



**HAL**  
open science

## **IGO, Integration from Genomes to Organisms**

Sandra S. Derozier, Helene H. Chiapello, Thomas Lacroix, Cyprien Guerin Guérin, Valentin Loux, Anne Goelzer, Jean-François Gibrat, Franck F. Samson, Annie Gendrault-Jacquemard Gendrault

► **To cite this version:**

Sandra S. Derozier, Helene H. Chiapello, Thomas Lacroix, Cyprien Guerin Guérin, Valentin Loux, et al.. IGO, Integration from Genomes to Organisms. JOBIM 2013, Jul 2013, Toulouse, France. , pp.246, 2013, JOBIM TOULOUSE 2013 - RÉSUMÉS COURTS (affiches). hal-02746442

**HAL Id: hal-02746442**

**<https://hal.inrae.fr/hal-02746442>**

Submitted on 3 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# IGO, Integration from Genomes to Organisms

Sandra DEROZIER<sup>1</sup>, Hélène CHIAPELLO<sup>2</sup>, Thomas LACROIX<sup>1</sup>, Cyprien GUERIN<sup>1</sup>, Valentin LOUX<sup>1</sup>,  
Anne GOELZER<sup>1</sup>, Jean-François GIBRAT<sup>1</sup>, Franck SAMSON<sup>1</sup>, Annie GENDRAULT<sup>1</sup>

<sup>1</sup> UR 1077 Mathématique Informatique et Génome (MIG), INRA, Domaine de Vilvert, 78352, Jouy-en-Josas, France  
{sandra.derozier, thomas.lacroix, cyprien.guerin, valentin.loux, anne.goelzer,  
jean-francois.gibrat, franck.samson, annie.gendrault}@jouy.inra.fr

<sup>2</sup> UR 0875 Mathématique et Informatique Appliquées Toulouse (MIAT), INRA, chemin de Borde-Rouge, 31326,  
Castanet-Tolosan, France  
helene.chiapello@toulouse.inra.fr

**Abstract** *IGO portal is a web interface implemented with the Google Web Toolkit. It is a gateway to some of the resources developed at MIG laboratory. Currently five web interfaces and 389 complete prokaryotic genomes are integrated. IGO is accessible at <http://migale.jouy.inra.fr/igo>.*

**Keywords** genomic, integration, online resource.

IGO (Integration from Genomes to Organisms) is a web portal that interconnects different online tools developed at MIG laboratory in various domains : Mosaic [1] for the comparison of genomes, Prose [2] for the protein sequences, Bacillus subtilis Expression Data Browser [3] for the expression of B. Subtilis data, Pareo [4] for metabolic pathways and Insyght [5] for homologies/synteny.

The integration of the different web tools relies on passing arguments as URL parameters. The benefit of this approach is that web applications from different technological backgrounds can be interconnected without major additional development. However, the web tools must support URL parameters such as the organism name, the strain, the gene name or the locus tag.

The IGO web interface has been developed using the Google Web Toolkit (GWT). Data are stored in a PostgreSQL relational database which currently contains 389 bacterial organisms.

The database contains three categories of data :

- primary data such as genomic annotations that are extracted from EBI Genome Reviews files,
- secondary data corresponding to the cross comparisons, using BLAST, of all the genes of the stored bacterial genomes,
- tertiary data such as orthologs that are computed from the primary and secondary data.

The entry point in IGO is an organism, a strain and the chromosome or plasmid of interest. A table lists all the genes and various information such as the locus tag, genomic positions, EC number, product and Gene Ontology.

A search functionality is implemented to facilitate the retrieval of genes of interest. Different types of filters are available : genomic locations, presence/absence of homology, identifier (gene name/synonymous, locus tag, EMBL or Uniprot IDs), Gene Ontology terms (biological process, function, cellular component), EC number and product.

By selecting the gene name of interest, a summary of its information is available along with links pointing to the various resources developed at MIG laboratory (Fig. 1 shows an example).

The screenshot shows the IGO portal interface. At the top, the IGO logo (a spider-like figure) and the text "IGO - Integration from Genomes to Organisms" are displayed. To the right is the MIG logo (Mathématique Informatique & Génome). Below the header, there are navigation links for "B. subtilis Expression Data Browser", "MOSAIC", "PATHWAY RELATIONAL ORGANISATION", "Insyght", and "PROSE". A central panel titled "Summary information about the selected gene :" contains a table of gene details. At the bottom, the INRA logo (SCIENCE & IMPACT) and contact information are visible, along with the mig:ole logo.

Summary information about the selected gene :	
Organism	Staphylococcus aureus strain TW20 = 0582 [taxonId=663951]
Element	chromosome FN433596_GR
Name	dnaN
Locus_Tag	<a href="#">SATW20_00020</a>
EC Number	<a href="#">2.7.7.7</a>
Start - Stop (Strand)	2156 - 3289 (1)
Function	<a href="#">DNA binding</a> / <a href="#">3'-5' exonuclease activity</a> / <a href="#">DNA-directed DNA polymerase activity</a>
Product	<a href="#">DNA polymerase III, beta chain</a>
Biological Process	<a href="#">DNA replication</a>
Cellular Component	<a href="#">cytoplasm</a>

**Figure1.** Summary information about the selected gene.

The IGO portal will be gradually enriched by integrating more MIG laboratory resources and all of available complete bacterial genomes.

### Acknowledgements

The authors would like to thank Julien Jourde, Philippe Veber and Leslie Aïchaoui for their valuable comments.

### References

- [1] <http://genome.jouy.inra.fr/mosaic/>
- [2] <http://genome.jouy.inra.fr/prose/>
- [3] <http://genome.jouy.inra.fr/cgi-bin/seb/index.py>
- [4] <http://genome.jouy.inra.fr/pareo/>
- [5] <http://genome.jouy.inra.fr/Insyght/>