Journal of Immunological Methods 387 (2013) 71–80

Contents lists available at SciVerse ScienceDirect

Journal of Immunological Methods

journal homepage: www.elsevier.com/locate/jim

Research paper

Decoding dendritic cell function through module and network analysis ☆

Gaurav Pandey [a,*], Ariella Cohain [a], Jennifer Miller [b], Miriam Merad [b]

[a] Institute for Genomics and Multiscale Biology and Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY, USA
[b] The Immunology Institute, Tisch Cancer Institute and Department of Oncological Sciences, Mount Sinai School of Medicine, New York, NY, USA

ARTICLE INFO

Article history:
Received 17 September 2012
Accepted 17 September 2012
Available online 23 October 2012

Keywords:
Dendritic cells
Gene expression data
Clustering
Network analysis
Transcription factors
Regulatory networks

ABSTRACT

Systems biology approaches that utilize large genomic data sets hold great potential for deciphering complex immunological process. In this paper, we propose such an approach to derive informative modules and networks from large gene expression data sets. Our approach starts with the clustering of such data sets to derive groups of tightly co-expressed genes, also known as co-expression modules. These modules are then converted into co-expression networks, and combined with transcriptional regulatory and protein interaction data to generate integrated networks that can help decipher the regulatory structure of these modules. We use this approach to derive the first set of modules and networks focused on dendritic cells (DCs). These cells are responsible for sampling the local environment to inform the adaptive immune system about peripheral stimuli, thus leading to the induction of an immune response. Using the ImmGen gene expression data set, we derive co-expression modules and integrated networks for the pDC, cDC and CD8+ DC subsets. In addition to recapitulating genes known to regulate the functions of these subsets, these networks reveal several novel genes and interactions that might have important roles in DC biology.

1. Introduction

Immune responses result from a complex interaction that relies on an elaborate and dynamic communications network that exists among the many different immune cell types that patrol the body. Although several of the cellular and molecular cues that control the induction of successful immune responses have been identified, there is an immense need for a systems-level understanding of how the different components of immune cells interact in the steady state and in response to different stimuli. To address this need, several groups have started utilizing the recent wave of biotechnologies to profile the immune system at the molecular level and analyze the related data to obtain novel insights (Gardy et al., 2009; Germain et al., 2011). In particular, the ImmGen consortium has profiled the genome-wide expression patterns in all the cell types in the immune system of *Mus musculus* (mouse), thus making available an unprecedented resource for such studies (Heng and Painter, 2008). This and other data sets have been utilized by some recent rigorous computational systems biology approaches that have built models of how the different components of the immune system function individually and in concert with the others (Amit et al., 2009; Germain et al., 2011; Novershtern et al., 2011; Benichou et al., 2012). However, the findings of these studies have largely been limited to the immune cells where rich data sets are available, such as T- and B-cells.

An important component of the immune system whose understanding has not benefitted much from these studies is dendritic cells (DCs) (Banchereau and Steinman, 1998). DCs are one of two types of mononuclear phagocytes that populate most tissues, the other being macrophages. The term "phagocyte" derives from the Greek word "phago", meaning "to devour", and reflects the ability of DCs and macrophages to capture exogenous proteins and damaged or dying cells. In

☆ This paper was presented at the 3rd Immunoinformatics and Computational Immunology Workshop (ICIW2012), Orlando, Florida, USA October 7, 2012.

* Corresponding author at: Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, One Gustave L. Levy Place, Box 1498, New York, NY 10029, USA. Tel.: +1 212 659 8535; fax: +1 646 537 8660.
E-mail address: gaurav.pandey@mssm.edu (G. Pandey).

0022-1759/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jim.2012.09.012

contrast to macrophages, whose main role is to scavenge phagocytosed material, DCs sample the local environment to inform the adaptive immune system about peripheral cues. They constantly transport environmental proteins broken down into small peptides termed "antigens" to the lymph node. There, they present self and foreign antigens on MHC-class I- and MHC-class II peptide complexes on the cell surface to resident lymphocytes and produce large amounts of activating cytokines (Guermonprez et al., 2002; Trombetta and Mellman, 2005) to promote the differentiation of antigen-specific effector immune responses (Steinman and Banchereau, 2007). In the case of the presentation of self-antigens, DCs cause the differentiation of antigen-specific T regulatory cells or the depletion of auto-reactive T cells 1(Steinman et al., 2003). MHC-class I and MHC-class II peptides are presented by DCs to induce a CD8+ or CD4+ T cell response respectively. CD8 + T cells are cytotoxic T cells, which specialize in the elimination of infected cells and thus are geared to respond to intracellular pathogens, while CD4+ T cells initiate antibody production of antigen-specific antibodies by B cells to respond to extracellular pathogens. Clearly, DCs play a key role in directing effective immune responses. However, the study of DCs has been hampered due to their rarity within tissues and, until recently, the inability to distinguish DCs from other tissue phagocytes such as macrophages.

Recent data have established that DCs consist of distinct subsets with different abilities to process antigens, respond to environmental stimuli and engage distinct effector lymphocytes (Heath and Carbone, 2009). The DC population can be divided into the following subsets based on ontogeny and function: plasmacytoid DCs (pDCs) and classical DCs (cDCs). These cells arise from different origins in the immune cell lineage and serve specialized immunological functions. pDCs secrete large amounts of the antiviral interferon alpha (IFN-α) cytokine in response to the stimulation of pathogen recognition receptors TLR7 and TLR9 to initiate T cell immunity against viral antigens (Reizis et al., 2011). These cells express low levels of MHC-II and the co-stimulatory cytokines needed to activate T cells in steady state tissue. In contrast, cDCs express high levels of MHC as well as co-stimulatory molecules and are the only hematopoietic cell population with the ability to stimulate naïve T cells in the steady state. Other hematopoietic populations can only stimulate T cells that have already been exposed to antigen, or "memory T cells".

In lymphoid tissue, cDCs consist of two main subsets, namely the CD8+ and CD8– DCs. CD8+ cDCs excel in the cross-presentation of cell-associated antigens and are most potent at stimulating $CD8^+$ T cells to induce a Th1 response (Coombes and Powrie, 2008). This population relies on the cytokine receptor Flt3 and the transcription factors ID2, Batf3, and Irf8 for development. In contrast, CD8– cDCs are most potent at inducing CD4+ T cells to induce a Th2 response (Heath and Carbone, 2009). This population requires Irf4 for their development (Reizis et al., 2011). Recent data established that CD8– DCs are very likely heterogeneous and include at least two main populations that are differentially controlled by Notch2 signaling (Lewis et al., 2011), thus making them very difficult to study.

Owing to the low numbers of DCs in tissues, the difficulty of isolating them from peripheral tissues, and the general expense of these procedures, most DC studies have been limited to the spleen with a limited number of replicates. Through targeted, generally low-throughput, studies, several genes have been identified to be involved in the functioning of DCs and their response to antigens. These genes, several of which are known regulators, include Relb, Irf8, Id2 and Flt3 (Shortman and Heath, 2010). Recent studies have employed high-throughput technologies, such as microarrays, to understand DC biology in vivo. This has greatly accelerated the study of DCs by 1) identifying subset-specific regulators, including Batf3 (Hildner et al., 2008), and most recently, Zbtb46 (Meredith et al., 2012; Satpathy et al., 2012), 2) showing that DC subsets differentially express important surface receptors and regulators (Edwards et al., 2008) and 3) that genes characteristic of the various DC subsets are conserved (Contreras et al., 2010). However, these studies utilize single gene analyses, such as measuring differential expression, to identify genes important for the functioning of DCs (Bar-On et al., 2010; Crozat et al., 2010; Manicassamy et al., 2010; Chevrier et al., 2011). Clearly, such approaches do not reveal the interactions between genes that are equally critical for this problem, as has been done for other immune cell types by systems biology approaches (Amit et al., 2009; Germain et al., 2011; Novershtern et al., 2011; Benichou et al., 2012).

Motivated by the need to build models for DC function that reveal cellular interactions in addition to important genes, we propose a systematic approach that derives detailed modules and networks from large-scale gene expression data sets. For this, we use the WGCNA algorithm (Langfelder and Horvath, 2008) to cluster the relevant portion of the ImmGen gene expression data into groups of tightly co-expressed genes, also known as co-expression modules.[1] We further convert these modules into co-expression networks, and integrate them with transcriptional regulatory data from the Molecular Signature Database (MSigDB) (Subramanian et al., 2005) and protein interaction data from BioGRID (Stark et al., 2011) to generate integrated networks that can help decipher the regulatory structure of these modules.

We use this approach to derive the first set of DC-focused modules and networks (to the best of our knowledge). For this, we build on Miller et al.'s (2012)'s work, where several insights were revealed about DC subsets, specifically cDCs, pDCs and CD8+'s, as well as overall DC functioning. Using their proposed core signatures for these subsets, we conduct an extensive evaluation of our pipeline to identify the most enriched modules that reveal informative integrated networks consisting of many genes and their co-expression and regulatory interactions. A detailed examination of these networks and modules highlights several novel genes, as well as interactions, that may explain the functioning of cDC, pDC and CD8+ cells, and thus add valuable knowledge to DC biology.

In summary, through the example of dendritic cells, we demonstrate how established algorithms and data sources can help generate actionable hypotheses about critical immunological processes, especially involving cell types that are under-represented in data sets.

[1] The terms "cluster" and "module" will be used interchangeably in the rest of this paper.