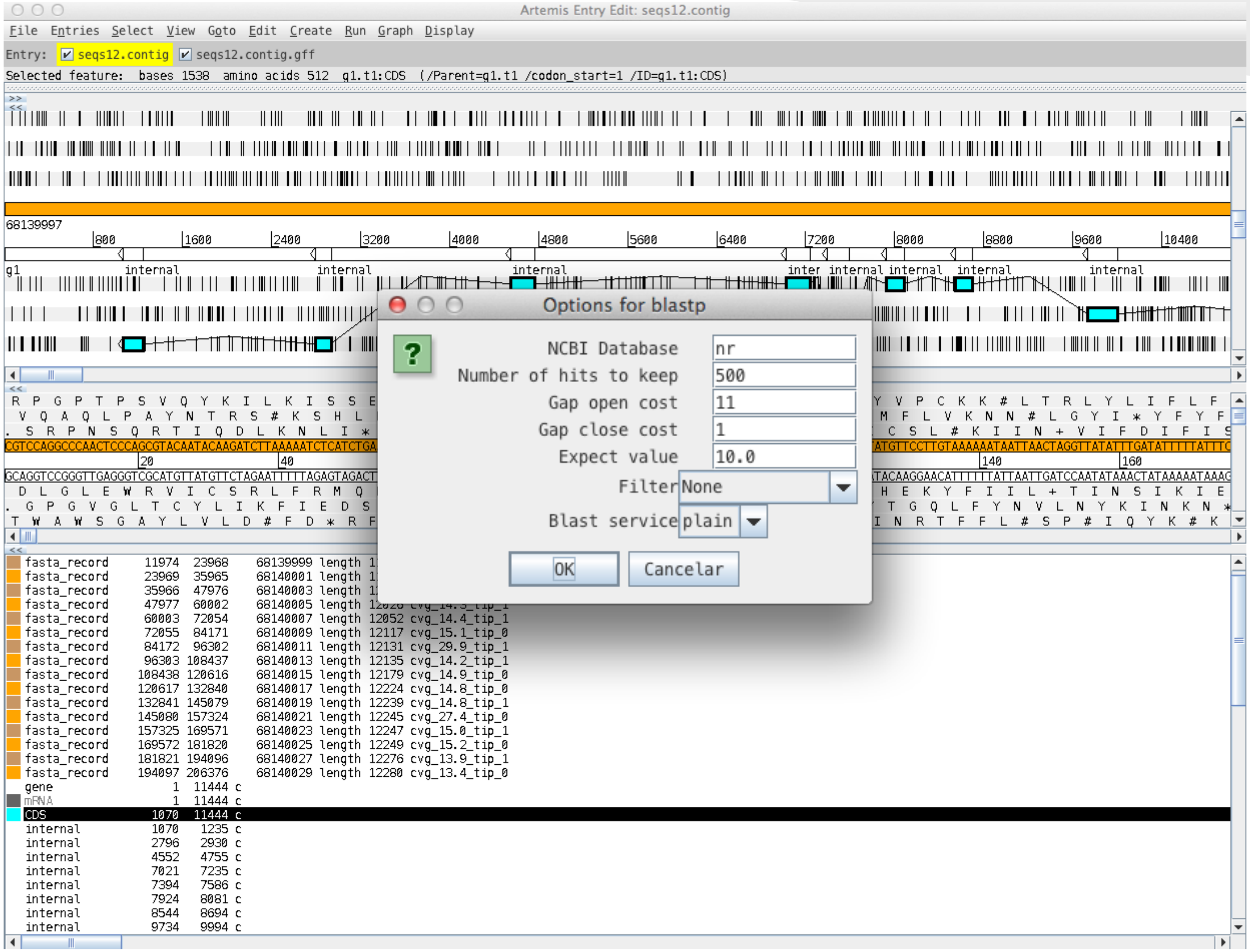


Entry: ☑ seqs12.contig ☑ seqs12.contig.gff

Selected feature: bases 1538 amino acids 512 g1.t1:CDS (/Parent=g1.t1 /codon_start=1 /ID=g1.t1:CDS)

```
fasta_record    145080 157324    68140021 length 12245 cvg_27.4_tip_0
fasta_record    157325 169571    68140023 length 12247 cvg_15.0_tip_1
fasta_record    169572 181820    68140025 length 12249 cvg_15.2_tip_0
fasta_record    181821 194096    68140027 length 12276 cvg_13.9_tip_1
fasta_record    194097 206376    68140029 length 12280 cvg_13.4_tip_0
gene                 1  11444 c
mRNA                 1  11444 c
CDS               1070  11444 c
internal          1070   1235 c
internal          2796   2930 c
internal          4552   4755 c
internal          7021   7235 c
internal          7394   7586 c
internal          7924   8081 c
internal          8544   8694 c
internal          9734   9994 c
initial          11390  11444 c
start_codon      11442  11444 c
gene            105872 108398 c
CDS             105872 108398 c
stop_codon      105872 105874 c
terminal        105872 106017 c
mRNA            105872 108398 c
internal        106330 106369 c
internal        107791 107946 c
initial         108384 108398 c
start_codon     108396 108398 c
```

Anotação de genes no Artemis



Entry: ☑ seqs12.contig ☑ seqs12.contig.gff

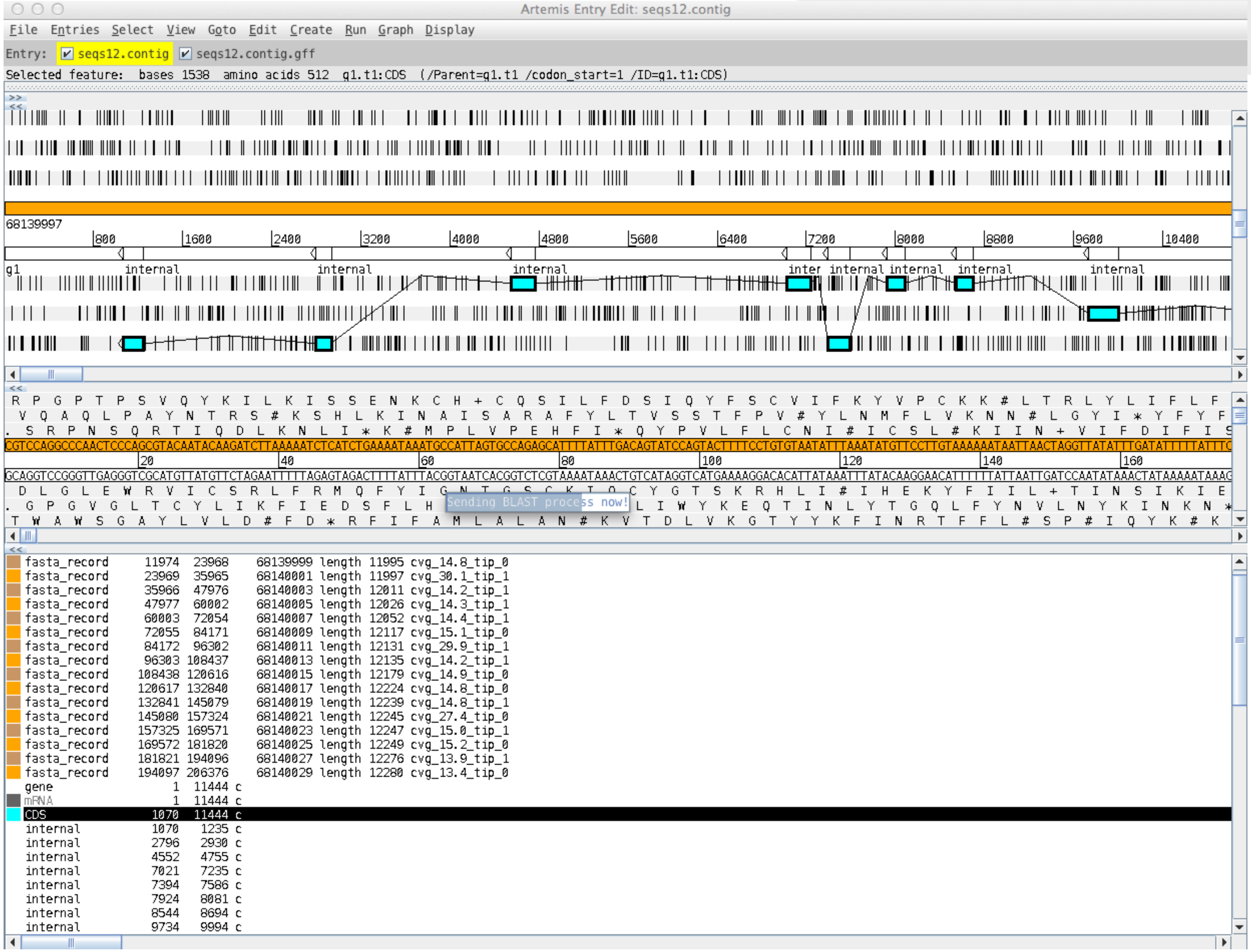
Selected feature: bases 1538 amino acids 512 g1.t1:CDS (/Parent=g1.t1 /codon_start=1 /ID=g1.t1:CDS)

Options for blastp

NCBI Database: nr
Number of hits to keep: 500
Gap open cost: 11
Gap close cost: 1
Expect value: 10.0
Filter: None
Blast service: plain

```
fasta_record     11974  23968     68139999 length 1
fasta_record     23969  35965     68140001 length 1
fasta_record     35966  47976     68140003 length 1
fasta_record     47977  60002     68140005 length 12028 cvg_14.3_tip_1
fasta_record     60003  72054     68140007 length 12052 cvg_14.4_tip_1
fasta_record     72055  84171     68140009 length 12117 cvg_15.1_tip_0
fasta_record     84172  96302     68140011 length 12131 cvg_29.9_tip_1
fasta_record     96303 108437     68140013 length 12135 cvg_14.2_tip_1
fasta_record    108438 120616     68140015 length 12179 cvg_14.9_tip_0
fasta_record    120617 132840     68140017 length 12224 cvg_14.8_tip_0
fasta_record    132841 145079     68140019 length 12239 cvg_14.8_tip_1
fasta_record    145080 157324     68140021 length 12245 cvg_27.4_tip_0
fasta_record    157325 169571     68140023 length 12247 cvg_15.0_tip_1
fasta_record    169572 181820     68140025 length 12249 cvg_15.2_tip_0
fasta_record    181821 194096     68140027 length 12276 cvg_13.9_tip_1
fasta_record    194097 206376     68140029 length 12280 cvg_13.4_tip_0
gene                 1  11444 c
mRNA                 1  11444 c
CDS               1070  11444 c
internal          1070   1235 c
internal          2796   2930 c
internal          4552   4755 c
internal          7021   7235 c
internal          7394   7586 c
internal          7924   8081 c
internal          8544   8694 c
internal          9734   9994 c
```

Anotação de genes no Artemis



Artemis Entry Edit: seqs12.contig

Entry: ☑ seqs12.contig ☑ seqs12.contig.gff

```
Selected feature:  bases 1538  amino acids 512  g1.t1:CDS  (/Parent=g1.t1 /codon_start=1 /ID=g1.t1:CDS)
```

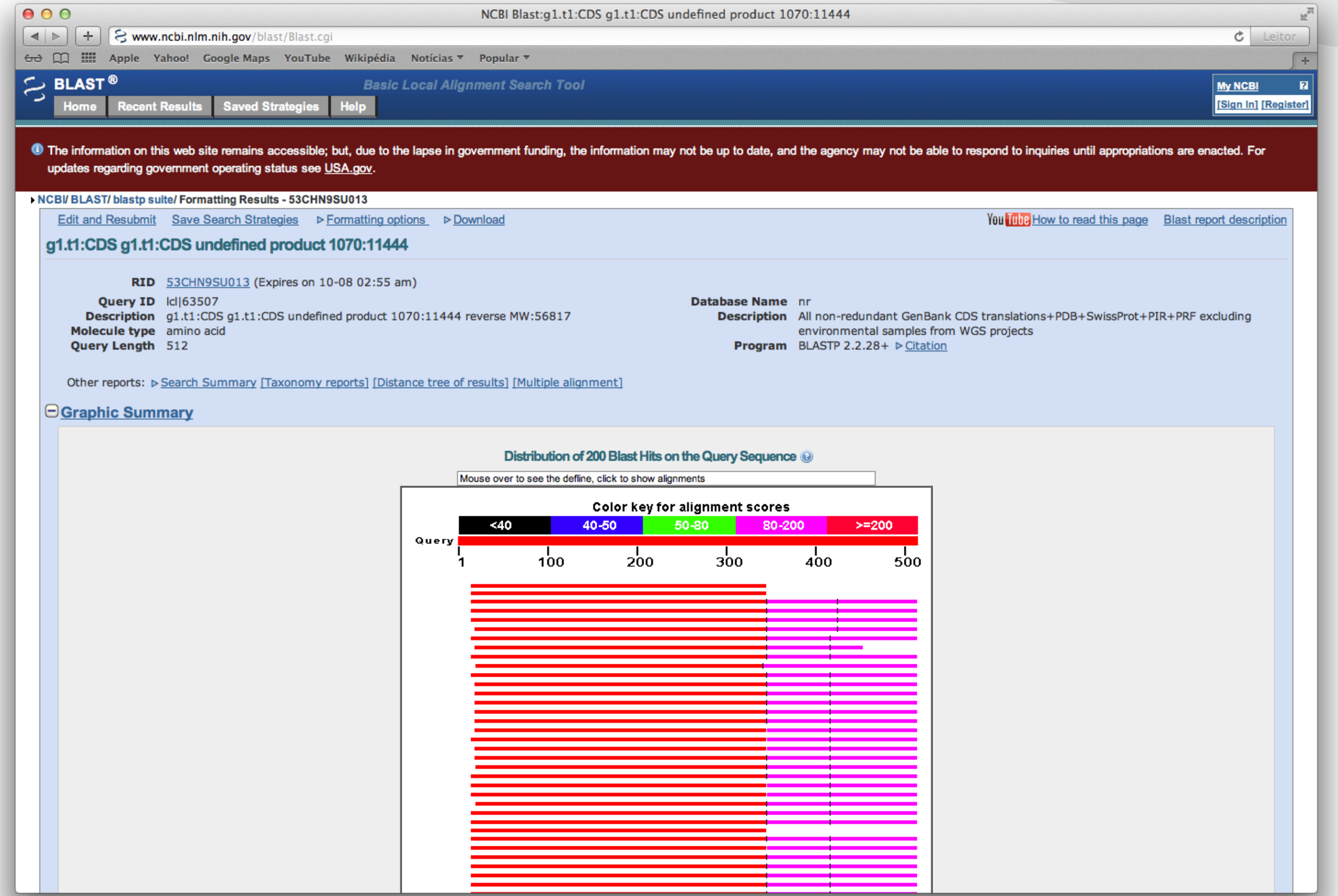
```
 R P G P T P S V Q Y K I L K I S S E N K C H + C Q S I L F D S I Q Y F S C V I F K Y V P C K K # L T R L Y L I F L F
  V Q A Q L P A Y N T R S # K S H L K I N A I S A R A F Y L T V S S T F P V # Y L N M F L V K N N # L G Y I * Y F Y F
. S R P N S Q R T I Q D L K N L I * K # M P L V P E H F I * Q Y P V L F L C N I # I C S L # K I I N + V I F D I F I S
CGTCCAGGCCCAACTCCCAGCGTACAATACAAGATCTTAAAAATCTCATCTGAAAATAAATGCCATTAGTGCCAGAGCATTTTATTTGACAGTATCCAGTACTTTTCCTGTGTAATATTTAAATATGTTCCTTGTAAAAAATAATTAACTAGGTTATATTTGATATTTTTATTTC
                   20                  40                  60                  80                  100                 120                 140                 160
GCAGGTCCGGGTTGAGGGTCGCATGTTATGTTCTAGAATTTTTAGAGTAGACTTTTATTTACGGTAATCACGGTCTCGTAAAATAAACTGTCATAGGTCATGAAAAGGACACATTATAAATTTATACAAGGAACATTTTTTATTAATTGATCCAATATAAACTATAAAAATAAAG
  D L G L E W R V I C S R L F R M Q F Y I G N T G S C K I Q C Y G T S K R H L I # I H E K Y F I I L + T I N S I K I E
. G P G V G L T C Y L I K F I E D S F L H Sending BLAST process now! L I W Y K E Q T I N L Y T G Q L F Y N V L N Y K I N K N *
 T W A W S G A Y L V L D # F D * R F I F A M L A L A N # K V T D L V K G T Y Y K F I N R T F F L # S P # I Q Y K # K
```

```
fasta_record    11974  23968   68139999 length 11995 cvg_14.8_tip_0
fasta_record    23969  35965   68140001 length 11997 cvg_30.1_tip_1
fasta_record    35966  47976   68140003 length 12011 cvg_14.2_tip_1
fasta_record    47977  60002   68140005 length 12026 cvg_14.3_tip_1
fasta_record    60003  72054   68140007 length 12052 cvg_14.4_tip_1
fasta_record    72055  84171   68140009 length 12117 cvg_15.1_tip_0
fasta_record    84172  96302   68140011 length 12131 cvg_29.9_tip_1
fasta_record    96303 108437   68140013 length 12135 cvg_14.2_tip_1
fasta_record   108438 120616   68140015 length 12179 cvg_14.9_tip_0
fasta_record   120617 132840   68140017 length 12224 cvg_14.8_tip_0
fasta_record   132841 145079   68140019 length 12239 cvg_14.8_tip_1
fasta_record   145080 157324   68140021 length 12245 cvg_27.4_tip_0
fasta_record   157325 169571   68140023 length 12247 cvg_15.0_tip_1
fasta_record   169572 181820   68140025 length 12249 cvg_15.2_tip_0
fasta_record   181821 194096   68140027 length 12276 cvg_13.9_tip_1
fasta_record   194097 206376   68140029 length 12280 cvg_13.4_tip_0
gene                1  11444 c
mRNA                1  11444 c
CDS              1070  11444 c
internal         1070   1235 c
internal         2796   2930 c
internal         4552   4755 c
internal         7021   7235 c
internal         7394   7586 c
internal         7924   8081 c
internal         8544   8694 c
internal         9734   9994 c
```

Anotação de genes no Artemis

O resultado do BLAST em formato HTML é automaticamente aberto pelo Browser (não confundir com Bowser).

Anotação de genes no Artemis



The information on this web site remains accessible; but, due to the lapse in government funding, the information may not be up to date, and the agency may not be able to respond to inquiries until appropriations are enacted. For updates regarding government operating status see USA.gov.

NCBI/ BLAST/ blastp suite/ Formatting Results - 53CHN9SU013

g1.t1:CDS g1.t1:CDS undefined product 1070:11444

RID 53CHN9SU013 (Expires on 10-08 02:55 am)
Query ID lcl|63507
Description g1.t1:CDS g1.t1:CDS undefined product 1070:11444 reverse MW:56817
Molecule type amino acid
Query Length 512

Database Name nr
Description All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples from WGS projects
Program BLASTP 2.2.28+ Citation

Other reports: Search Summary [Taxonomy reports] [Distance tree of results] [Multiple alignment]

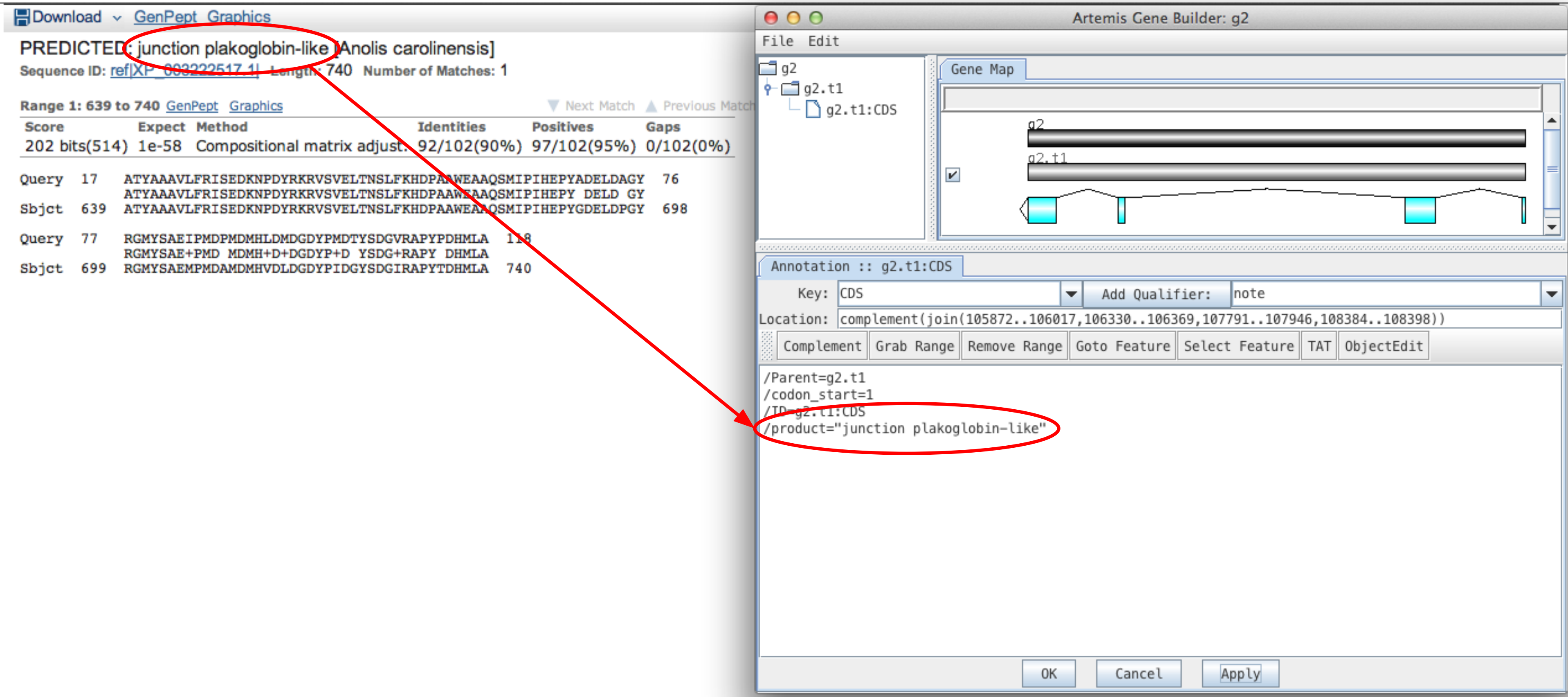
Graphic Summary

Distribution of 200 Blast Hits on the Query Sequence

Mouse over to see the define, click to show alignments

Anotação de genes no Artemis

O resultado do BLAST pode ser utilizado para caracterizar o gene através do Artemis;

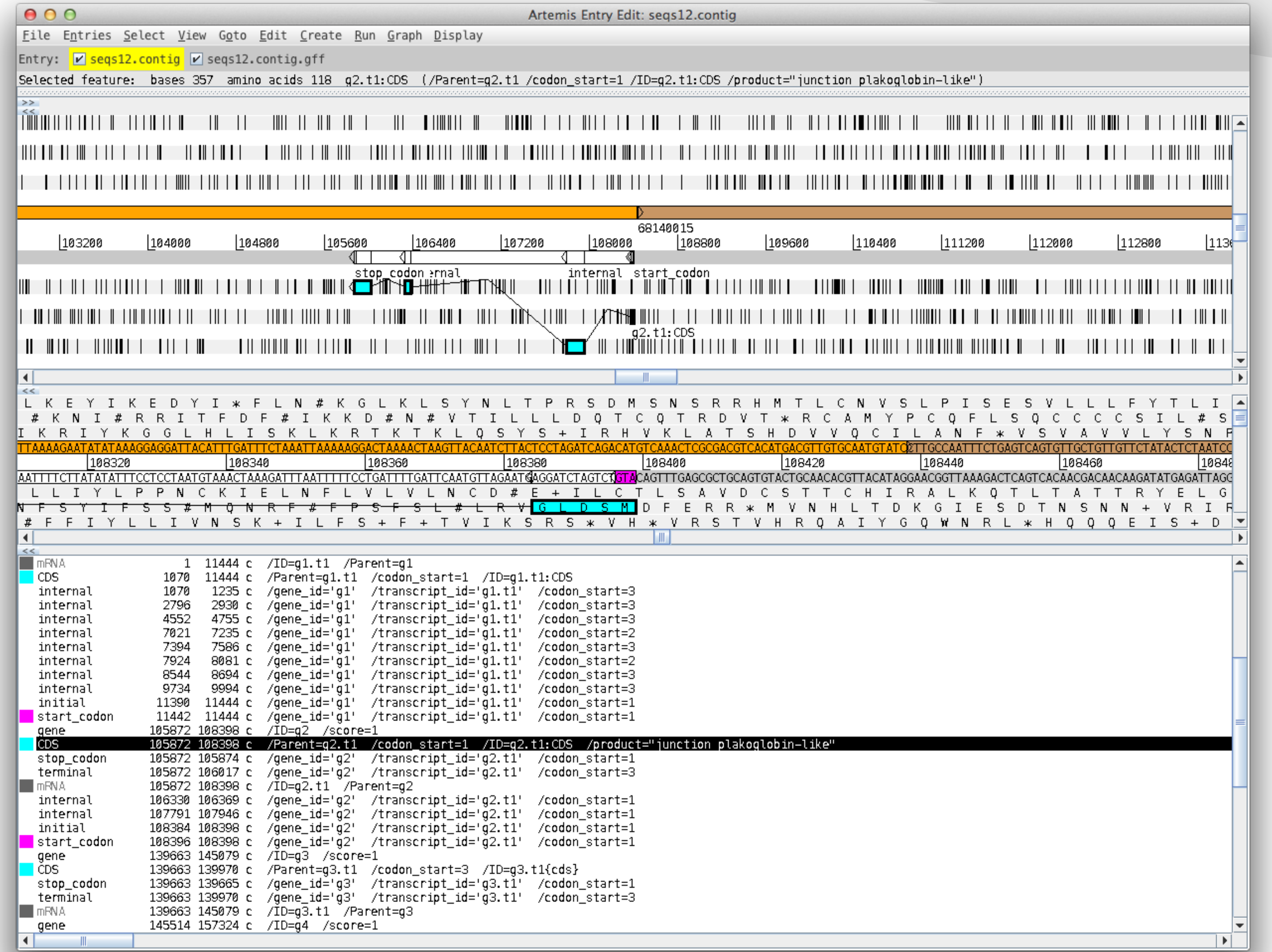


```
PREDICTED: junction plakoglobin-like [Anolis carolinensis]
Sequence ID: ref|XP_003222517.1| Length: 740   Number of Matches: 1

Range 1: 639 to 740 GenPept  Graphics
 Score          Expect  Method                         Identities     Positives      Gaps
 202 bits(514)  1e-58   Compositional matrix adjust.   92/102(90%)    97/102(95%)    0/102(0%)

Query  17    ATYAAAVLFRISEDKNPDYRKRVSVELTNSLFKHDPAAWEAAQSMIPIHEPYADELDAGY  76
             ATYAAAVLFRISEDKNPDYRKRVSVELTNSLFKHDPAAWEAAQSMIPIHEPY DELD GY
Sbjct  639   ATYAAAVLFRISEDKNPDYRKRVSVELTNSLFKHDPAAWEAAQSMIPIHEPYGDELDPGY  698

Query  77    RGMYSAEIPMDPMDMHLDMDGDYPMDTYSDGVRAPYPDHMLA  118
             RGMYSAE+PMD MDMH+D+DGDYP+D YSDG+RAPY DHMLA
Sbjct  699   RGMYSAEMPMDAMDMHVDLDGDYPIDGYSDGIRAPYTDHMLA  740
```

Key: CDS
Add Qualifier: note
Location: complement(join(105872..106017,106330..106369,107791..107946,108384..108398))

```
/Parent=g2.t1
/codon_start=1
/ID=g2.t1:CDS
/product="junction plakoglobin-like"
```

Anotação de genes no Artemis



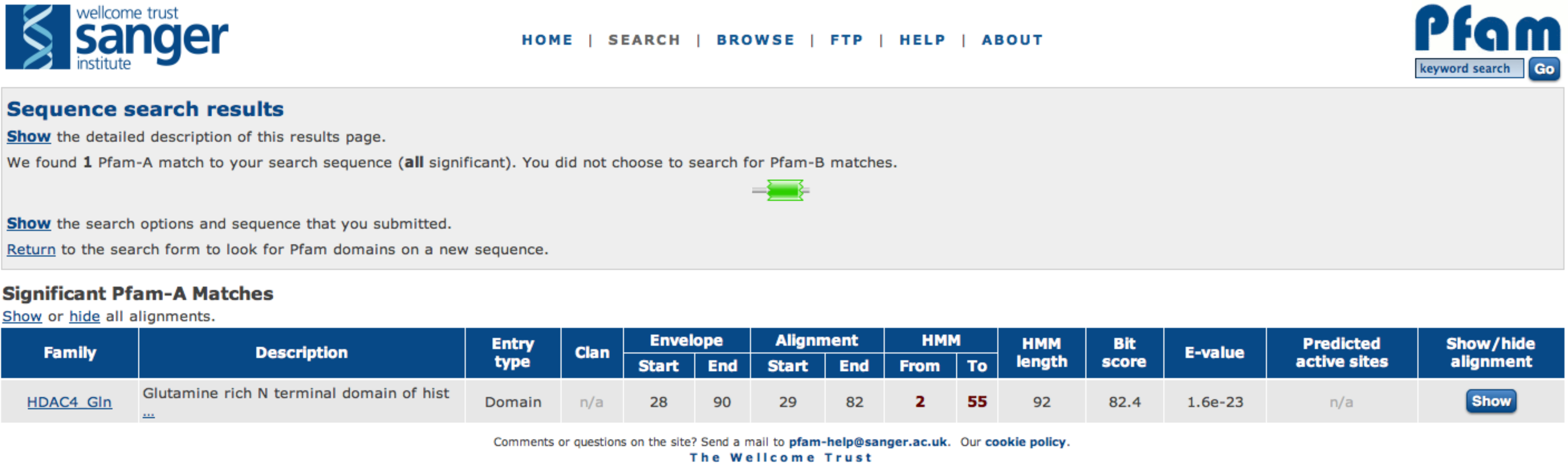
Entry: ☑ seqs12.contig ☑ seqs12.contig.gff

```
Selected feature:  bases 357  amino acids 118  g2.t1:CDS  (/Parent=g2.t1 /codon_start=1 /ID=g2.t1:CDS /product="junction plakoglobin-like")
```

```
mRNA              1  11444 c  /ID=g1.t1  /Parent=g1
CDS            1070  11444 c  /Parent=g1.t1  /codon_start=1  /ID=g1.t1:CDS
internal       1070   1235 c  /gene_id='g1'  /transcript_id='g1.t1'  /codon_start=3
internal       2796   2930 c  /gene_id='g1'  /transcript_id='g1.t1'  /codon_start=3
internal       4552   4755 c  /gene_id='g1'  /transcript_id='g1.t1'  /codon_start=3
internal       7021   7235 c  /gene_id='g1'  /transcript_id='g1.t1'  /codon_start=2
internal       7394   7586 c  /gene_id='g1'  /transcript_id='g1.t1'  /codon_start=3
internal       7924   8081 c  /gene_id='g1'  /transcript_id='g1.t1'  /codon_start=2
internal       8544   8694 c  /gene_id='g1'  /transcript_id='g1.t1'  /codon_start=3
internal       9734   9994 c  /gene_id='g1'  /transcript_id='g1.t1'  /codon_start=3
initial       11390  11444 c  /gene_id='g1'  /transcript_id='g1.t1'  /codon_start=1
start_codon   11442  11444 c  /gene_id='g1'  /transcript_id='g1.t1'  /codon_start=1
gene         105872 108398 c  /ID=g2  /score=1
CDS          105872 108398 c  /Parent=g2.t1  /codon_start=1  /ID=g2.t1:CDS  /product="junction plakoglobin-like"
stop_codon   105872 105874 c  /gene_id='g2'  /transcript_id='g2.t1'  /codon_start=1
terminal     105872 106017 c  /gene_id='g2'  /transcript_id='g2.t1'  /codon_start=3
mRNA         105872 108398 c  /ID=g2.t1  /Parent=g2
internal     106330 106369 c  /gene_id='g2'  /transcript_id='g2.t1'  /codon_start=1
internal     107791 107946 c  /gene_id='g2'  /transcript_id='g2.t1'  /codon_start=1
initial      108384 108398 c  /gene_id='g2'  /transcript_id='g2.t1'  /codon_start=1
start_codon  108396 108398 c  /gene_id='g2'  /transcript_id='g2.t1'  /codon_start=1
gene         139663 145079 c  /ID=g3  /score=1
CDS          139663 139970 c  /Parent=g3.t1  /codon_start=3  /ID=g3.t1{cds}
stop_codon   139663 139665 c  /gene_id='g3'  /transcript_id='g3.t1'  /codon_start=1
terminal     139663 139970 c  /gene_id='g3'  /transcript_id='g3.t1'  /codon_start=3
mRNA         139663 145079 c  /ID=g3.t1  /Parent=g3
gene         145514 157324 c  /ID=g4  /score=1
```

Anotação de genes no Artemis

Também é possivel fazer uma busca por domínios conservados do Pfam diretamente pelo Artemis.



Sequence search results

Show the detailed description of this results page.

We found **1** Pfam-A match to your search sequence (**all** significant). You did not choose to search for Pfam-B matches.

Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

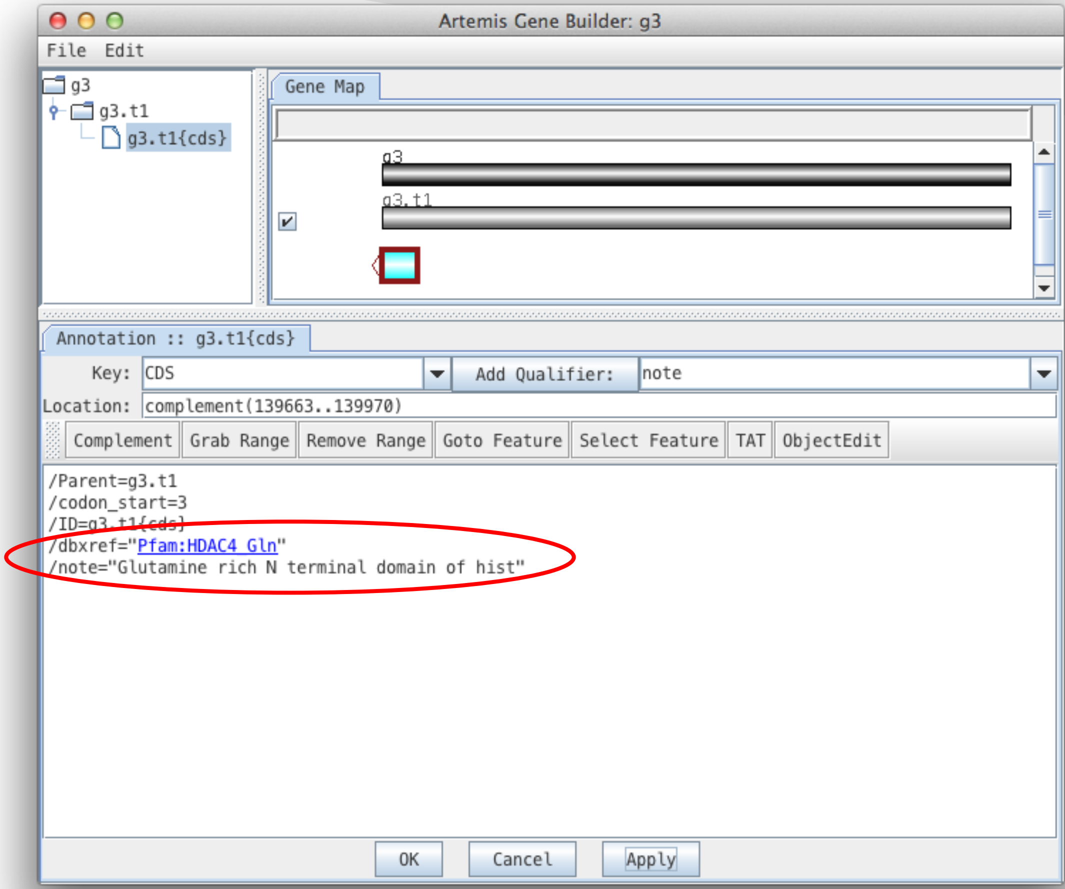
Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
HDAC4_Gln	Glutamine rich N terminal domain of hist ...	Domain	n/a	28	90	29	82	2	55	92	82.4	1.6e-23	n/a	Show

Comments or questions on the site? Send a mail to **pfam-help@sanger.ac.uk**. Our **cookie policy**.
The Wellcome Trust

Anotação de genes no Artemis

As sequências podem ser caracterizadas com diversas informações diferentes, e essas informações podem ser adicionadas diretamente pelo Artemis (assim como feito no BLAST).



Key: CDS

Add Qualifier: note

Location: complement(139663..139970)

```
/Parent=g3.t1
/codon_start=3
/ID=g3.t1{cds}
/dbxref="Pfam:HDAC4_Gln"
/note="Glutamine rich N terminal domain of hist"
```