

Genome analysis

CiVi: circular genome visualization with unique features to analyze sequence elements

Lex Overmars^{1,2,3,*}, Sacha A. F. T. van Hijum^{1,2,4}, Roland J. Siezen^{1,2} and Christof Francke^{1,2,5}

¹CMBI, Radboud University Medical Centre, Nijmegen 6525GA, The Netherlands, ²Netherlands Bioinformatics Centre, Nijmegen 6500HB, The Netherlands, ³Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands, ⁴NIZO Food Research B.V., Ede, The Netherlands and ⁵Hogeschool Arnhem en Nijmegen BioCentre, 6525EM Nijmegen, The Netherlands

*To whom correspondence should be addressed.

Associate Editor: John Hancock

Received on November 28, 2014; revised on April 17, 2015; accepted on April 20, 2015

Abstract

Summary: We have developed CiVi, a user-friendly web-based tool to create custom circular maps to aid the analysis of microbial genomes and sequence elements. Sequence related data such as gene-name, COG class, PFAM domain, GC%, and subcellular location can be comprehensively viewed. Quantitative gene-related data (e.g. expression ratios or read counts) as well as predicted sequence elements (e.g. regulatory sequences) can be uploaded and visualized. CiVi accommodates the analysis of genomic elements by allowing a visual interpretation in the context of: (i) their genome-wide distribution, (ii) provided experimental data and (iii) the local orientation and location with respect to neighboring genes. CiVi thus enables both experts and non-experts to conveniently integrate public genome data with the results of genome analyses in circular genome maps suitable for publication.

Contact: L.Overmars@gmail.com

Supplementary information: [Supplementary data](#) are available at *Bioinformatics* online.

Availability and implementation: CiVi is freely available at <http://civi.cmbi.ru.nl>

1 Introduction

Circular genome representations provide an excellent way to comprehensively inspect genome-wide data. Various tools have been developed to generate these circular visualizations including CGView (Stothard and Wishart, 2005), GenomeVx (Conant and Wolfe, 2008), GeneWiz (Hallin *et al.*, 2009) and DNAPlotter (Carver *et al.*, 2009). Tools such as BRIG (Alikhan *et al.*, 2011), CGView Comparison Tool (Grant *et al.*, 2012), Circos (Krzywinski *et al.*, 2009) and Circleator (Crabtree *et al.*, 2014) can also visualize genome comparisons and in some cases, visualize links between genome sequence and other types of information. The tools described above are well-suited for the visualization of high-throughput genomic data like sequence-similarities or read counts, but have limited functionality in relation to smaller-scale activities such as reconstructing transcription networks and finding gene functions associated to

genetic elements. Moreover, with the exception of GeneWiz (Hallin *et al.*, 2009), they require laborious and sometimes complex uploads of genome and annotation data. In 2013 we published MGcV (Overmars *et al.*, 2013), a linear-genomic context visualization tool tailored to provide a simple and quick visual access to the publicly available genomic data from NCBI (Pruitt *et al.*, 2012). The tool incorporated the capabilities of the earlier MGv (Kerkhoven *et al.*, 2004), and extended the visualization with data export options to advance the gene-specific analysis of microbial genomes. We now present CiVi, which has been developed using the same philosophy, to extend the circular viewing options provided by the original MGv. The extensions include the possibility to display annotation data like COG category, PFAM domain or subcellular location directly on the genome map and to reveal the position of selected annotations through a keyword search option. CiVi also enables the

upload and visualization of custom data, such as the positions of genomic elements and the export of associated data. The latter include the function annotations and/or sequences of neighboring genes, as well as information on the distance distribution of the elements with respect to the genes. The resulting circular maps can be edited and the pictures exported in svg-, png- and pdf-format. Finally, CiVi offers a completely new interface and back-end with enhanced usability, interactivity (via mouse-over) and speed.

2 Usage and implementation

CiVi enables users to create custom circular microbial genome maps in a simple step-wise fashion, adding data ring by ring. The interface consists of a panel on the right in which the map is displayed and four panels on the left related to the different menus, labeled: (i) 'Genome and data selection'; (ii) 'On display'; (iii) 'Data import'; and (iv) 'Elements and genomic context'. For every ring the user can subsequently set (panel i): the organism and genome of interest, the type of information, the coloring and the radius of the ring. The types of information that can be included directly are: the location of the genes on the +strand and -strand, COG categories (NCBI RefSeq), GC%, GC-skew and AT-skew (calculated; Overmars et al., 2013), PFAM domains, and subcellular location predictions (PSORTdb; Nancy et al., 2011). A keyword matching option allows highlighting genes whose gene product, gene name, COG code or PFAM ID match a query. In addition, the coordinates and a title or background coloring can be added, and gene-associated quantitative data can be represented by either (bar graph like-) spikes or by a red-to-green gradient. The categorical data have been linked to fixed colors. Map additions are tracked in the 'On display'-panel, in which added rings can also be removed. Users can upload three types of data in the 'Data import'-menu: quantitative data (e.g. expression ratios), the predicted position of sequence elements (e.g. regulatory elements) and/or custom color schemes to designate any genomic region in the genome map. Different genomes can be included in a single circular map, but as synteny is not determined this feature should only be used with very closely related genomes.

2.1 Analysis and visualization of sequence elements

An integrated view of genome-wide experimental data and the predicted location of particular regulatory elements can be very allusive in the analysis of transcriptional networks (as illustrated in Fig. 1 and supplementary file S1). The position of any particular genomic element with respect to the location and orientation of the surrounding genes can hint at the biological role of that element. CiVi generates plots in the 'Elements and genomic context'-panel for each uploaded set of sequence elements, in which both the distance to the neighboring genes and the orientation with respect to the genes is summarised (Fig. 1B). Similarly, the biological role of a particular element may be derived from the functional characteristics of the gene context. The user can download the positions and the annotation data for the gene context for subsequent analysis using the 'Generate context table'-link.

2.2 Implementation

CiVi is a web-application developed using a combination of python, javascript, MySQL and SVG. CiVi was implemented as a single page application; the front-end makes server side calls through JQuery and AJAX and receives a response from the server. The maps can be downloaded in SVG, PNG or PDF format, where conversions are done using 'Batik Rasterizer'. The maps in SVG-format can be

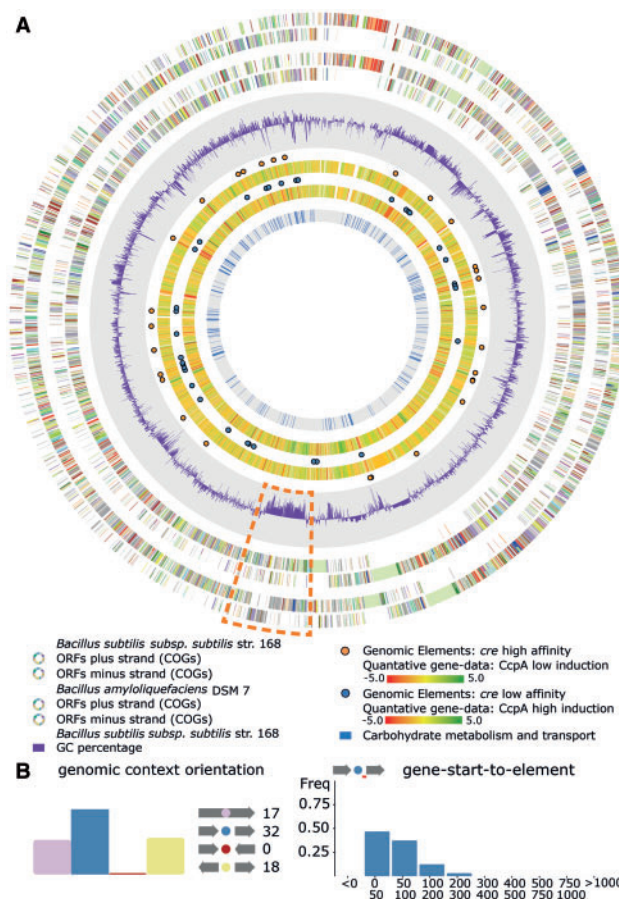


Fig. 1. (A) Application of CiVi in the analysis of genome structure and gene regulation. A comparative visualization of the genomes of the closely related *Bacillus* species *B. subtilis* str. 168 (genes and COG annotation on outer two rings) and *B. amyloliquefaciens* DSM 7 (next two rings) shows that the former genome contains a region that seems to have been inserted (indicated by orange box) and which is clearly associated with a deviating GC percentage (5th ring). In *Bacillus* species carbohydrate uptake and metabolism is governed by a phenomenon called 'carbon catabolite repression' which is mediated by CcpA. CcpA binds to a characteristic DNA operator sequence called *cre*, located just upstream of the regulated gene(s). Variability between the operator sequences was assumed to cause a variability in the response depending on the level of repression (Francke et al., 2008). The postulated operator dependent variable response was tested experimentally and confirmed (Marciniak et al., 2012). All of the characteristic properties of CcpA mediated regulation are apparent in an integrative visualization of the transcription factor binding site predictions (6th and 8th ring), gene expression data (7th and 9th ring) and the gene annotation (COG category 'carbohydrate transport and metabolism'; 10th ring). **(B)** The provided analysis of the genomic context of the predicted CcpA binding sites confirms the validity of the input motif; as it is mainly found upstream of genes and close to the translation start

edited in programs such as Adobe Illustrator. CiVi is operable in Firefox, Chrome and Internet Explorer.

3 Conclusion

CiVi is a versatile and easy-to-use web-application to create custom circular genome maps. It provides a visual integration of publicly available genomic data and additional provided data, the latter including e.g. gene expression data and genomic elements. The functional analysis of latter elements is aided by the characterization of their genomic context, a feature that is unique to CiVi.

Funding

This work was supported by the BioRange programme of the Netherlands Bioinformatics Centre (NBIC), which is supported by the Netherlands Genomics Initiative (NGI).

Conflict of Interest: none declared.

References

- Alikhan, N.F. *et al.* (2011) BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics*, **12**, 402.
- Carver, T. *et al.* (2009) DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics*, **25**, 119–120.
- Conant, G.C. and Wolfe, K.H. (2008) GenomeVx: simple web-based creation of editable circular chromosome maps. *Bioinformatics*, **24**, 861–862.
- Crabtree, J. *et al.* (2014) Circleator: flexible circular visualization of genome-associated data with BioPerl and SVG. *Bioinformatics*, **30**, 3125–3127.
- Francke, C. *et al.* (2008) A generic approach to identify Transcription Factor-specific operator motifs; Inferences for LacI-family mediated regulation in *Lactobacillus plantarum* WCFS1. *BMC Genomics*, **9**, 145.
- Grant, J.R. *et al.* (2012) Comparing thousands of circular genomes using the CGView Comparison Tool. *BMC Genomics*, **13**, 202.
- Hallin, P.F. *et al.* (2009) GeneWiz browser: an interactive tool for visualizing sequenced chromosomes. *Standards Genomic Sci.*, **1**, 204.
- Kerkhoven, R. *et al.* (2004) Visualization for genomics: the Microbial Genome Viewer. *Bioinformatics*, **20**, 1812–1814.
- Krzywinski, M. *et al.* (2009) Circos: an information aesthetic for comparative genomics. *Genome Res.*, **19**, 1639–1645.
- Marciniak, B.C. *et al.* (2012) High- and low-affinity CRE boxes for CcpA binding in *Bacillus subtilis* revealed by genome-wide analysis. *BMC Genomics*, **13**, 401.
- Nancy, Y.Y. *et al.* (2011) PSORTdb—an expanded, auto-updated, user-friendly protein subcellular localization database for Bacteria and Archaea. *Nucleic Acids Res.*, **39**, D241–D244.
- Overmars, L. *et al.* (2013) MGcV: the microbial genomic context viewer for comparative genome analysis. *BMC Genomics*, **14**, 209.
- Pruitt, K.D. *et al.* (2012) NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res.*, **40**, D130–D135.
- Stothard, P. and Wishart, D.S. (2005) Circular genome visualization and exploration using CGView. *Bioinformatics*, **21**, 537–539.