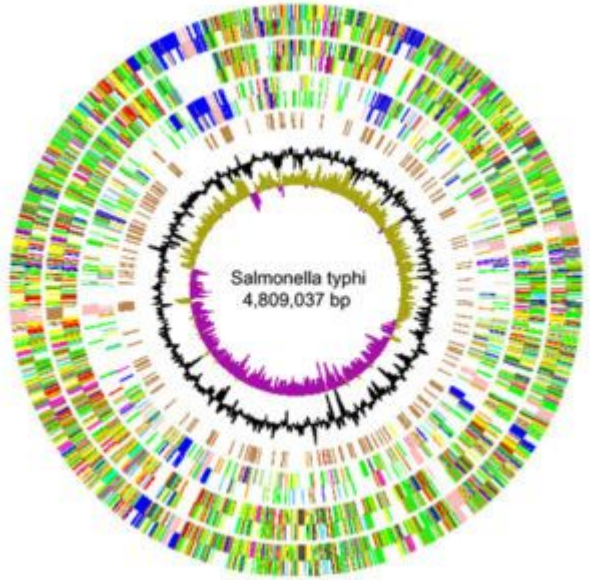
Module 1: Artemis



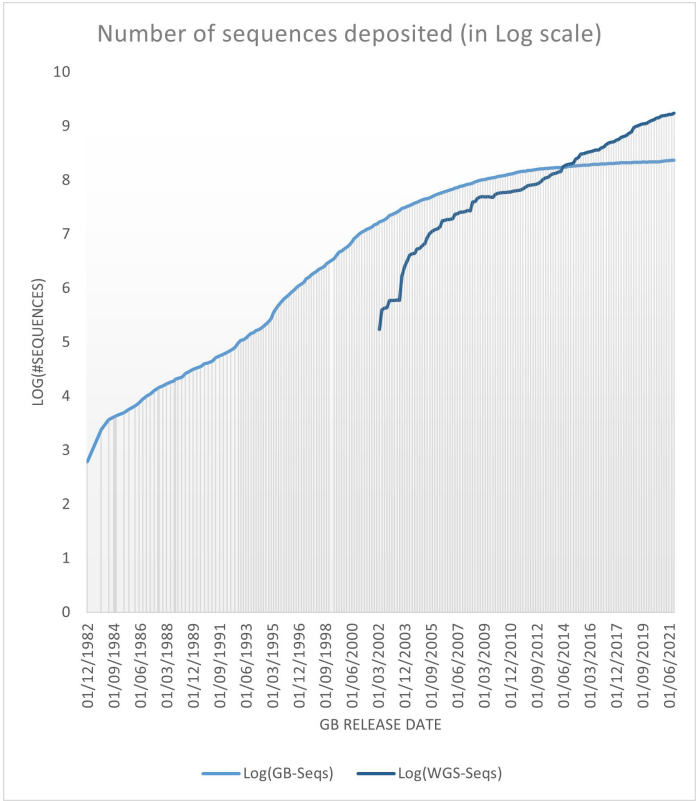
Lionel Uran Landaburu
Instituto de Investigaciones Biotecnológicas
lionel.u.l@iib.unsam.edu.ar

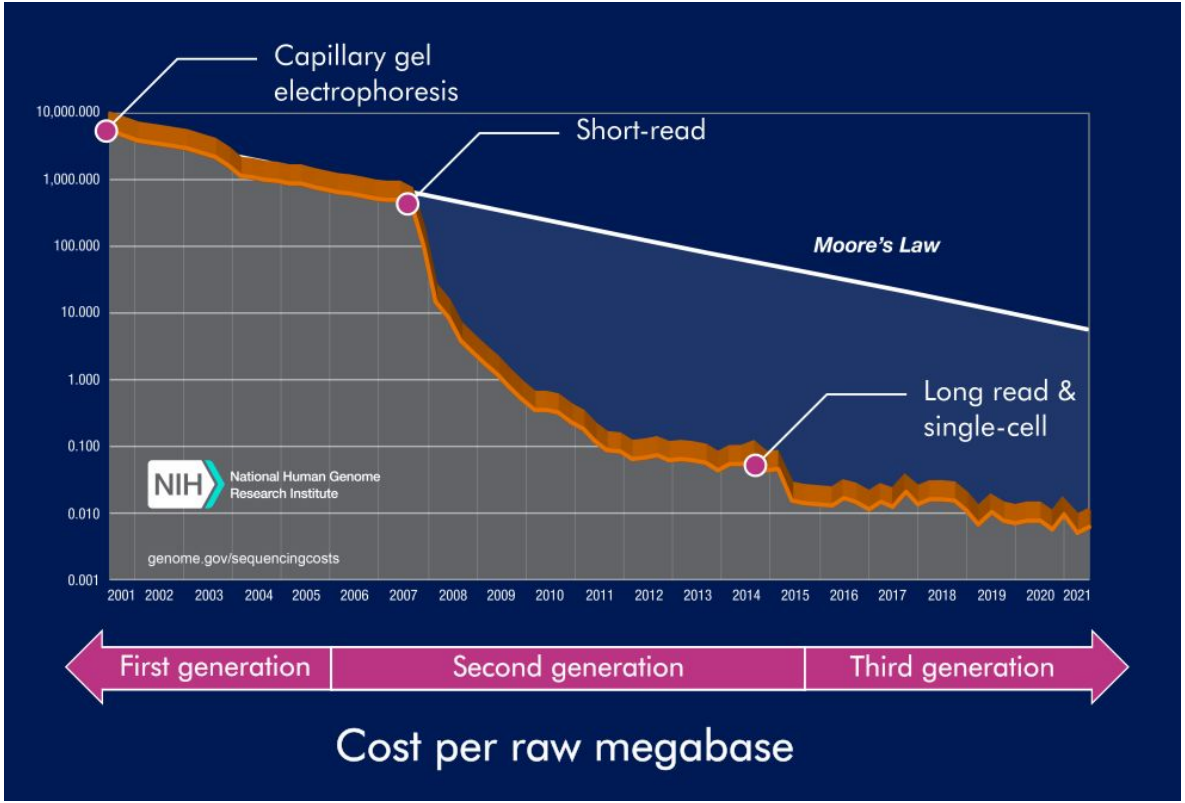
Rodrigo Quiroga
Instituto de Investigaciones Biotecnológicas
lionel.u.l@iib.unsam.edu.ar

Working with Pathogen Genomes 7-11th, February 2022.

wellcome connecting science

Next generation sequencing, a paradigm shift

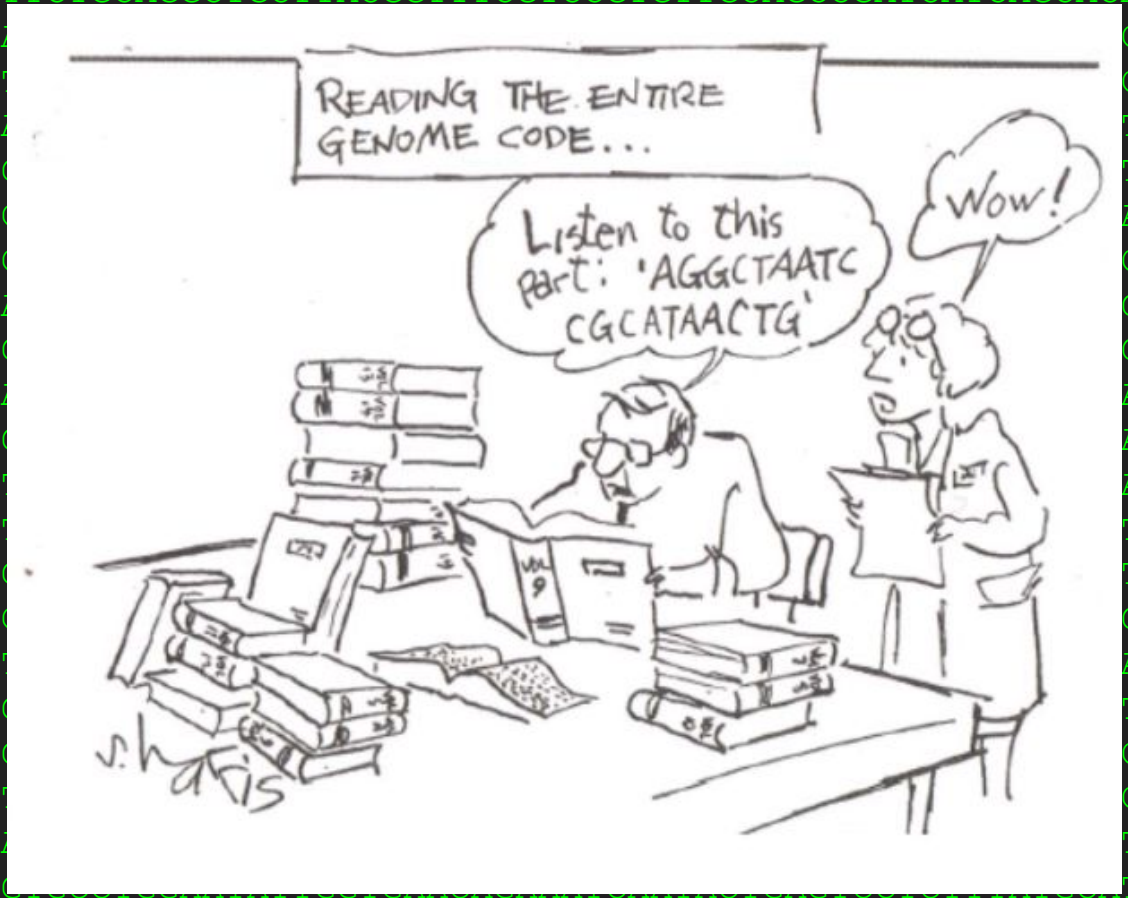




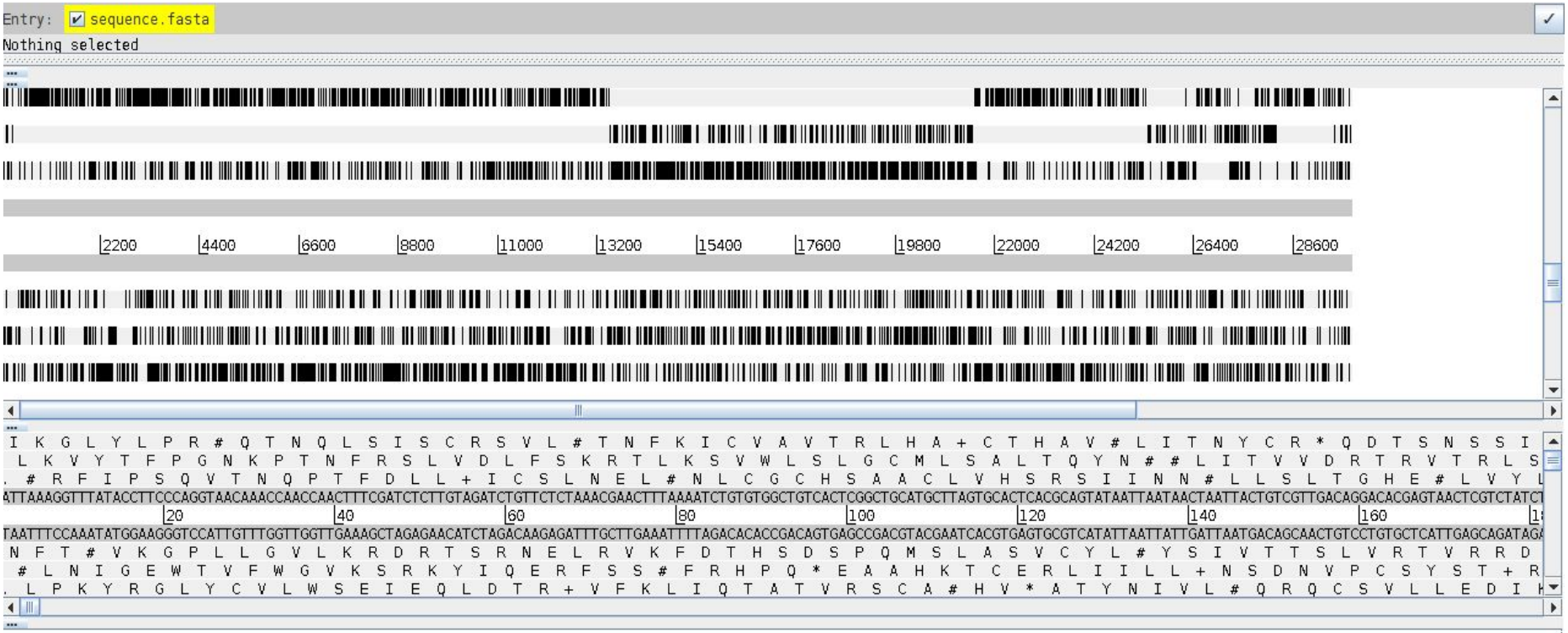
Source: GenBank statistics (https://www.ncbi.nlm.nih.gov/genbank/statistics/)

Adapted from genome.gov/sequencingcosts & Athina Gkazi (2021). An Overview of Next-Generation Sequencing. Technology Networks: Genomics Research.

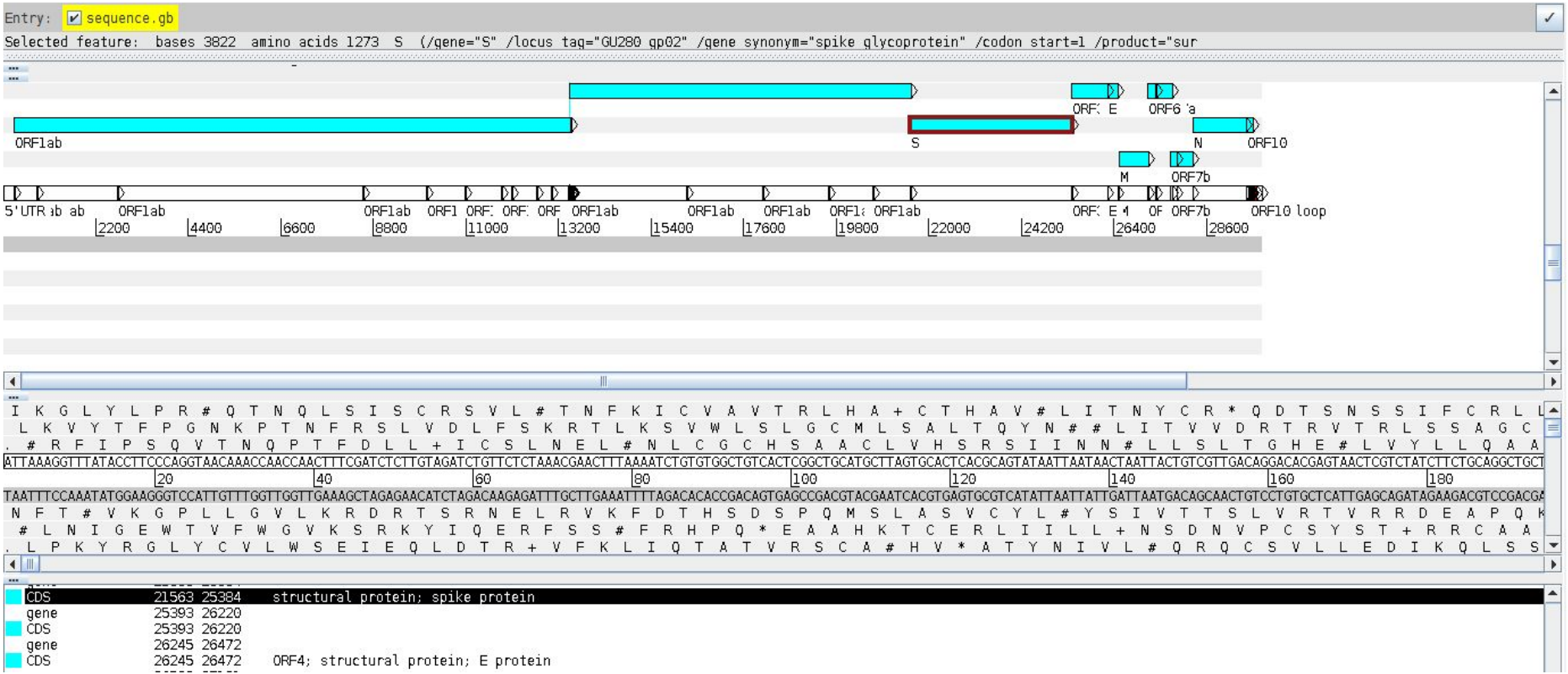
We have sequences… now what?



See anything yet?



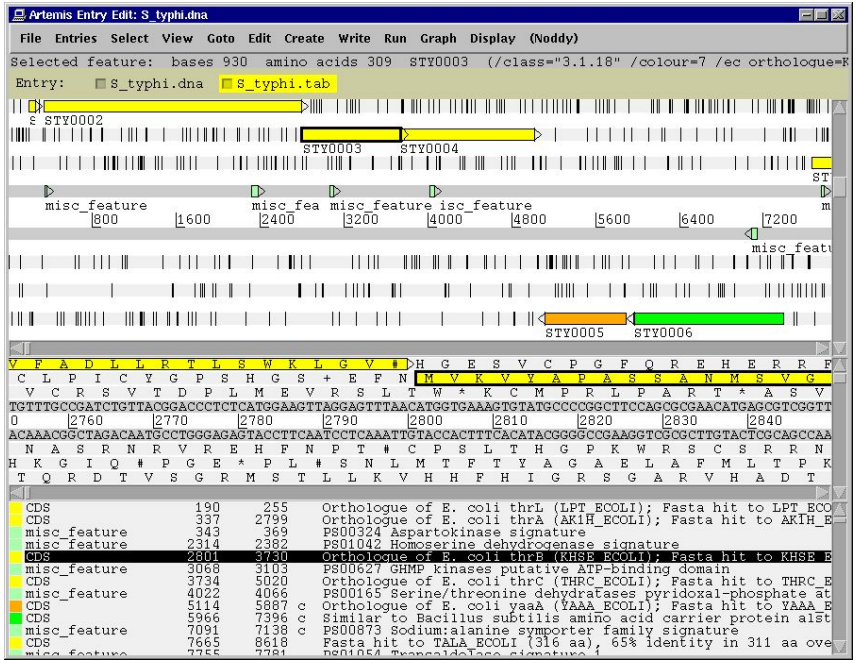
How about now? Does it ring a bell?



Genome browser and annotation tool

Artemis

- Visualization of sequence
  - DNA
  - six frame translation
  - Panoramic and sequence view
- Annotation
  - Features
  - Mapped and listed
  - Editable
  - In layers (entry)
- Perform and view analysis
  - basic analysis
  - Basic stats & index can be plotted
  - import and view the results of other searches/analysis
  - Different lines of evidence can be seen together



wellcome sanger institute

SCIENCE | Programmes | Groups | Collaborations | Tools | Data | Facilities

Home » Science » Tools and Software » Annotation

Artemis

Overview | Download | Learn | License | Contact | Related | Publications

Overview

Artemis is a free genome browser and annotation tool that allows visualisation of sequence features, next generation data and the results of analyses within the context of the sequence, and also its six-frame translation.

Download and Installation

Please see our GitHub page for download and installation instructions.

Learn and Support

For information and advice on using this software please see the our GitHub page.

In addition, an email discussion list called artemis-users is available and posts to the list since September 2001 are archived at mail-archive.com

Chado

Artemis and ACT can be used to connect to Chado databases. They are being developed to read from and write to the database and perform the same functions as the standard Artemis and ACT.

- Chado overview
- Chado admin
- Chado practical guide
- Data storage

Genome browser and annotation tool (Rutherford et al, 2000)

Free and regularly updated software

- (https://www.sanger.ac.uk/science/tools/artemis)

Developed in JAVA in 1998

Available for UNIX, macOS and Windows

Developed for annotation and analysis of prokaryotic and small eukaryotic genomes

Can be used simply as a sequence viewer allowing the visualization of sequence and annotation taken directly from EMBL or GeneBank.

Files in Artemis

Sequence	Annotation
.fasta	.tab
.seq	
.dna	
.embl	

It can read several file formats (FASTA, EMBL, GENEBANK, GFF).

EMBL

Two-character line code indicates the type of information contained in the line

Feature Key

Qualifier

```
ID   ECRSMA     standard; DNA; PRO; 500 BP.
XX
AC   L40173;
XX
SV   L40173.1
XX
DT   10-AUG-1995 (Rel. 44, Created)
DT   04-MAR-2000 (Rel. 63, Last updated, Version 4)
XX
DE   Erwinia carotovora repressor (rsmA) gene, complete cds.
XX
KW   repressor; rsmA gene.
XX
OS   Pectobacterium carotovorum
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriaceae;
OC   Pectobacterium.
XX
RN   [1]
RP   1-500
RA   Cui Y., Chatterjee A., Liu Y., Dumenyo C.K., Chatterjee A.K.;
RT   "Identification of a global repressor gene, rsmA, of Erwinia carotovora
RT   subsp. carotovora that controls extracellular enzymes,
RT   N-(3-oxohexanoyl)-L-homoserine lactone, and pathogenicity in soft-rotting
RT   Erwinia spp";
RL   J. Bacteriol. 177(17):0-0(1995).
XX
DR   GOA; Q47620; Q47620.
DR   SWISS-PROT; Q47620; CSRA_ERWCA.
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..500
FT                   /db_xref="taxon:554"
FT                   /organism="Pectobacterium carotovorum"
FT                   /strain="71"
FT                   /sub_species="carotovora"
FT   -10_signal      107..112
FT                   /gene="rsmA"
FT   RBS             235..239
FT                   /gene="rsmA"
FT   CDS             246..431
FT                   /codon_start=1
FT                   EIYQRIQAEKSQPTSY"
XX
SQ   Sequence 500 BP; 140 A; 101 C; 120 G; 139 T; 0 other;
     ggatccggca agcaggatag aaagtgtgtt accttcagat attctgaagc tttacatgct        60
     cagttctgtt gttgtgataa caaaagcaca agctactgat atcgactaaa ctaacaagta       120
     gtgacaaacc ggagtgtgat ggtgtggtta taccatcgtc taggtttacg ttttcacagc       180
     acatgatgga taatggcggg gagacagaga gacccgactc tttataatct ttcaaggagc       240
     aaagaatgct tattttgact cgtcgagttg gcgaaaccct catcatcggc gatgaggtaa       300
     cggttaccgt attaggagtg aaaggcaacc aggtgcgtat tggtgttaat gcacctaaag       360
     aggtttctgt ccaccgtgaa gagatctatc agcgtattca ggccgaaaaa tctcaaccaa       420
     cgtcatattg attgacaatg cgtctcgtgt tcgcgggacg caattgttat ttccggtttt       480
     tcccccacac atttatcgat                                                  500
//
```

Genbank

Header

Annotation

Sequence

```
LOCUS       ERWRSMA                 500 bp    DNA     linear   BCT 19-AUG-19
DEFINITION  Erwinia carotovora repressor (rsmA) gene, complete cds.
ACCESSION   L40173
VERSION     L40173.1  GI:927031
KEYWORDS    repressor; rsmA gene.
SOURCE      Pectobacterium carotovorum
  ORGANISM  Pectobacterium carotovorum
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriaceae;
            Pectobacterium.
REFERENCE   1  (bases 1 to 500)
  AUTHORS   Cui,Y., Chatterjee,A., Liu,Y., Dumenyo,C.K. and Chatterjee,A.K.
  TITLE     Identification of a global repressor gene, rsmA, of Erwinia
            carotovora subsp. carotovora that controls extracellular enzymes,
            N-(3-oxohexanoyl)-L-homoserine lactone, and pathogenicity in
            soft-rotting Erwinia spp
  JOURNAL   J. Bacteriol. 177(17) (1995) In press
COMMENT     Original source text: Erwinia carotovora (strain 71, sub_species
            carotovora) DNA.
FEATURES             Location/Qualifiers
     source          1..500
                     /organism="Pectobacterium carotovorum"
                     /strain="71"
                     /sub_species="carotovora"
                     /db_xref="taxon:554"
     gene            107..431
                     /gene="rsmA"
     -10_signal      107..112
                     /gene="rsmA"
     RBS             235..239
                     /gene="rsmA"
     CDS             246..431
                     /gene="rsmA"
                     /function="global repressor"
                     /note="putative"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA74502.1"
                     /db_xref="GI:927032"
                     /translation="MLILTRRVGETLIIGDEVTVTVLGVKGNQVRIGVNAPKEVSVHR
                     EEIYQRIQAEKSQPTSY"
BASE COUNT      140 a    101 c    120 g    139 t
ORIGIN
        1 ggatccggca agcaggatag aaagtgtgtt accttcagat attctgaagc tttacatgct
       61 cagttctgtt gttgtgataa caaaagcaca agctactgat atcgactaaa ctaacaagta
      121 gtgacaaacc ggagtgtgat ggtgtggtta taccatcgtc taggtttacg ttttcacagc
      181 acatgatgga taatggcggg gagacagaga gacccgactc tttataatct ttcaaggagc
      241 aaagaatgct tattttgact cgtcgagttg gcgaaaccct catcatcggc gatgaggtaa
      301 cggttaccgt attaggagtg aaaggcaacc aggtgcgtat tggtgttaat gcacctaaag
      361 aggtttctgt ccaccgtgaa gagatctatc agcgtattca ggccgaaaaa tctcaaccaa
      421 cgtcatattg attgacaatg cgtctcgtgt tcgcgggacg caattgttat ttccggtttt
      481 tcccccacac atttatcgat
//
```

Artemis panels & navigation

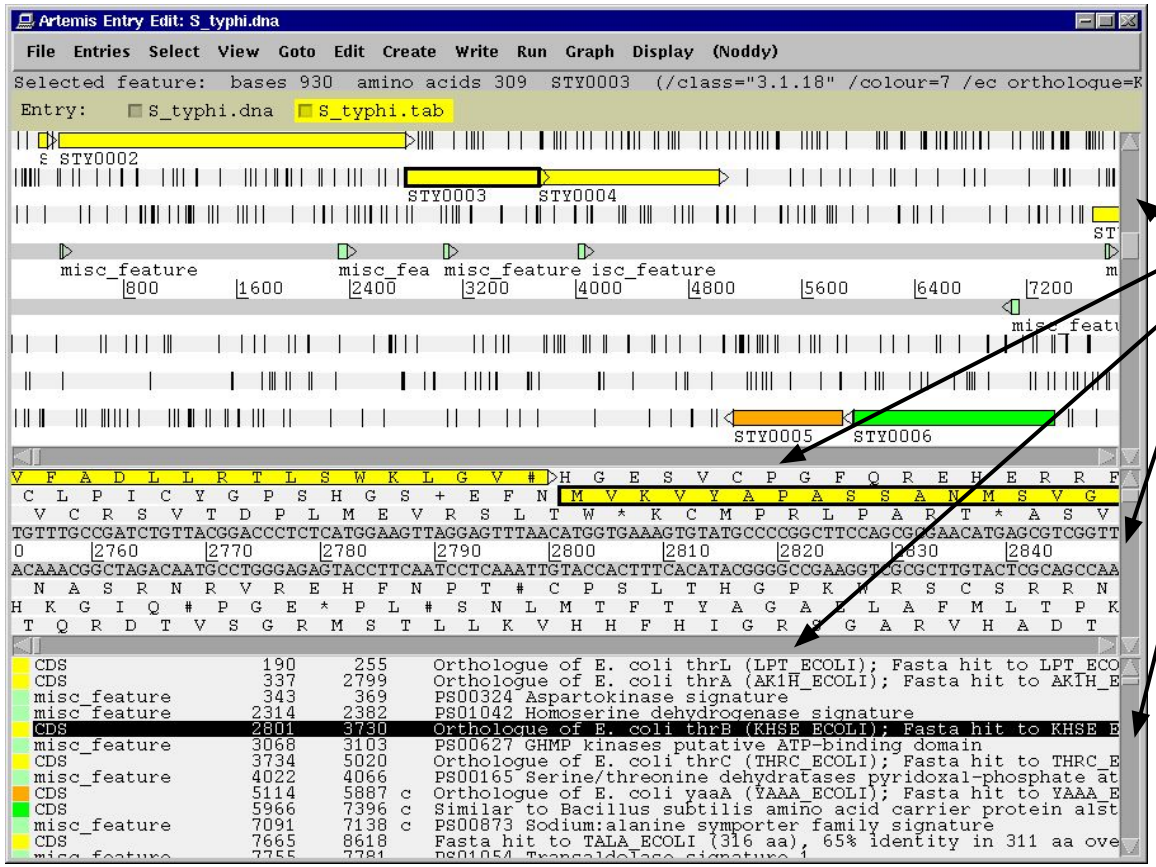
Drop Down Menus

Entry Button Line

Main Sequence View Panel

Magnified Sequence View Panel

Feature Menu



A word on genome browsers

	Artemis	GIVE	IGB	IGV	Jbrowse	Tablet	UCSC
LOCAL INSTALLATION							
Native app	●		●	●	●	●	
Web app		●●		●●	●●		●
Creation date	1999	2017	2001	2008	2009	2007	2014
Development status (2019)	stalled	early	stalled	mature/ early for web	active	mature	active
Software License	GPL3	Apache 2.0	Common Public License 1.0	MIT License	GNU LGPL v2.1	BSD-2 Clause	Copyright © 2001 UC Regents
PUBLIC WEB INSTANCE		●		●			●
Creation date		2017		2018			2000
Development status (2019)		early		early			active

BONARDI, Franck (2019) *Which genome browser to use for my data.* Bilille, Institut Français de Bioinformatique

Features of the Artemis module

Exercise 1: Familiarize with Artemis

- Loading sequences and annotation files
- Changing the view
- Searching and getting around

Exercise 2: Find insight within a genome

- Graphs and plots

Exercise 3: Analyze a genome

- Basic analysis
- Generating features

Exercise 4: Feature editing

- Adding and Modifying annotations
- Finding evidence : Database searches