CRISPR SYSTEM
Applications of CRISPR systems analysis in the context of metagenomics

Components of the CRISPR-Cas systems

(3-33)

(2-588 repeats)

Leader

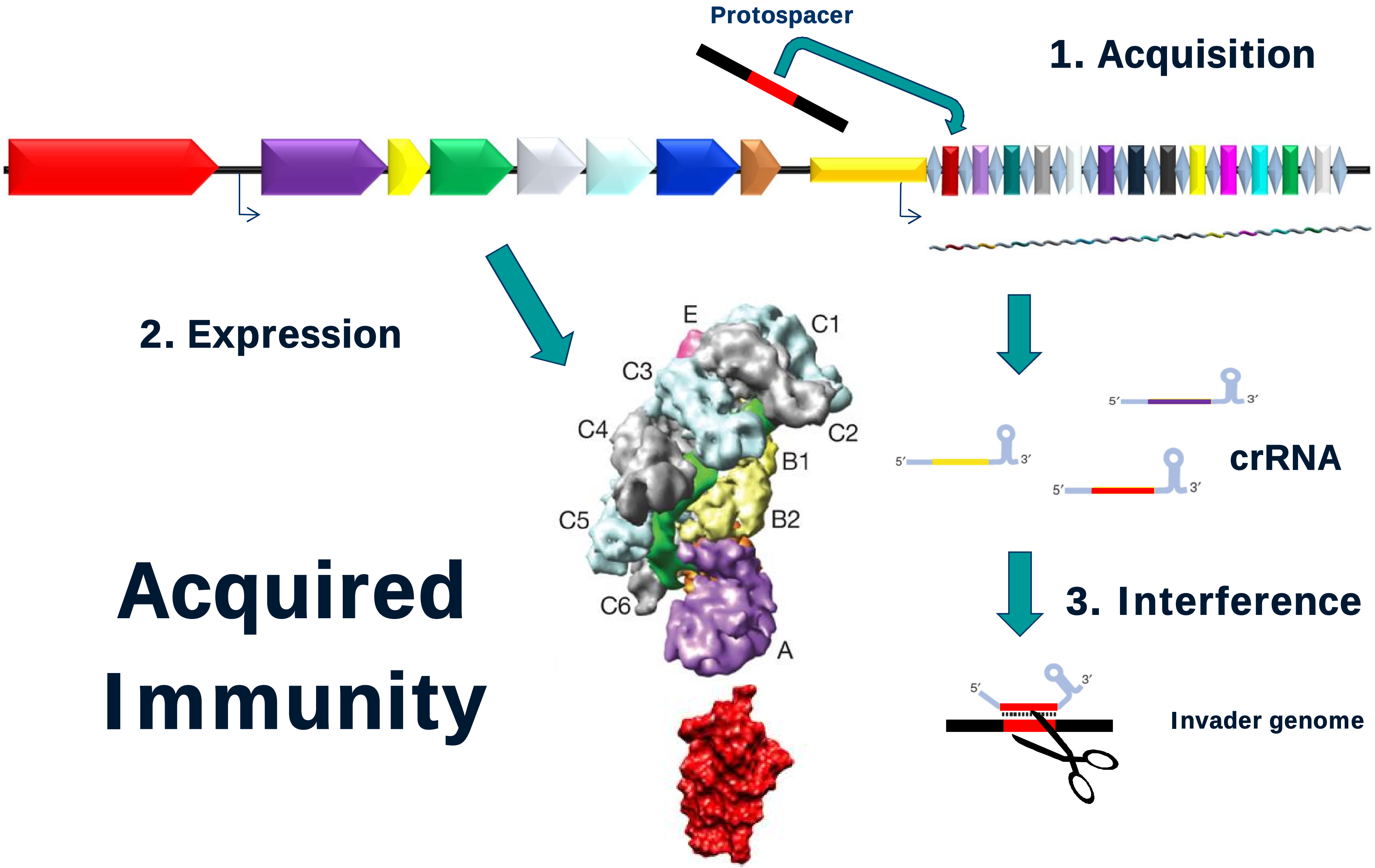
***cas* genes**

(*CRISPR associated*)

♦ **CRISPR**

Clustered Regularly Interspaced Short Palindromic Repeat

The CRISPR-Cas pathway



Protospacer

1. Acquisition

2. Expression

crRNA

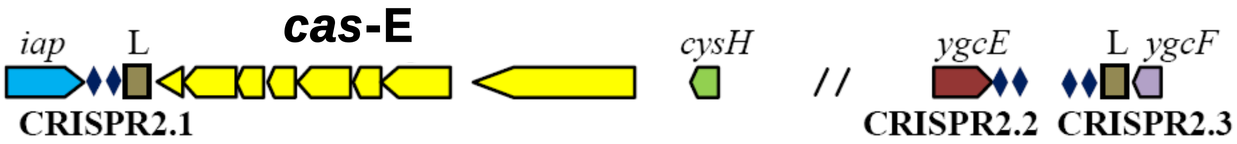
3. Interference

Invader genome

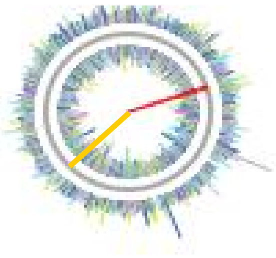
Acquired Immunity

Multiple systems in a genome

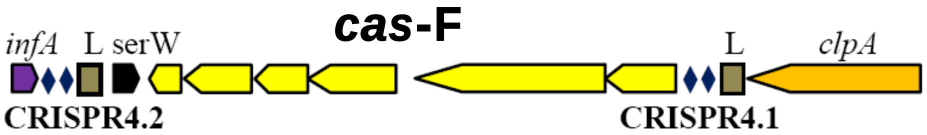
CRISPR-Cas I-E



CGGTTTATCCCCGCTGGCGCGGGGAAACWC



CRISPR-Cas I-F

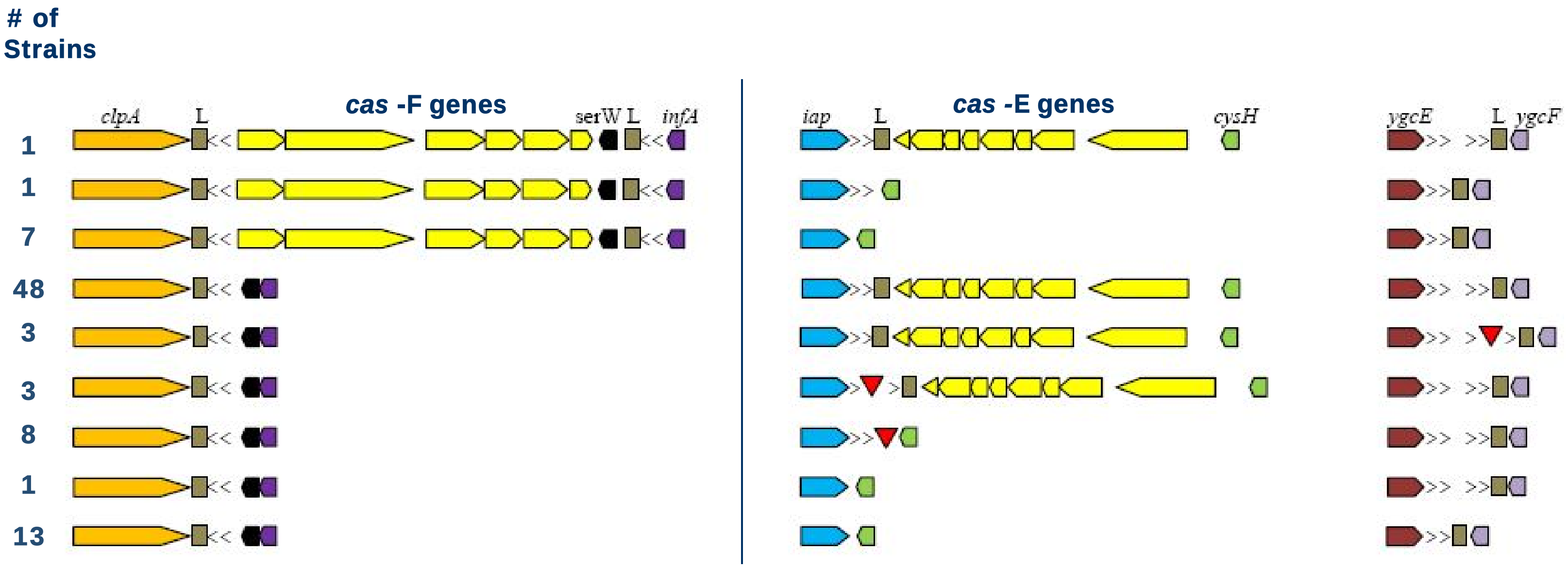


TTTCTAAGCTGCCTGTACGGCAGTGAAC

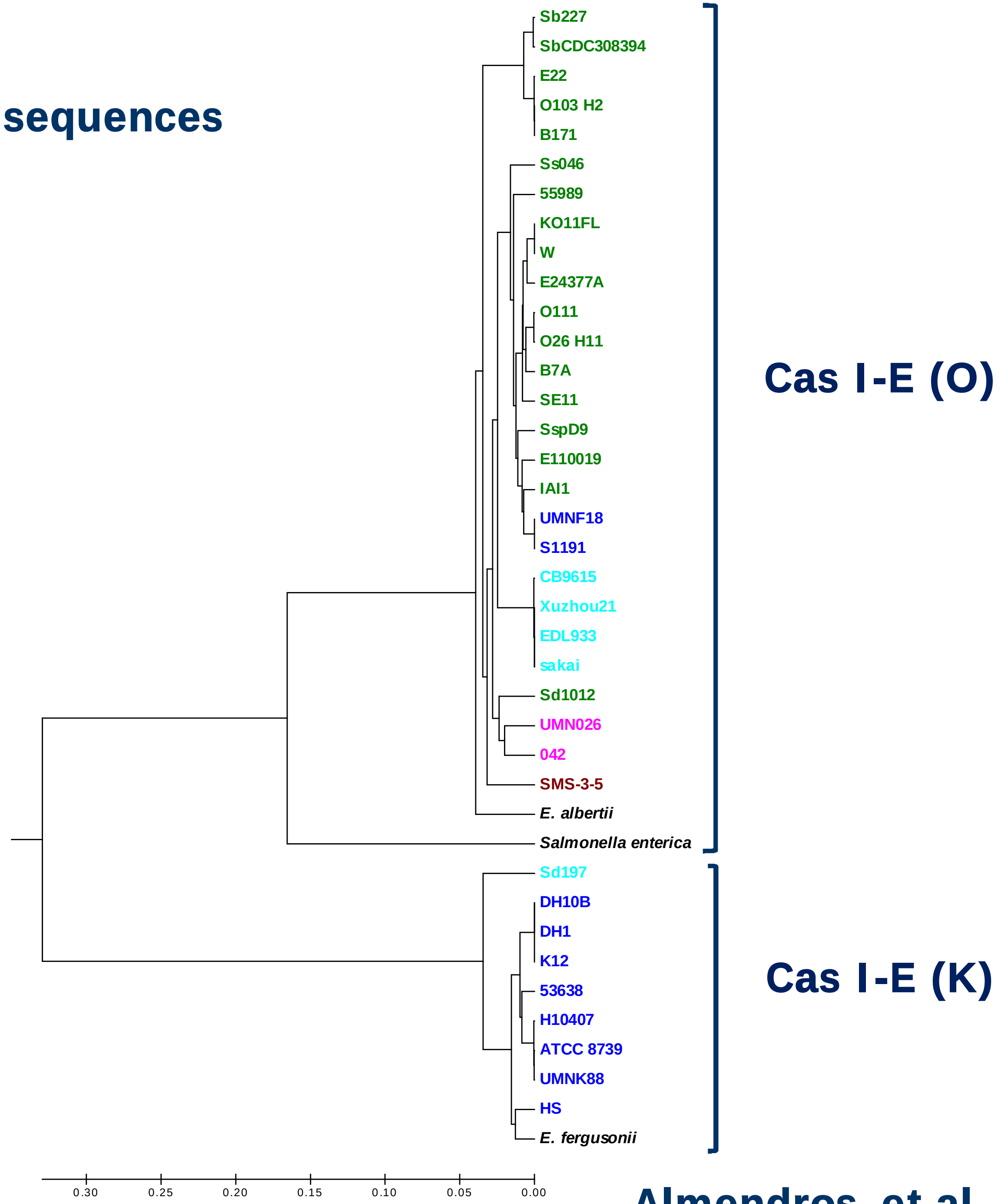
Intraspecific diversity

CRISPR-Cas I-F

CRISPR-Cas I-E

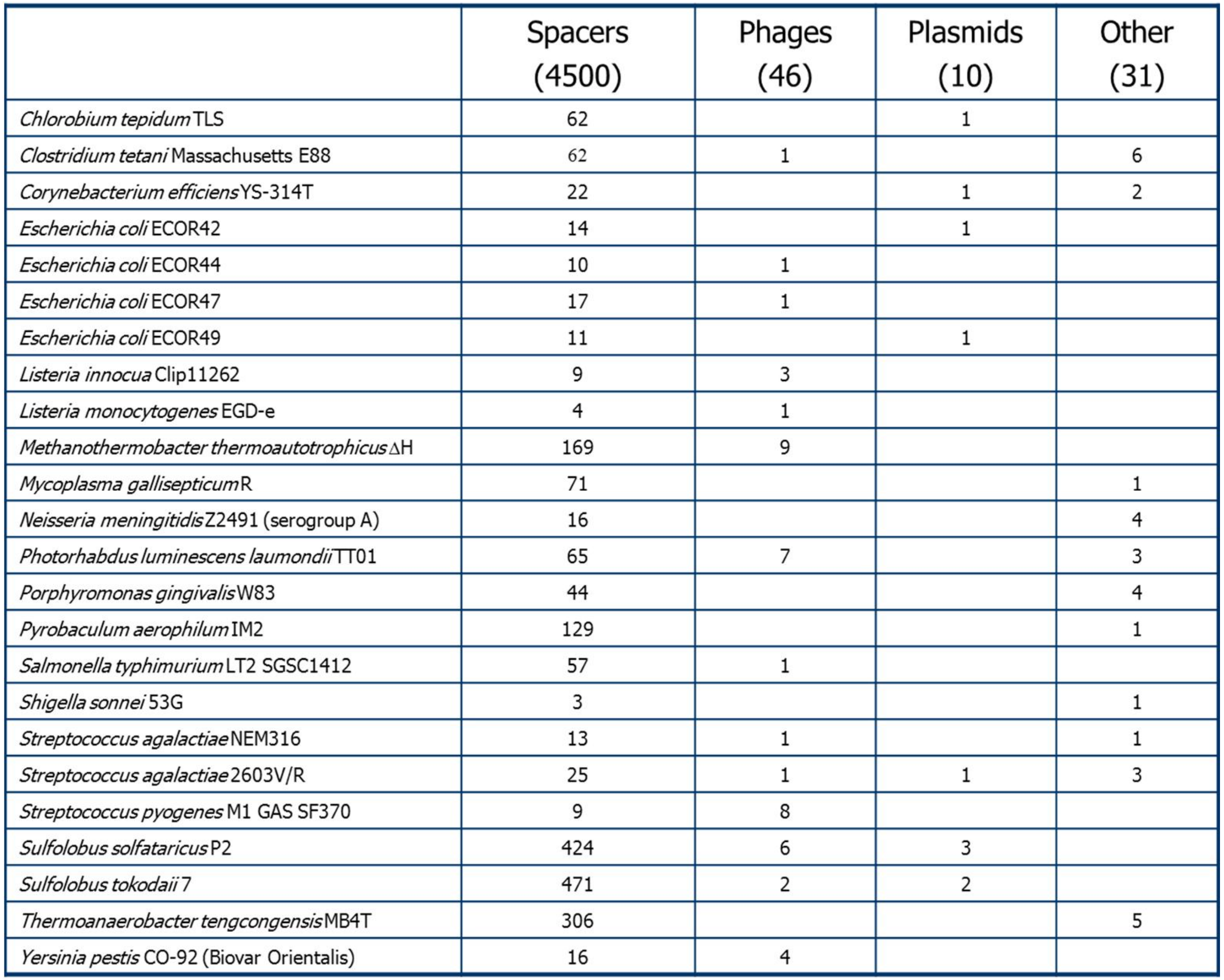


Cas I-E sequences



Almendros et al. *mBIO.* In press

Origin of Protospacers



	Spacers (4500)	Phages (46)	Plasmids (10)	Other (31)
Chlorobium tepidum TLS	62		1	
Clostridium tetani Massachusetts E88	62	1		6
Corynebacterium efficiens YS-314T	22		1	2
Escherichia coli ECOR42	14		1	
Escherichia coli ECOR44	10	1		
Escherichia coli ECOR47	17	1		
Escherichia coli ECOR49	11		1	
Listeria innocua Clip11262	9	3		
Listeria monocytogenes EGD-e	4	1		
Methanothermobacter thermoautotrophicus ΔH	169	9		
Mycoplasma gallisepticum R	71			1
Neisseria meningitidis Z2491 (serogroup A)	16			4
Photorhabdus luminescens laumondii TT01	65	7		3
Porphyromonas gingivalis W83	44			4
Pyrobaculum aerophilum IM2	129			1
Salmonella typhimurium LT2 SGSC1412	57	1		
Shigella sonnei 53G	3			1
Streptococcus agalactiae NEM316	13	1		1
Streptococcus agalactiae 2603V/R	25	1	1	3
Streptococcus pyogenes M1 GAS SF370	9	8		
Sulfolobus solfataricus P2	424	6	3	
Sulfolobus tokodaii 7	471	2	2	
Thermoanaerobacter tengcongensis MB4T	306			5
Yersinia pestis CO-92 (Biovar Orientalis)	16	4		

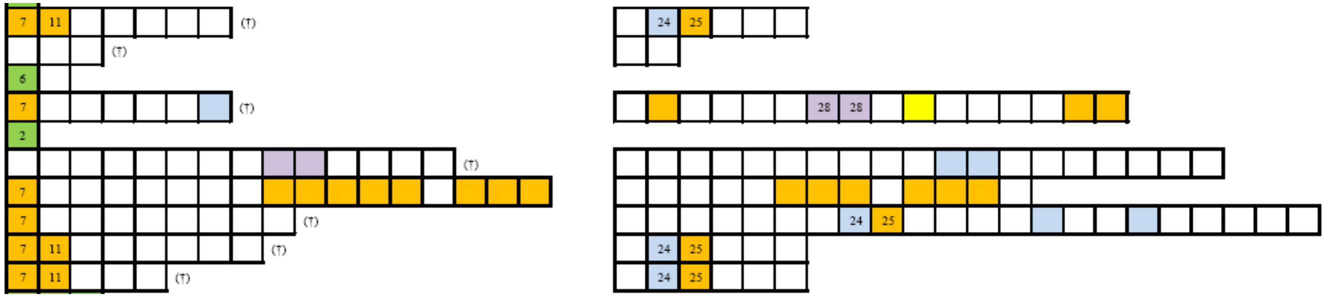
Mojica et al. *J. Mol. Evol.* 2005

Origin of Protospacers

LEGEND

E24377-SE11 Phage	🟪
Other Phages	🟦
Plasmid	🟧
Non-Mobile	🟨
CAS Gene	🟩

CRISPR-Cas I-F

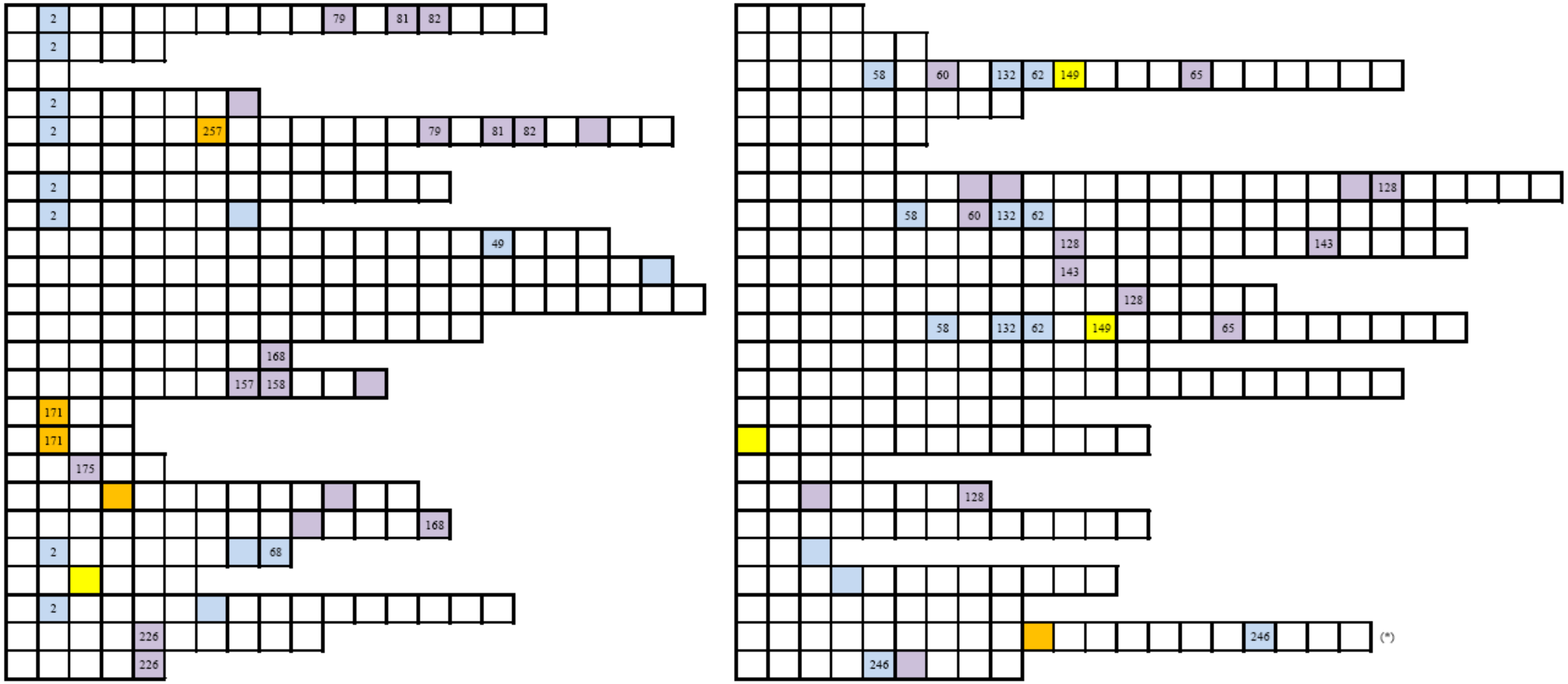


Origin of Protospacers

CRISPR-Cas I-E

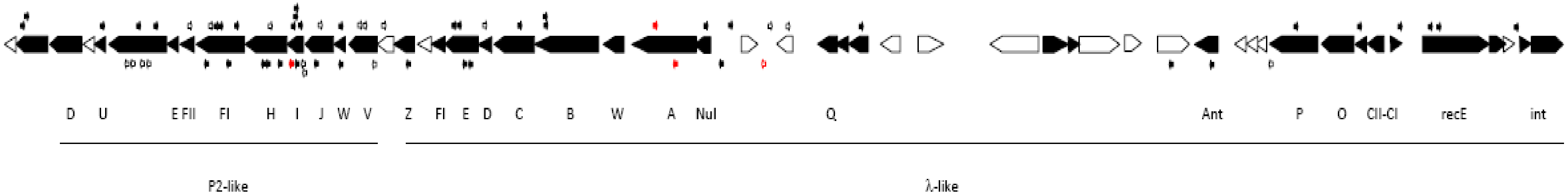
LEGEND

E24377-SE11 Phage	🟪
Other Phages	🟦
Plasmid	🟧
Non-Mobile	🟨
CAS Gene	🟩



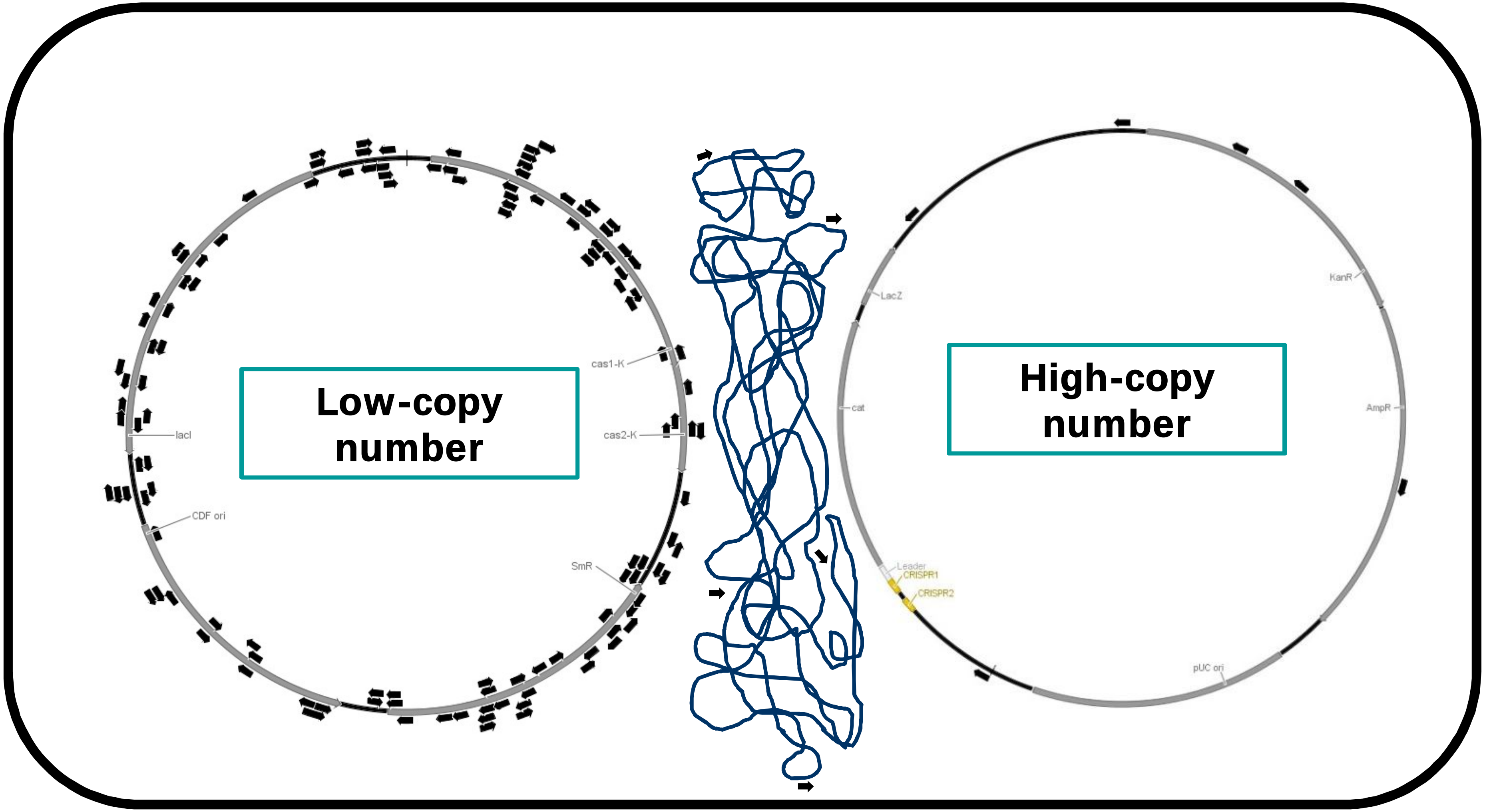
Incidence of spacer donors

Genetic Element	> 90% Identity	% Spacers (950)
Enterophages	67	7%
Prophage SE11-6	**42**	**63%**
Other phages	25	37%
Plasmids	29	3%
Other	16	2%
Total	**112**	**12%**



SE11-6

Preference for spacer donors



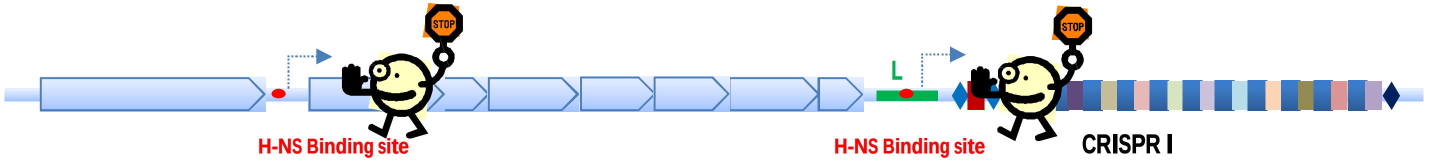
Low-copy number

High-copy number

Díez-Villaseñor et al. *RNA Biology* 2013

Activity

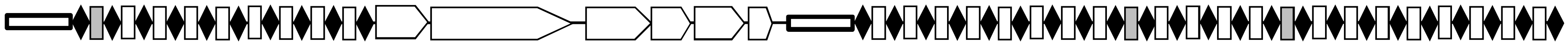
Silencing of I-E



The on-off switch of CRISPR immunity against phages in *E. coli* . Mojica and Díez-Villaseñor. *Mol Microbiol*. 2010

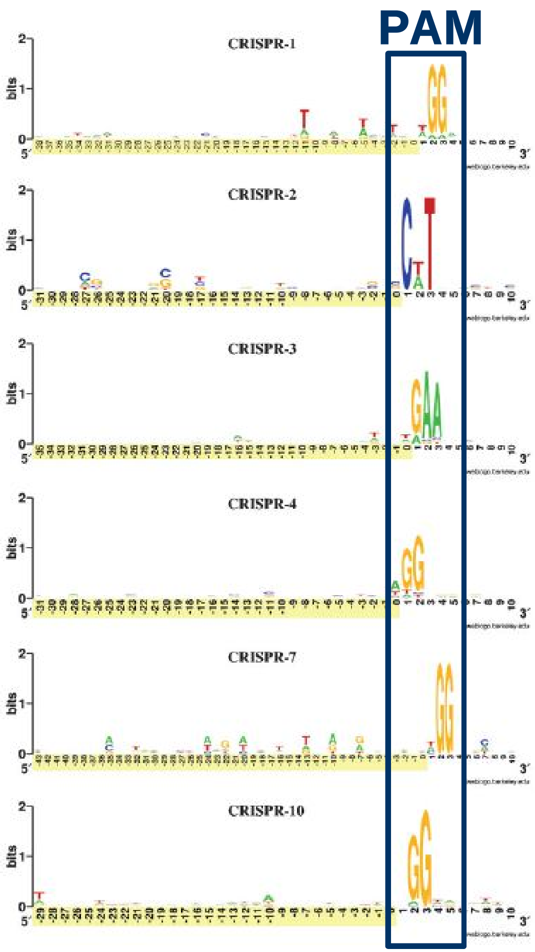
Constitutive expression of I-F

CRISPR4.1 array ***cas*-F genes** **CRISPR4.2 array**



Target motifs affecting natural immunity by a constitutive CRISPR-Cas system in *E. coli*
Almendros, *et al*. *PLoS ONE* 2012

The Protospacer Motifs

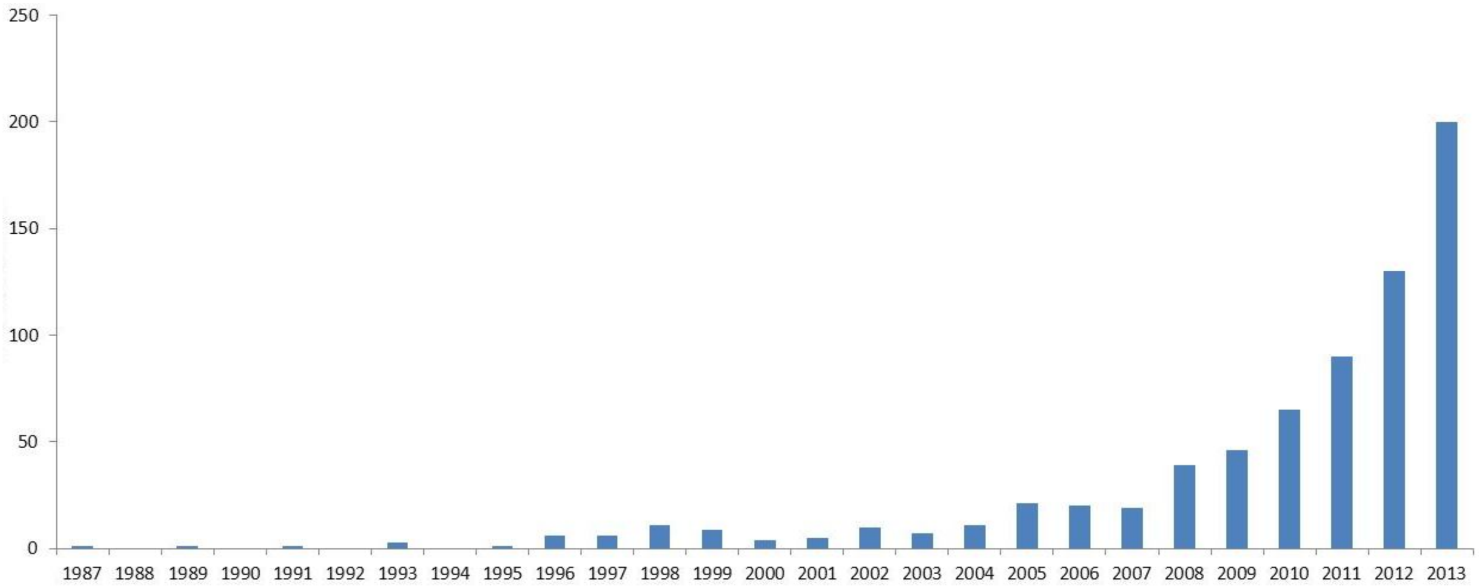


Mojica et al. *Microbiology*. 2009

Required for **S**pacer **A**cquisition: **SAM**

Recognized during **T**arget **I**nterference: **TIM**

CRISPR-Cas publications



Applications

- Molecular markers for Typing
    - Epidemiology
- Generation of virus-resistant strains
    - Targeting specific viruses
- Increase chromosomal stability and integrity
    - Targeting transposons
- Reduce the harmful potential of pathogens
    - Vaccination against antibiotic resistance uptake
- Regulation of gene expression
    - RNA-targeting
- Genome engineering: specific programmable cleavage of genomic DNA
    - Gene disruption, replacement.
- Insights into microbial population diversity and dynamics
    - Metagenomes

Metagenome/CRISPR analysis

- **Diversity of CRISPR-Cas elements**

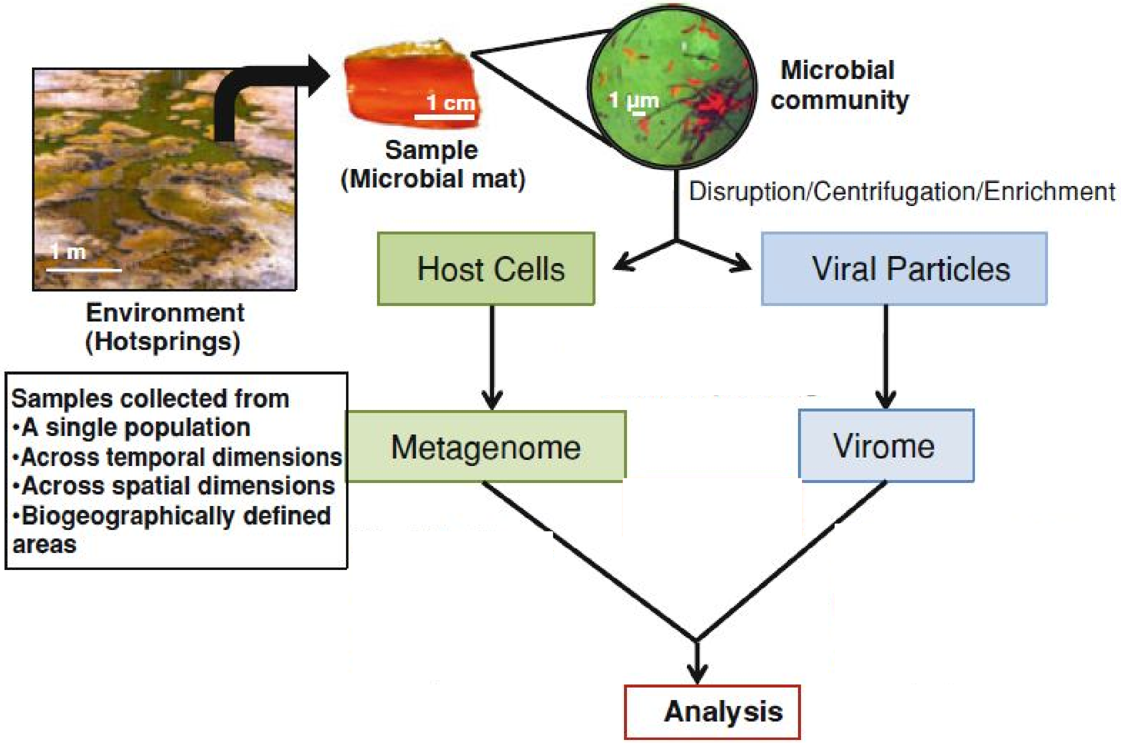
- **Fast evolution of spacer content**
  - **Snapshots of invaders**
  - **Chronological record**
  - **Adaptation to resistant mutants**

Metagenomic Information

Host and transmisible elements diversity

Host-virus interactions

General process



Environment
(Hotsprings)

Sample
(Microbial mat)

Microbial
community

Disruption/Centrifugation/Enrichment

Host Cells

Viral Particles

Metagenome

Virome

Samples collected from
- A single population
- Across temporal dimensions
- Across spatial dimensions
- Biogeographically defined areas

Analysis

Identification of
- **Cas**
- **CRISPR**
- **Spacers**
- **Protospacers**
- **PAMs**

CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats.

Grissa et al. Nucleic Acids Res. 2007

Friday November 29th 2013

CRISPRfinder program online

Welcome to the CRISPRs web service. This page enables the easy detection of CRISPRs in user-submitted sequence data (allows sequences up to 67,000,000 bp).

Input format :

Your data must be a DNA sequence (or many DNA sequences) in FASTA format.
Download sample1 Aquifex VF5 or sample2 P. aeruginosa PA14 or samples YP1 Yersinia

Submit your Own data

Enter sequence (DNA) below using copy and paste:

Or upload a local file: No se ha seleccionado ningún archivo.

Data Summary

	Genomes analysed	CRISPRs found (*)
Archea	150	657(125)
Bacteria	2480	3255(1126)
Total	2630	3912(1251)

*number of convincing CRISPR structures (number of genomes with such CRISPR)

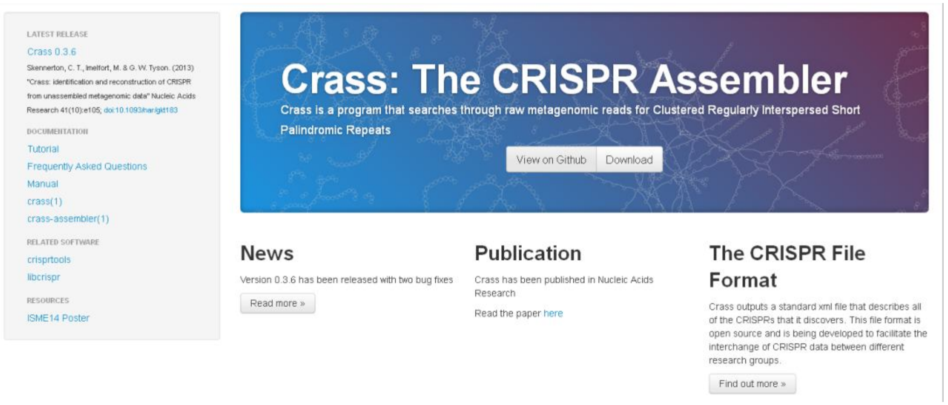
Database status:
Last update : 2013-09-18

Contact:
Christine POURCEL

```
757015  AATGGGGGCTCCACGAACGGCTTGAGGGGGGTGCAAC     GGGATGGGTGCGCTGGTGCGCCAACACGGATGGCAA
757088  GTTTCAGTATCCACGAACGGATCGAGGGGGATGCAAC     TCGAAGAACCGTCATTTATTCTTGAAATTGCGTTTTG
757162  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     ATCCTGGTAACTTTCCACGAAGAGATTAACCGCGAA
757235  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     GACGAACAGCATCCTATCCTGCAGTACATGATAAAT
757308  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     CCGGTGATGACCAGCATAAAAAACAGCGCGCCAGCGAT
757383  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     ACCGCAGCCGCATCCCCCCGATTCGTTACCGGATAG
757456  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     AGAATGTATTCAATCGTGAACGAATTGTTTTCTTCT
757529  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     CCTATTAGGGAGGCAGCGGGTGGCCTGCCCCGAAAG
757602  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     TTTGAAGTACCCATTTTCGGGCGCGTGAATCTTTACA
757676  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     TTTGAAGTACCCATTTTCGGGCGCGTGAATCTTTACA
757750  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     CATCCACTCGCCCACAGGTGCTTCTCAGAGCACCTG
757823  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     GCACCTGTCCTTATTTTAGCCGGTTTTCAGCCATTCTT
757898  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     AAAATACGGGCGCAGGACTTCTTCAACGCGCCACAC
757971  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     AAATACAGGCGCAGGACTTCTTCAACGCGCCACAC
758043  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     GAATGGCGAGAGTCTATGATAGTCACGGGTTTACTA
758116  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     GGAGTAATTTCAAACCCCTTGATTGTCCAGTATGA
758188  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     CCACAAGTGCTTAAGTACCTTGTTGCCGAGGGCGAA
758261  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     TGGGGGTTTGGTGCCGCATGCCCGTTTGAGTGGCCGTA
758336  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     ATCAGCAGAAACGAAATAGAAGATTTGTTCAAGTCGT
758410  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     TAAACCAGGCATGCCAAAGTGTAAATGATAATGGTT
758483  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     AAAGTTGACCAGTGCAAAGGTGAGAACCAGCACGAC
758556  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     CGTTGAGCCGCCGTCCAGCGAGAGTTGCGCTTCGGT
758629  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     AAATTCTTGCTGGAAGTACTTATGCTTTTGCGGCTGAA
758704  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     TGCGTCTGGCGGCGATCCGCTCTCTCTTGCGCCACCT
758778  GTTTCAGTATCCACTAACGGATCGAGGGGGATGCAAC     TCATCGGCGGGTGACTTCATCAAGGCGTTTGTCGAG
```

Crass: identification and reconstruction of CRISPR from unassembled metagenomic data.

Skennerton et al. Nucleic Acids Res. 2013



Crass 0.3.6

Skennerton, C. T., Imelfort, M. & G. W. Tyson. (2013) "Crass: identification and reconstruction of CRISPR from unassembled metagenomic data" Nucleic Acids Research 41(10):e105; doi:10.1093/nar/gkt183

Crass: The CRISPR Assembler

Crass is a program that searches through raw metagenomic reads for Clustered Regularly Interspersed Short Palindromic Repeats

News

Version 0.3.6 has been released with two bug fixes

Publication

Crass has been published in Nucleic Acids Research

Read the paper here

The CRISPR File Format

Crass outputs a standard xml file that describes all of the CRISPRs that it discovers. This file format is open source and is being developed to facilitate the interchange of CRISPR data between different research groups.

CRISPR and Metagenomics

- Biogeographic and temporal diversity: “Strains from different locations are distinct” “Hosts and viruses evolve rapidly”
  - Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. Tyson GW, Banfield JF. Environ Microbiol. 2008
  - Diverse CRISPRs evolving in human microbiomes. Rho M, Wu YW et al. PLoS Genet. 2012
  - Analysis of streptococcal CRISPRs from human saliva reveals substantial sequence diversity within and between subjects over time. Pride DT et al. . Genome Res. 2011
- Evidence for epidemic infections by a single virus: “independently acquired spacers from a virus”
  - CRISPR associated diversity within a population of Sulfolobus islandicus. Held NL et al. PLoS One. 2010
  - Virus population dynamics and acquired virus resistance in natural microbial communities. Andersson AF, Banfield JF. Science. 2008
- Identification of novel mobile genetic elements
  - Metagenomic analyses of novel viruses and plasmids from a cultured environmental sample of hyperthermophilic neutrophiles. Garrett RA et al. . Environ Microbiol. 2010
  - Use of cellular CRISPR (clusters of regularly interspaced short palindromic repeats) spacer-based microarrays for detection of viruses in environmental samples. Snyder et al. Appl. Environ. Microbiol. 2010
- Identify virus-host pairs
  - Using CRISPRs as a metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. Anderson RE et al. FEMS Microbiol Ecol. 2011
  - Reconstructing viral genomes from the environment using fosmid clones: the case of haloviruses. García-Heredia et al. PLoS One. 2012
- Virus-host interactions
  - Germ warfare in a microbial mat community: CRISPRs provide insights into the co-evolution of host and viral genomes. Heidelberg JF, et al. PLoS One. 2009
  - Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. Tyson GW, Banfield JF. Environ Microbiol. 2008
  - Evolutionary dynamics of clustered irregularly interspaced short palindromic repeat systems in the ocean metagenome. Sorokin VA et al. Appl Environ Microbiol. 2010

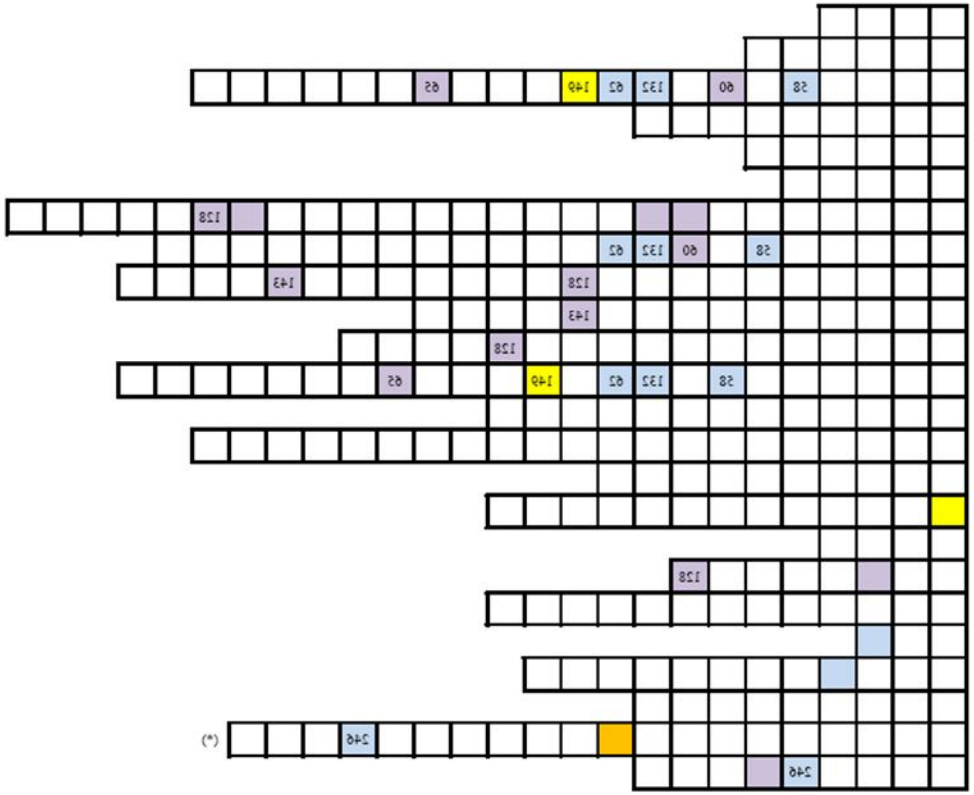
“Low-abundance viruses are targeted by conserved spacers”

↓

“They are more stable than (highly dynamic) higher abundance viruses”

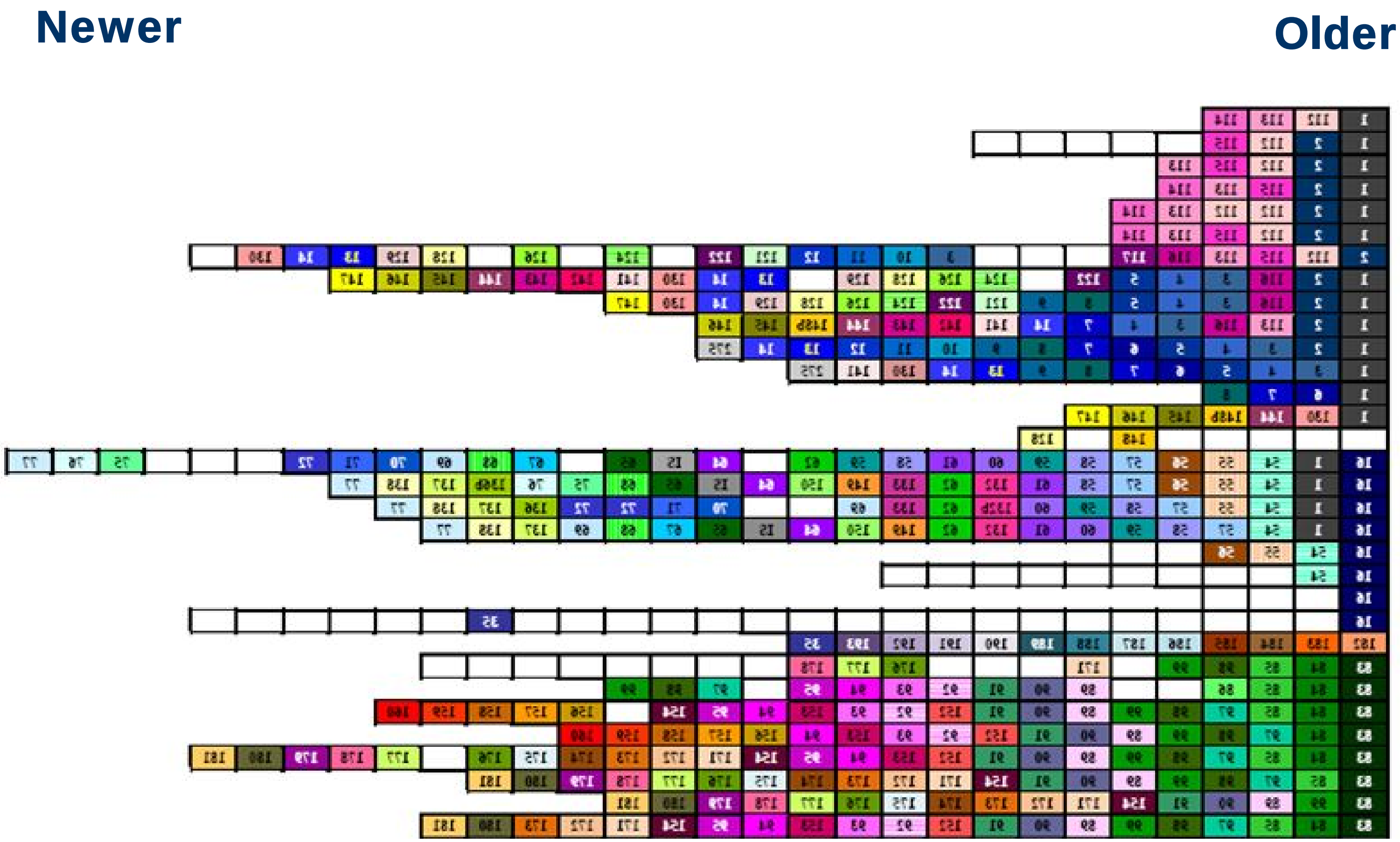
Virus-host and CRISPR dynamics in Archaea-dominated hypersaline Lake Tyrrell, Victoria, Australia.

Emerson et al. *Archaea*, 2013



Newer

Older


Newer
Older

CONCLUDING REMARKS

- CRISPR reveal relevant features about host and virus communities.
- They have contributed to shed light on host–viral interactions.
- These studies will be enhanced by metagenomic and virome data:
  - from a variety of environments
  - across spatial scales
  - over time
- Environments with low microbial complexity are ideal.
- The missive that CRISPR-Cas are a defence system against viruses is an oversimplification.