



GIPSy: Genomic island prediction software



Siomar C. Soares^{a,b,*}, Hakan Geyik^{b,c}, Rommel T.J. Ramos^d, Pablo H.C.G. de Sá^d,
Eudes G.V. Barbosa^{b,e}, Jan Baumbach^e, Henrique C.P. Figueiredo^f, Anderson Miyoshi^b,
Andreas Tauch^c, Artur Silva^d, Vasco Azevedo^b

^a Department of Immunology, Microbiology and Parasitology, Federal University of Triângulo Mineiro, Uberaba, Minas Gerais, Brazil

^b Department of General Biology, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

^c Institute for Genome Research and Systems Biology, Center for Biotechnology (CeBiTec), Bielefeld University, Bielefeld, Germany

^d Department of Genetics, Federal University of Pará, Belém, Pará, Brazil

^e Computational Biology laboratory, Institute for Mathematics and Computer Science, University of Southern Denmark, Odense, Denmark

^f AQUACEN, National Reference Laboratory for Aquatic Animal Diseases, Ministry of Fisheries and Aquaculture, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

ARTICLE INFO

Article history:

Received 6 May 2015

Received in revised form 28 August 2015

Accepted 11 September 2015

Available online 12 September 2015

Keywords:

Genomic islands

Pathogenicity islands

Resistance islands

Symbiotic islands

Metabolic islands

Horizontal gene transfer

Genome plasticity

ABSTRACT

Bacteria are highly diverse organisms that are able to adapt to a broad range of environments and hosts due to their high genomic plasticity. Horizontal gene transfer plays a pivotal role in this genome plasticity and in evolution by leaps through the incorporation of large blocks of genome sequences, ordinarily known as genomic islands (GEIs). GEIs may harbor genes encoding virulence, metabolism, antibiotic resistance and symbiosis-related functions, namely pathogenicity islands (PAIs), metabolic islands (MIs), resistance islands (RIs) and symbiotic islands (SIs). Although many software for the prediction of GEIs exist, they only focus on PAI prediction and present other limitations, such as complicated installation and inconvenient user interfaces. Here, we present GIPSy, the genomic island prediction software, a standalone and user-friendly software for the prediction of GEIs, built on our previously developed pathogenicity island prediction software (PIPS). We also present four application cases in which we crosslink data from literature to PAIs, MIs, RIs and SIs predicted by GIPSy. Briefly, GIPSy correctly predicted the following previously described GEIs: 13 PAIs larger than 30 kb in *Escherichia coli* CFT073; 1 MI for *Burkholderia pseudomallei* K96243, which seems to be a miscellaneous island; 1 RI of *Acinetobacter baumannii* AYE, named AbaR1; and, 1 SI of *Mesorhizobium loti* MAFF303099 presenting a mosaic structure. GIPSy is the first life-style-specific genomic island prediction software to perform analyses of PAIs, MIs, RIs and SIs, opening a door for a better understanding of bacterial genome plasticity and the adaptation to new traits.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Bacteria are highly diverse organisms broadly distributed over the world (Oren, 2004). The adaptability of bacteria to different

environments mainly results from their high genomic plasticity upon which selection will finally act allowing evolution to occur (Barbosa et al., 2014; Dobrindt and Hacker, 2001). Genome plasticity is achieved through diverse mechanisms like point mutations, rearrangements (inversions and translocations), deletions and insertion of DNA from other organisms (including plasmids, prophages, transposons and noncanonical classes of mobile genetic elements) (Bellanger et al., 2014; Brüßow et al., 2004; Dobrindt and Hacker, 2001; Schmidt and Hensel, 2004). In this context, genomic islands (GEIs) play an important role. They are large regions in the DNA sequence (~6–200 kb) acquired from other organisms through one of the above mentioned mechanisms. They promote the bacterial evolution by providing blocks of genes involved in

* Corresponding author.

E-mail addresses: siomars@gmail.com, siomar@icbn.ufcm.edu.br (S.C. Soares), hakan.geyik@uni-bielefeld.de (H. Geyik), rommelramos@ufpa.br (R.T.J. Ramos), pablogomesdesa@gmail.com (P.H.C.G. de Sá), eudesgvb@gmail.com (E.G.V. Barbosa), jan.baumbach@imada.sdu.dk (J. Baumbach), figueiredoh@yahoo.com (H.C.P. Figueiredo), miyoshi@icb.ufmg.br (A. Miyoshi), tauch@cebitec.uni-bielefeld.de (A. Tauch), asilva@ufpa.br (A. Silva), vascoariston@gmail.com (V. Azevedo).

The figure displays three sequential screenshots of the GIPSY (Genomic Island Prediction Software) Java-based graphical user interface.

Top Screenshot (Step 1): The interface shows the 'Dependencies' window with a progress bar at the top indicating Step 1 of 8. The main area prompts the user to 'Prepare files from selected organisms' and provides two sets of buttons: 'Open query' and 'Create files' for the genome to be analysed, and 'Open Subject' and 'Create files' for the reference genome. A central graphic features the 'GIPSY' logo and a DNA double helix. A contact email for the developer, Siomar Soares (siomars@gmail.com), is provided. A 'Next >' button is at the bottom right.

Middle Screenshot (Step 5): The interface shows Step 5 of 7. The left panel lists search categories: 'Search for specific factors', 'Virulence Factors (Pathogenicity Islands)', 'Antibiotic Resistance (Resistance Islands)', 'Metabolism (Metabolic Islands)', and 'Symbiosis (Symbiotic Islands)'. Below these is a table of Query IDs and Subject IDs. The right panel shows a BLAST search window with parameters: 'BLASTP 2.2.26 [Sep-21-2011]', a reference to Altschul et al. (1997), a query 'c0011 R 7147:8577 Putative transporter yaaJ (476 letters)', and a database of 55,193 sequences. A table at the bottom shows search results for various query IDs, with 'c0011' highlighted, showing a 30% hit rate and other metrics.

Bottom Screenshot (Step 8): The interface shows Step 8 of 8. The main area is titled 'Predict Genomic Islands based on previously identified features'. It has four sections: 'Pathogenicity Islands', 'Resistance Islands', 'Metabolic Islands', and 'Symbiotic Islands', each with 'Run' and 'Visualize' buttons. Below is a detailed table of predicted islands with columns for Putative Genomic Island, G+C Dev., Codon U., Virulence, Hypothet., Gene Composition, Position, and Predictio. The table lists various islands and their characteristics.

Putative Genomic Island	G+C Dev.	Codon U.	Virulence	Hypothet.	Gene Composition	Position	Predictio.
Genome	10%	24%	44%	41%	NA	NA	NA
Putative Genomic Island 1	66%	100%	33%	66%	c0133-c0139	126688..132406	NA
Putative Pathogenicity Island 1	21%	66%	64%	76%	c0253-c0368	248421..348625	Strong
Putative Pathogenicity Island 2	37%	37%	75%	50%	c0391-c0398	370109..378117	Normal
Putative Pathogenicity Island 3	0%	66%	100%	33%	c0409-c0414	388978..395711	Normal
Putative Pathogenicity Island 4	20%	37%	77%	41%	c0932-c0979	908853..942509	Normal
Putative Pathogenicity Island 5	13%	68%	100%	59%	c1165-c1293	1127423..1241384	Strong
Putative Pathogenicity Island 6	21%	72%	67%	57%	c1400-c1475	1327847..1373052	Strong
Putative Genomic Island 2	14%	59%	14%	74%	c1481-c1507	1377890..1389209	NA
Putative Pathogenicity Island 7	17%	58%	62%	39%	c1515-c1602	1395726..1452493	Normal
Putative Genomic Island 3	53%	100%	15%	100%	c1881-c1893	1715525..1728047	NA
Putative Pathogenicity Island 8	0%	77%	77%	44%	c1935-c1943	1777183..1787317	Normal
Putative Genomic Island 4	33%	100%	16%	0%	c1955-c1960	1796112..1802390	NA
Putative Pathogenicity Island 9	61%	61%	91%	63%	c2302-c2438	2193646..2250680	Strong

Fig. 1. Screenshots of GIPSY's Java-based graphical user interface.

In the top of the figure is step 1, where the GenBank or EMBL files for the query and subject genomes are provided for the software to create the additional files. In the middle of the figure is step 5, where the software searches the query genome for GEI specific factors (*i.e.*, virulence, antibiotic resistance, metabolism-related or symbiosis-related factors). In the bottom of the figure is step 8, where the software predicts the specific class of GEI (pathogenicity, resistance, metabolic or symbiotic island) based on the resulting files from previous steps.

Download English Version:

<https://daneshyari.com/en/article/22531>

Download Persian Version:

<https://daneshyari.com/article/22531>

[Daneshyari.com](https://daneshyari.com)