

Insyght: Using symbols to visualize homologies, conserved syntenies and genomic insertions across multiple genomes

Thomas Lacroix¹, Valentin Loux¹, Annie Gendrault¹, Mark Hoebeke² and Jean-François Gibrat¹

¹ Mathématique, Informatique et Génome, INRA, Jouy-en-Josas, 78352 France {thomas.lacroix, valentin.loux, annie.gendrault, jean-francois.gibrat}@jouy.inra.fr

² CNRS, UPMC, FR2424, ABIIMS, Station Biologique, 29680, Roscoff, France mark.hoebeke@sb-roscoff.fr

Insyght proposes a new way to explore the landscape of conserved and idiosyncratic genomic regions across multiple genomes and their rearrangements throughout evolution. Its unique display consists of a symbolic representation tightly integrated with a proportional view. Insyght is suitable for a variety of analyses: genome-wide inference of gene function, detection of evolutionary events, phylogenetic profiling and investigation of the core genome or niche-specific genes.

Browse public dataset: 389 prokaryotes genomes to date

Alternatively, you can use a virtual machine to compare your own genomes (private dataset). The VM includes :

- the database
- the pipeline to process the data: blast, bi-directional best hit for orthology, dynamic programming for synteny
- the visualization tool

Select a reference organism

Visualize the genomic organization with a view that combines a symbolic and a proportional representation. The symbols highlight a region of interest and provide legibility while the proportional view simultaneously allows grasping genomic locations and complex rearrangements scattered across the genomes and occurring at different scales.

Choose your gene set and analyze the presence, absence, or multiple copies of homologs. A functionality based on filters facilitates the retrieval of genes of interest and allows the formulation of relevant biological questions, such as finding niche-specific or core genome genes that match a few particular functions or biological processes.

A typical genomic organisation view :

Comparison windows (full size)
Reference genome (hidden)
Navigate results

Detailed information and display options
Selection and contextual menu

The filter functionality :

Filter genes by:
Presence / absence homology
presence
of homologs in organism(s) :
Aeropyrum pernix strain K1 [taxonid=272557]
Function
containing
binding
Add a new filter

Filtered genes list:
- dnaA (292-1632)
- dnaN1 (1811-2950)
- BCA_0003 (3078-3290)
- recF (3303-4430)
- gyrB (4469-6391)
- gyrA (6480-8951)
- BCA_0013 (14239-15240)
- guaB (15356-16819)
- dacA (16924-18243)
- BCA_0016 (18404-19291)
- BCA_0017 (19310-19900)
- serS (20228-21502)
- BCA_0020 (21918-22310)
- BCA_0021 (22346-23014)

The homolog table view :

#223 (score=9) Burkholderia cenocepacia strain AU 1054 [taxonid=331271]
#224 (score=9) Burkholderia mallei strain ATCC 23344 [taxonid=243160]
#225 (score=9) Burkholderia mallei strain NCTC 10247 [taxonid=320389]
#226 (score=9) Cyanobacteria sp. strain ATCC 29155 = PCC 7424 = PCC 7424 [taxonid=65393]
#227 (score=9) Corynebacterium efficiens strain DSM 44549 = NBRC 100395 = JCM 11189 = YS-314 = AJ 12310 [taxonid=196164]

Zoom on the symbolic (A) and proportionnal (B) representation :

#8 (score=6357) Bacillus cereus strain 03BB102 [taxonid=572264]
15 genes
63 bp
2 genes
1226 bp
B. anthracis A0248
B. cereus 03BB102

Go back and forth between the 2 views



Learn more at

<http://genome.jouy.inra.fr/Insyght/>