Integrative Multi-omics Module Network Inference with Lemon-Tree

Eric Bonnet, Laurence Calzone, Tom Michoel

▶ **To cite this version:**

Eric Bonnet, Laurence Calzone, Tom Michoel. Integrative Multi-omics Module Network Inference with Lemon-Tree. PLoS Computational Biology, 2015, 11 (2), pp.e1003983. 10.1371/journal.pcbi.1003983 . inserm-02140297

HAL Id: inserm-02140297

https://inserm.hal.science/inserm-02140297

Submitted on 27 May 2019

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.

PLOS | COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Integrative Multi-omics Module Network Inference with Lemon-Tree

Eric Bonnet[1,2,3]*, **Laurence Calzone**[1,2,3], **Tom Michoel**[4]*

1 Institut Curie, Paris, France, **2** INSERM U900, Paris, France, **3** Mines ParisTech, Fontainebleau, France, **4** Division of Genetics & Genomics, The Roslin Institute, The University of Edinburgh, Easter Bush, Midlothian, United Kingdom

* eric.bonnet@curie.fr (EB); tom.michoel@roslin.ed.ac.uk (TM)

CrossMark
click for updates

OPEN ACCESS

Citation: Bonnet E, Calzone L, Michoel T (2015) Integrative Multi-omics Module Network Inference with Lemon-Tree. PLoS Comput Biol 11(2): e1003983. doi:10.1371/journal.pcbi.1003983

Editor: Paul P Gardner, University of Canterbury, New Zealand

Received: July 29, 2014

Accepted: October 14, 2014

Published: February 13, 2015

Copyright: © 2015 Bonnet et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data used for this research were obtained from the public domain and no new data were generated as part of this research. Formatted test data and software source code are available from http://lemon-tree.googlecode.com.

Funding: This research was supported by Roslin Institute Strategic Grant funding from the BBSRC (TM). The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement number FP7-HEALTH-2010-259348 (ASSET project). EB and LC are members of the team "Computational Systems Biology of Cancer", Equipe labellisée par la Ligue Nationale

Abstract

Module network inference is an established statistical method to reconstruct co-expression modules and their upstream regulatory programs from integrated multi-omics datasets measuring the activity levels of various cellular components across different individuals, experimental conditions or time points of a dynamic process. We have developed Lemon-Tree, an open-source, platform-independent, modular, extensible software package implementing state-of-the-art ensemble methods for module network inference. We benchmarked Lemon-Tree using large-scale tumor datasets and showed that Lemon-Tree algorithms compare favorably with state-of-the-art module network inference software. We also analyzed a large dataset of somatic copy-number alterations and gene expression levels measured in glioblastoma samples from The Cancer Genome Atlas and found that Lemon-Tree correctly identifies known glioblastoma oncogenes and tumor suppressors as master regulators in the inferred module network. Novel candidate driver genes predicted by Lemon-Tree were validated using tumor pathway and survival analyses. Lemon-Tree is available from http://lemon-tree.googlecode.com under the GNU General Public License version 2.0.

This is a *PLOS Computational Biology* Software Article

Introduction

Recent years have witnessed a dramatic increase in new technologies for interrogating the activity levels of various cellular components on a genome-wide scale, including genomic, epigenomic, transcriptomic, and proteomic information [1]. It is generally acknowledged that integrating these heterogeneous datasets will provide more biological insights than performing separate analyses. For instance, in 2005, Garraway and colleagues combined SNP-based genetic maps and expression data to identify a novel transcription factor involved in melanoma progression [2]. More recently, international consortia such as The Cancer Genome Atlas (TCGA) or the International Cancer Genome Consortium (ICGC) have launched large-scale initiatives to characterize multiple types of cancer at different levels (genomic, transcriptomic, epigenomic, etc.) on several hundreds of samples. These integrative studies have already led to the identification of novel cancer genes [3, 4].

PubMed

Format: Abstract

Full text links

PLOS COMPUTATIONAL BIOLOGY PMC FREE Full text Services BibloInserm

PLoS Comput Biol. 2015 Feb 13;11(2):e1003983. doi:

Integrative multi-omics module network inference with Lemon-Tree.

Bonnet E[1], Calzone L[1], Michoel T[2].

Author information

1 Institut Curie, Paris, France; **INSERM** U900, Paris, France; Mines ParisTech, Fontainebleau, France.

2 Division of Genetics & Genomics, The Roslin Institute, The University of Edinburgh, Easter Bush, Midlothian, United Kingdom.

Abstract

Module network inference is an established statistical method to reconstruct co-expression modules and their upstream regulatory programs from integrated multi-omics datasets measuring the activity levels of various cellular components across different individuals, experimental conditions or time points of a dynamic process. We have developed Lemon-Tree, an open-source, platform-independent, modular, extensible software package implementing state-of-the-art ensemble methods for module network inference. We benchmarked Lemon-Tree using large-scale tumor datasets and showed that Lemon-Tree algorithms compare favorably with state-of-the-art module network inference software. We also analyzed a large dataset of somatic copy-number alterations and gene expression levels measured in glioblastoma samples from The Cancer Genome Atlas and found that Lemon-Tree correctly identifies known glioblastoma oncogenes and tumor suppressors as master regulators in the inferred module network. Novel candidate driver genes predicted by Lemon-Tree were validated using tumor pathway and survival analyses. Lemon-Tree is available from http://lemon-tree.googlecode.com under the GNU General Public License version 2.0.

PMID: 25679508 PMCID: PMC4332478 DOI: 10.1371/journal.pcbi.1003983

[Indexed for MEDLINE] **Free PMC Article**

Images from this publication. See all images (4) Free text