

# Differential Network Analysis of Longitudinal Gene Expression in Response to Perturbations.\*

Shuyue Xue<sup>a,b,1</sup>, Lavida R.K. Rogers<sup>c</sup>, Minzhang Zheng<sup>b</sup>, Jin He<sup>d</sup>, Carlo Piermarocchi<sup>a,\*</sup>, George I. Mias<sup>a,b,d,\*</sup>

<sup>a</sup>*Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA*

<sup>b</sup>*Institute for Quantitative Health Science and Engineering, Michigan State University, East Lansing, MI 48824, USA*

<sup>c</sup>*Department of Biological Sciences, University of the Virgin Islands, Virgin Islands*

<sup>d</sup>*Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA*

---

## Abstract

Understanding changes in gene expression under the effects of a perturbation is a key goal of systems biology. A powerful approach to address this goal uses gene networks and describes the perturbation's effects as a rewiring of each gene's connections. This approach is known as differential network (DN) analysis. Here, we used DNAs to analyze RNA-sequencing time series datasets, focusing on expression changes: (i) In the saliva of a human subject after vaccination with a pneumococcal vaccine (PPSV23), and (ii) in B cells treated *ex vivo* with a monoclonal antibody drug (Rituximab). Using network community detection, we revealed the collective behavior of clusters of genes, and detected communities of genes based on their longitudinal behavior, and corresponding pathway activations. We identified biological pathways consistent with the mechanism of action of the vaccine and with Rituximab's targets. The approach may be useful in drug development by providing an effective analysis of expressing changes in response to a drug.

*Keywords:* time series,, networks, gene expression

---

\*This project has been funded by the Translational Institute for Space Health Grant T0412.

\*Corresponding authors

Email addresses: [piermaro@msu.edu](mailto:piermaro@msu.edu) (Carlo Piermarocchi), [gmi@msu.edu](mailto:gmi@msu.edu) (George I. Mias)

<sup>1</sup>Lead Contact [xueshuy1@msu.edu](mailto:xueshuy1@msu.edu) (Shuyue Xue).

2020 MSC: 92-08, 92C42

---

## 1. Introduction

Recent advances in RNA-sequencing (RNA-seq) technologies have enabled researchers to incorporate time in the analysis of large-scale gene expression data from biological systems. These time course (TC) experiments can generate 5 longitudinal RNA-seq data as time series, with time-steps ranging from hours to weeks. This not only permits capturing the transcriptome's dynamic regulation over time, but also generates large amounts of expression data that need to be analyzed and interpreted. While TC experiments can measure time-dependent gene expression changes, a good data analysis scheme is key for translating the 10 experimental observations into meaningful information about the underlying biological dynamics [1, 2, 3, 4, 5, 6].

There are currently a wide variety of well-established data analysis methods for RNA-seq data. However, this is not the case for TC RNA-seq, as in this case another dimension, time, is present. Some methods originally developed 15 for microarray sequencing have been adapted to RNA-seq time series data [3, 6, 7, 8]. Network-based analysis and, in particular, Differential Network (DN) analysis methods, have been shown to be very useful in the analysis of the dynamics of gene expression under the effect of an external perturbation [9] and could provide insightful interpretations of the RNA-seq data. DN analysis 20 is a method based on the subtraction of one network from another, and has been adopted in many genomics studies in the past decade [10, 11, 12]. In a gene-gene correlation network (co-expression network), vertices represent genes while edges represent the correlation coefficient of the expression of two genes. Typically, these co-expression networks are weighted by the strength and the 25 sign of the correlation between two genes. The DN analysis method uses a pairwise cancellation of nodes and edges common to two networks that describe the expression data before and after a given perturbation. In doing so, the process leaves behind interaction variations that describe the network rewiring induced

by the perturbation. For instance, in gene expression studies, the DN analysis  
30 method successfully separated gene expressions under specific drug responses  
from generic stress responses [13]. It also aided researchers in investigating dys-  
functional regulatory networks in unhealthy states, providing insights into the  
genetic basis of diseases [14]. By focusing on the structural difference between  
35 two networks, the DN analysis method has demonstrated its effectiveness in  
identifying biological activities in different states. In addition, this graph-based  
model offers an advantage in representing the architecture of a gene network's  
overall changes where the emphasis is on the nature of interactions rather than  
the quantitative predictions in time.

In the present study, we applied a DN approach to RNA-seq time series  
40 datasets retrieved from two longitudinal TC RNA-seq experiments: (i) The  
first dataset (GSE108664) was generated from saliva samples from a healthy  
individual before and after the administration of the Pneumococcal Polysac-  
charide Vaccine (PPSV23) [15]. The primary goal of this study was to gain  
insights on the adaptive immune responses to PPSV23 through saliva profiling.  
45 Due to its convenience in processing relative to blood samples, saliva draws  
much interest for diagnostics as well as health monitoring applications. Saliva  
analysis can produce results in a timely manner, its collection is minimally inva-  
sive, and little training is required for saliva sampling, even for non-medically  
trained professionals. (ii) The second dataset (GSE100441) was generated from  
50 a time course experiment on primary B cells, where one set was treated with  
Rituximab and another used as an untreated control. Rituximab is known for  
its therapeutic use in targeting B cells [16] to treat cancers such as lymphomas  
and leukemias. This drug has a history of safe and effective usage since 1997  
[17], and the World Health Organization (WHO) place Rituximab on their list  
55 of essential medicines [18]. Rituximab binds with CD20, expressed on pre-B  
and mature B cells, but not on stem cells [19]. The binding causes perturba-  
tions to intracellular signaling and membrane structure [20], mediating the cell  
depletion. It is worthwhile to mention that the B cell pathways of Rituximab  
activation have been experimentally validated [21, 22, 23, 24], which facilitates

60 the evaluation of the effectiveness of the DN method we utilize in this work. Both the saliva and primary B cell experiments involve drug-treated samples (treatment sets) and untreated samples (control sets) monitored over time.

65 For both datasets, we started with building gene networks, one for each of the control and the treatment sets. We used gene-gene correlations between time series signals, over 24 hours in saliva and 15 hours in B cells, to evaluate pairwise gene connections. Graphically, the time series correlation networks built from the treatment sets summarized a system-wide pathway activation due to the perturbation, whereas the networks from the controls sets acted as the baseline. Within the DN analysis framework, we subtracted the baseline 70 network from the one obtained using the treatment data, arriving at the final differential network.

75 The presence of modules, also known as communities, describes a topological property of networks [25, 26, 27, 28]. One community is a group of densely connected nodes. In the context of a biological system, nodes in the same community are assumed to be close in biological functions [29, 30, 31, 32, 33]. We exploited this property of the differential network to observe fine details 80 of gene groups affected by the perturbation. Specifically, we employed one of the most established module detection algorithms, the Louvain method [34], to identify communities in our final differential network. We then performed Reactome [35] pathway enrichment analysis on individual communities, and finally examined the corresponding heatmaps for each community.

## 2. Results

85 Our RNA-seq time series raw data were retrieved from the Gene Expression Omnibus database under accessions GSE108664 and GSE100441 for the saliva and B cell experiments, respectively. The study of the immune response to the PPSV23 vaccine in saliva probed the expression of a potential 84647 gene identifiers (GENCODE annotation[36]) at 24 time points [15]. The other study of drug activation by Rituximab in B cells provided a dataset for 6 time points.

Since gene co-expression networks rely on correlations, our network analysis  
90 could be prone to spurious correlations, which we removed as described in the  
STAR Methods.

We constructed our saliva DN by subtracting the saliva network without  
vaccine from the network obtained using post-vaccine data. The B cell DN in  
response to Rituximab was generated in a similar manner. Next, we clustered  
95 the DNs into communities using the Louvain community detection method [34].  
We then conducted a Reactome Enrichment Analysis [35] using PyIOMica [37],  
on each community to identify significant pathways and associated genes. We  
also visualized the heatmaps of relative gene expression as a function of time  
for each community. Finally, we plotted the DNs and their major individual  
100 communities. The workflow is summarized in Figure 1, and additional details  
are provided in the STAR Methods.

### *2.1. Saliva DN*

Our saliva DN contains 1144 nodes (i.e., genes) and 13,775 edges. The  
Louvain algorithm identified 48 communities (modules) in total. 15 of the com-  
105 munities have a size of at least 4 nodes, while the remaining 33 are pairs or  
triplets. In the global saliva DN visualization, we excluded the communities  
with pairs or triplets, as none of them belonged to the three major connected  
components of the DN network. We also filtered the network to remove con-  
nected components with less than 4 genes. The global saliva DN is presented in  
110 Fig. 2a, where communities are visualized using different colors and encircled  
in loops. Furthermore, community labels are based on their size (largest to  
smallest, with C0 being the largest community, and C14 the smallest).

### *2.2. Pathway Enrichment of Saliva Communities*

In our pathway analysis, we queried individual communities to investigate  
115 how their highly co-expressed genes are functionally related. Our analysis is  
based on the Reactome pathway database [35, 38, 39].

Statistically significant enrichment of pathways (with False Discovery Rate (FDR) < 0.05) was identified in 6 communities, C0, C1, C2, C4, C8 and C9. The majority of statistically significant Reactome pathways were related to response to stimulus, immune response, and inflammatory response. Among the 120 six communities, C0 and C1 are the two largest communities. C0 comprises of 248 genes, colored in red in the global DN shown in Fig. 2a, whereas C1 contains 198 genes, colored in yellow in the same panel. We display the C0 and C1 in Fig. 2b as representative communities. Genes that belong to the statistically 125 significant biological pathways are highlighted in red in Fig. 2b.

In the C0 community, the Reactome enrichment analysis identified 15 statistically significant pathways (FDR < 0.05): (i) three pathways for interferon signaling, (ii) three related to the immune system, (iii) four related to antigen presentation, (iv) one associated with ER-Phagosomes, (v) one lymphoid-related, and (vi) three pertaining to interleukin-12 signaling. In particular, the 130 alpha, beta, and gamma signaling pathways all appear in the interferon signaling pathways. The immune system pathways include one cytokine signaling and one related to the adaptive immune system. Among the four antigen-related pathways, two are explicitly associated to the dependence of Class I MHC. The 135 Endosomal / Vacuolar pathway implies the involvement of the Class I MHC and of the Antigen processing-Cross presentation. Lastly, interleukin-12 plays a crucial role in the coordination of innate and adaptive immunity [40].

In the C1 community, the Reactome analysis identified 9 statistically significant pathways (FDR < 0.05). Two of these pathways are broadly related to 140 the immune system and cytokine signaling. Another two pathways, the NGF-stimulated transcription and the FOXO-mediated transcription pathways, modulate cell survival, growth, and differentiation. In Table 1 we have listed all the results of the Reactome pathway enrichment analysis for C0 and C1 with FDR < 0.05.

145 Of the communities we observed, the C0 community exhibits the strongest response to the stimulus and immune system, which is evidenced by the very low FDRs  $\sim \mathcal{O}(10^{-14})$ . The complete pathway enrichment analysis for all commu-

nities in saliva is provided in the online data files (ODFs), available on Zenodo, in the "Results/SLV\_results/reactome\_analysis" folder.

150 *2.3. Saliva Communities Temporal Visualization*

We further visualized each community's change over time with heatmaps within the DN network. This is shown for C0 and C1 in Fig.2c. Here, each row denotes a gene, while each column corresponds to a time point of post-treatment. The values plotted in the heatmaps are rescaled gene expression differences between the treated data and the control, and indicate the expression 155 at the particular time point relative to the first time point of the experiment, with rows normalized using Euclidean norm. Red indicates up-regulated genes, blue down-regulated genes, and white indicates unchanged expression. The hierarchical clustering dendrograms revealed relationships among genes at each time point based on the similarity of the gene expressions. The prominent 160 red columns show that genes are upregulated together at these time points. Note that the C0 has a pronounced peak at time point 6, making it an early responding module, while C1 is a late responding cluster, with a pronounced peak at time point 19, as illustrated in Fig. 2c.

165 Here we only show heatmaps for C0 and C1 as representative communities. However, we provide the other communities' heatmaps and with their corresponding Reactome pathway analysis in the ODFs in the folder named "Results/SLV\_results/network\_plots". Our saliva DN has a clear pattern of mostly discrete punctuated gene expression response times for each community. 170 As these punctuated response times, save for one exception (both C0 and C11 show maximized response at t5), are specific to each community, they reflect the biological signatures for individual groups. Most of our saliva DN communities have only one punctuated activation time, although C5 in the saliva DN has 3 up-regulation events at time points 15, 20, and 22 that do not overlap 175 with those of other communities. Between the communities, we observed strong temporally-specific relationships. Our heatmaps are suggestive of the presence of directional signaling between early-activation communities and subsequent

groups, with a potential sequential activation pattern as follows: C6, C9, C8, C2, C0 and C1, C3, C4, C10, C5, C1, and finally C5. At time points from t6 to t10, t14, and from t16 to t18, no communities activated.

#### 2.4. B Cell DN

Our B cell DN consists of 1,759 nodes (genes) and 10,421 edges that we classified into 145 communities using the Louvain algorithm. Similar to the saliva DN, most of these communities are small clusters on small components. Due to its larger size relative to the saliva DN and larger number of communities, our cutoff for plotting was increased to 8 nodes both for community and component size. The global B cell DN is presented in Fig. 3a, with 5 components and 14 communities. Here, we omitted the remaining 130 communities since they neither belong to any of the 5 main components, nor are they large enough for Reactome enrichment analysis. Like in the saliva DN, communities were ordered in descending size (largest to smallest, from C0 to C13 respectively), designated with different colors, and encircled by loops. Fig. 3 has the same format of Fig. 2. In this case, C2 and C4 are displayed in panel b, as magnified representations of the purple cluster and the green cluster, respectively, in panel a. Panel b's magnified perspective provides details about the communities' internal structures. In Fig 3b, for example, we observe that some of the genes highlighted in red form a clique.

#### 2.5. Pathway Enrichment of B Cell Communities

As for the saliva DN, we conducted a community-wise Reactome enrichment analysis for communities with at least 8 genes. 14 communities in the B cell DN were analyzed. This analysis found 9 communities with statistically significant pathway enrichment (FDR < 0.05.): C2, C4, C5, C6, C7, C9, C10, C12, and C13. Most of the pathways associated with genes in these communities centered around transcriptional regulation, protein metabolism, DNA binding ability, and signaling. Among its 111 statistically significant pathways, C4 was

found to be strongly enriched with genes in the FCERI-mediated NF- $\kappa$ B activation pathway, the B cell receptor (BCR) signaling pathway, and the Fc epsilon receptor (FCERI) signaling pathway. These pathways and others relevant to Rituximab mechanism of action, are listed in Table 2. The NF- $\kappa$ B pathway activation by FCERI leads to the production of cytokines during mast cell activation, making it important in allergic inflammatory diseases [41]. C4 also contained a significant number of genes in the B cell receptor pathway, an important pathway related to B cells. The Fc epsilon gene is expressed on antigen-presenting cells, and its signaling occurs on the plasma membrane. A comprehensive list of statistically significant pathways can be found in the ODFs in the "Results/Bcell\_results/reactome\_analysis" folder.

In summary, C4 contains the highest number of responsive pathways which are relevant to the B cell response to Rituximab. As our representative communities, we display the C2 and C4 in Fig. 3 b, our two largest among the 9 communities with significant pathways. Our top 10 pathways based on p-values from the Reactome enrichment analysis for C2 and C4 are listed in Table 2.

### 2.6. B Cell Communities Temporal Visualization

The heatmaps for the temporal behavior for the C2 and C4 communities of the B cell data are shown in Fig. 3c. The formatting of the heatmaps is the same as that of the saliva heatmaps; all values in the heatmaps refer to gene expression relative to time 0 in the treated dataset. The C4's blue column at time point 2 and the less prominent blue column for C2 at time point 15 identify patterns of down-regulation in the two community. While C2 shows a trend of initial up-regulation followed by a gradual diffusion, C4 exhibits an initial down-regulation, followed by later up-regulation.

Though C2 and C4 are our representative communities, we carried out heatmap visualization for all our 9 communities that demonstrated significant pathway enrichment. These heatmaps are available to view in the ODFs in the "Results/Bcell\_results/network\_plots/heatmaps" folder. Overall, in the B cell community heatmaps, we recognized three types of time patterns in terms of

240 collective behavior within an individual community. In the first pattern group, the majority of genes started with a moderate degree of down-regulation. By 7 hours, most instead displayed slight or moderate up-regulation. However, each of these timepoints contained a significant minority of genes with a small level of fluctuation, with the size of the deviating group differing in each heatmap. The second observed time pattern operated in reverse, with most genes beginning upregulated and shifting towards downregulation by the 15-hour mark. Finally, a third group remained consistent in its behavior, with genes trending one way or remaining unchanged across the entire time period.

245 **3. Discussion**

250 Our goal was to determine whether the DN method can identify the activation of biological processes caused by a perturbation. This study applied DN analysis, community identification and Reactome pathway analysis of the DN communities, and identified communities with highly statistically significant enrichment. We analyzed the DNs of two gene expression datasets where a perturbation was applied: (i) Saliva dataset ( PPSV23 vaccination as perturbation; 24 time points), (ii) Primary B-Cells dataset (*ex-vivo* Rituximab drug treatment as perturbation; 6 time points). In summary, our results from the saliva DN revealed pathway activation in immunological and inflammatory responses. In the B cell DN, significant pathways were activated in the regulation of transcription, immune cell survival, activation and differentiation, and inflammatory response. By using the DN method on two separate data sets and comparing our results to known mechanisms of action and target pathways, we can assess the approach's strengths and limitations. These are discussed further 255 below.

260

### 3.1. *PPSV23 Pathway Activations Following Perturbation in Saliva*

*Streptococcus Pneumoniae*'s virulence and associated host immunity have been extensively studied [42]. The PPSV23 is an inactivated vaccine that uses

purified capsular polysaccharides, and is typically administered to older adults  
265 (65+) and susceptible younger individuals [43, 44, 45, 46]. In our assessment we focused on the vaccine's potential pathways of action. Our initial saliva investigation in PPSV23 established that an immune response to the vaccination can be detected utilizing non-invasive saliva monitoring at the molecular level [15]. Since aggregate saliva was sampled, we expected that multiple the multiple  
270 immune cells contained therein are involved in the observed patterns and associated immune responses. Based on our previous findings and general vaccine responses, we anticipated the activation of pathways involved with antigen presentation and processing, regulation of IgM and B/T cells, Lymphoid cells, MHC molecules, and phagocytosis. We also expected the activation of pathways  
275 of general immune response to stimuli or inflammation. To evaluate whether the DN method was as effective as previous studies, we focused on identifying the specific pathways involved.

In our results, a number of expected pathways emerged. These included pathways associated with antigen presentation and processing, Class I MHC  
280 mediated antigen processing and presentation, and ER-phagocytosis, and pathways governing the immunoregulation of interactions between Lymphoid and non-Lymphoid cells [38]. Further results indicative of the participation of immune cells, included the CLEC inflammasome pathway in C4. This pathway is associated with enabling host immune system to mount a fungal/bacterial defense using T-Helper 17 cells (TH17) [47, 48]. Interferon signaling, cytokine  
285 signaling, immune/adaptive immune and interleukin stimulation and signaling are all part of a generalized immune response [49]. We found these more general pathways in the pathway enrichment analysis of C0, C1, C2, C9, and C10. Interferon signaling is crucial in antiviral defense, cell regulation and growth, and  
290 immune response modulation [50]. Our Reactome pathway analysis results are consistent with the results of our saliva multi-omics study [15], which observed that vaccination activates various immune response and regulation pathways, which are also identified in our present results, including ER-Phagosome pathway, Interferon alpha/beta and gamma signaling, cytokine signaling, and MHC

<sup>295</sup> antigen presentation.

### *3.2. Rituximab Pathway Activations Following Perturbation in Primary B Cells*

Regarding our primary B cell results, previous work [51] has established both the biological pathways and the mechanisms of action associated with Rituximab. These previous studies have demonstrated Rituximab's ability to cause <sup>300</sup> antibody-dependent and complement-dependent cellular cytotoxicity, growth inhibition and apoptosis, and regulation of the cell cycle. We also expected to observe Rituximab regulations of the B cell receptor (BCR) based on prior research. Particularly significant among our findings was the enrichment of the nuclear factor  $\kappa$ B (NF- $\kappa$ B) pathways. According to Jazirehi et al. (2005) [23] <sup>305</sup> and Bonavida (2005, 2007) [52, 53], treating NHL B cell lines with Rituximab inhibits NF- $\kappa$ B's signaling pathways by up-regulating RKIP and Raf-1 kinase inhibitors. RKIP has been found to antagonize signal transduction pathways that mediate the NF- $\kappa$ B activation [54].

Following NF- $\kappa$ B's down-regulation due to RKIP's up-regulation, the Bcl-<sup>310</sup> xl expression is also down-regulated. As a result, tumor cells become more chemosensitive. Rituximab also decreased the activity of NF- $\kappa$ B-inducing kinase, I $\kappa$ B kinase, and I $\kappa$ B-a phosphorylation. Finally, the introduction of Rituximab also decreased the activity of the IKK kinase and NF- $\kappa$ B binding to DNA from 3 to 6 hours after treatment [23].

<sup>315</sup> Among the more general enriched pathways observed are signaling pathways that play a role in the molecular mechanisms of chemosensitization, which are also impacted by Rituximab. In line with those effects, we anticipate impacts in the MAPK signaling pathway, the interleukin cytokine regulatory loop, and the Bcl-2 expression. Concerning the expression of genes involved in the healing process, research has uncovered Rituximab's role in affecting pathways associated <sup>320</sup> with immunoglobulin production, chemotaxis, immune response, cell development, and wound healing. Rituximab can also increase existing drug-induced apoptosis [51].

In our community of C4, for example, our Reactome analysis found 5 NF-  
325  $\kappa$ B related pathways with FDR < 0.05. Of these 5 pathways, one is shown in Table 1; the remaining are displayed in the comprehensive table in the "Results/Bcell\_results/reactome\_analysis" folder on Zenodo. Alongside these NF- $\kappa$ B pathways in C4 is the BCR pathway. Our results indicate that the C4 community response is highly relevant because of the activation of both NF- $\kappa$ B  
330 and BCR pathways.

Our C2 community appears to be involved with the metabolism of proteins and cellular responses to external stimuli. Rituximab targets the CD20 B cell transmembrane protein that is involved in B-cell development, activation and proliferation [51]. The C2 community captures cell development pathways which  
335 were included in our expectations of more generalized responses.

We also observed relevant responses in other communities. For example, the C8 community showed activity in the RAF/MAP kinase cascade pathway. In a similar fashion, C10 demonstrated CD22 mediated BCR regulation, classical antibody-mediated complement activation, FCGR activation, antigen activation  
340 of the BCR, and initial complement triggering, etc. The pathways that emerged in our results are thus consistent and highly overlap with established pathways from previous studies. This suggests the effectiveness of our DN method.

### 3.3. Perturbation Induces Temporal Responses

Communities aid in defining the genes' collective behavior, and observing  
345 the collective behavior of communities in the entire network can clarify relative trends between these collective behaviors. The generated heatmaps for each community depicted gene regulation for individual time points, and also displayed trends over time within the identified communities. The trends we observed in our saliva data were consistent with a time-dependent regulation. The  
350 results suggest a sequence of communities activations (up- and down-regulation) at individual timepoints, indicative of sequential immune system responses due to the PPSV23 vaccination. In the primary B cell data were less clear, as fewer time points were monitored, and also the network was more densely connected.

The B-cell heatmaps still indicate overall trends associated with Rituximab activation (both up- and down-regulation) within the first 7 hrs of the treatment.  
355 Our future work will focus on the possibility of establishing a causal chained signaling response, and associated pathways across these communities.

#### 3.4. Conclusion

Our analysis confirmed the applicability of a DN approach in evaluating time course RNA-seq data. Specifically, the DN method showed results in the saliva experiment data consistent with our previous work on profiling PPSV23 vaccination responses [15]. For the primary B cell responses to Rituximab, the DN has found the same signaling pathway as numerous prior experimental results, thus helping with our validation from a computational perspective. The  
360 DN approach complements prior studies by offering a systems-level network perspective of aggregate temporal changes due to drug activation. In future work we plan to address the identification of sequential activation of network communities, as well determining directionality/causality in such activations.

### 4. Limitations of the Study

370 Though our analysis identified multiple pathways relevant to Rituximab activation in the primary B cell data, heatmaps trends were not as distinct as those obtained from the saliva experiment, with weaker and less structured signals from the B cells. One major factor that may be contributing to these somewhat diffuse responses could be the nature of the *ex-vivo* experiment from which the  
375 data were obtained. Isolation in an *ex vivo* environment curtails interactions between Rituximab and aspects of the immune system that are difficult to measure efficiently using existing methods. This is in contrast to *in vivo* settings in which B cells have the ability to interact with those immune system factors. A key difference between the B cell data and the saliva data is that the latter  
380 were obtained under *in vivo* conditions , and thus reflect biological reality. In general, *ex vivo* experimental data are less accurate in summarizing the effects than those of *in vivo* experiments.

In addition, while the saliva experiment covered 24 time points over as many hours, the B cell experiment covered 6 time points over 15 hours. The sampling for the B cell dataset was both less frequent compared to the saliva dataset and unevenly spaced, thereby not accounting for the longer intervals during which 385 no data were recorded.

Regarding the results from the saliva experiment, the data describes a bulk behavior from a tissue containing a mixture of different cell types instead of 390 a single cell type. In principle, single cell RNA-seq data may provide better representation of dynamics and pathways involved in the response compared to a bulk RNA-seq dataset, and elucidate the temporal behavior of different individual cells involved (though currently such studies would also be limited to a pseudotime approach as each cell is only sampled once).

395 Finally, our DN method did not use the time information embedded in our time series dataset. Correlation based approaches do not respect time ordering. Other methods like the Bayesian method and the causality inference method and may be helpful in determining the directionality of the edges in the gene network. As discussed above, we anticipate future work utilizing such methods 400 may enable us to provide deeper information on the causal rewiring of the gene signaling network.

## 5. Author Contributions

Conceptualization, S.X, C.P. and G.I.M.; Methodology, S.X, C.P. and G.I.M.; 405 Software, S.X. and G.I.M.; Investigation, S.X., L.R.K.R., M.Z., J.H, C.P. and G.I.M.; Visualization, S.X.; Resources, S.X., L.R.K.R., C.P. and G.I.M.; Writing – Original Draft, S.X., C.P. and G.I.M.; Writing – Review & Editing, All authors; Funding Acquisition, C.P. and G.I.M.; Supervision, C.P. and G.I.M.

## 6. Acknowledgements

This work was supported by the Translational Research Institute for Space 410 Health through NASA Cooperative Agreement NNX16AO69A (project T0412).

C.P also acknowledges support from NIH/NIGMS (R01GM122085).

## **7. Declaration of Interests**

C.P. owns equity in Salgomed, Inc. G.I.M. has consulted for Colgate Palmolive North America. S.X., L.R.K.R., M.Z. and J.H. declare the absence of  
<sup>415</sup> any commercial or financial relationships that could be construed as a potential conflict of interest.

## **8. Figure titles and legends**

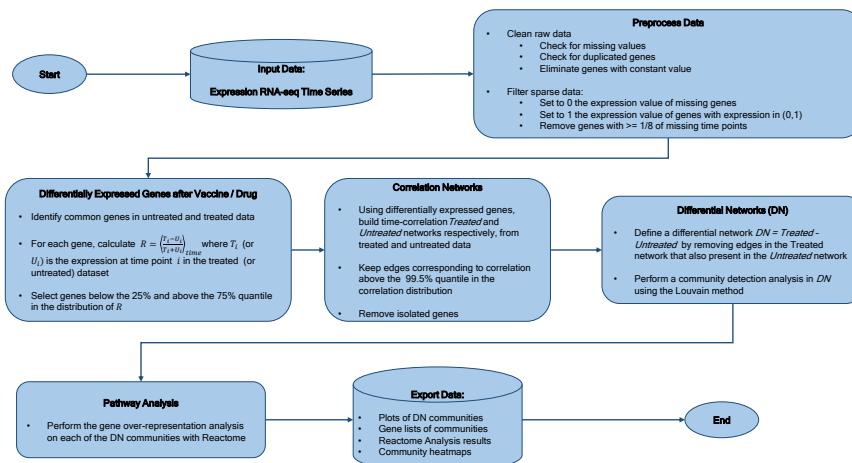
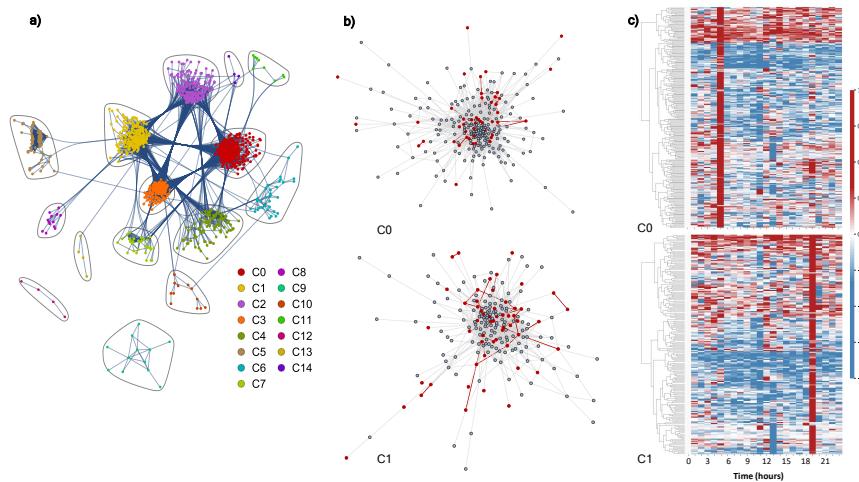
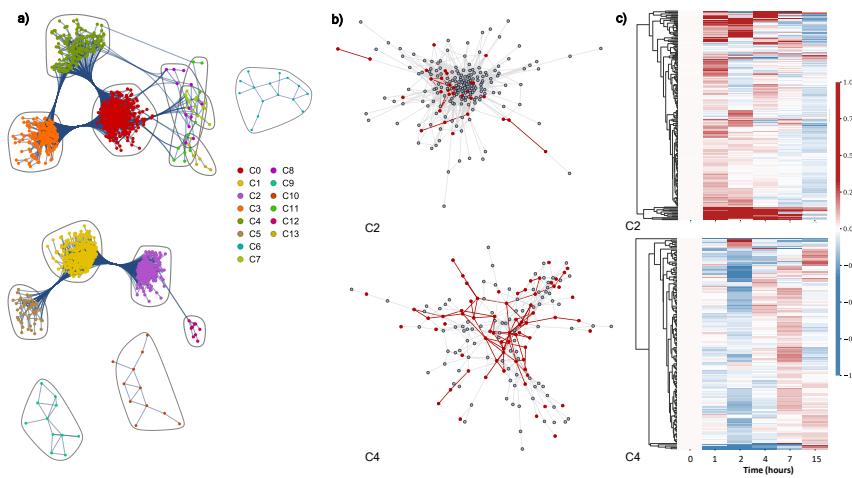


Figure 1: **Workflow Overview**. Our methodology, starts with time course experimental data, followed by network construction, differential network determination, community detection, pathway analyses of individual communities, and final results including analyses and temporal trend visualizations.



**Figure 2: Differential network analysis for the saliva experiment.** a) Differential network with community structure found by the Louvain community detection method. b) Isolated visualizations of C0 (top) and C1 (bottom) communities with red highlights corresponding to genes found in significant Reactome pathways. c) Heatmaps of C0 (top) and C1 (bottom) over 24 hours. Columns represent time points while rows denote gene identifiers. These row data demonstrate the difference in each entry's expression relative to time 0. The relative values were determined by subtracting the individual time points from time point 0 and then normalized using a Euclidean norm, so that each row ranges from -1 (down-regulation) to 1 (up-regulation). For the dendrogram clustering we used the complete-linkage method (also known as the Farthest Point Algorithm) [55, 56].



**Figure 3: Differential network analysis for the B cell experiment.** a) Differential network with community structure found by the Louvain community detection method. b) Isolated visualizations of C2 (top) and C4 (bottom) communities with red highlights corresponding to significant nodes (genes) and their edges (correlations). c) Heatmaps of C2 (top) and C4 (bottom) over 15 hours (6 time points). Columns represent time points while rows denote genes. These row data demonstrate the difference in each entry's expression relative to time 0. The relative values were determined by subtracting the individual time points from time point 0 and then normalized using a Euclidean norm, so that each row ranges from -1 (down-regulation) to 1 (up-regulation). For the dendrogram clustering we used the complete-linkage method (Farthest Point Algorithm) [55, 56].

## 9. Tables and legends

Saliva DN: C0		
Pathway name	Entities FDR	Submitted entities found
Antigen Presentation: Folding, assembly and peptide loading of class I MHC	1.2E-14	HLA-B;NAA15
Endosomal/Vacuolar pathway	1.2E-14	HLA-B
Interferon gamma signaling	1.2E-14	STAT1;IRF1;HLA-B;PTPN6
Class I MHC mediated antigen processing & presentation	1.2E-14	PSMD8;TLR1;CDH1;RPN1;GBF1;HLA-B;UBRA4;CYBA;NAA15;ELOC;FBXO32;FBXO11
ER-Phagosome pathway	1.2E-14	PSMD8;TLR1;RPN1;HLA-B
Interferon alpha/beta signaling	1.2E-14	STAT1;IRF1;HLA-B;PTPN6
Interferon Signaling	1.2E-14	EIF4A1;STAT1;IRF1;HLA-B;PTPN6;ARIH1
Antigen processing-Cross presentation	1.2E-14	PSMD8;TLR1;RPN1;HLA-B;CYBA
Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell	1.2E-14	CDH1;CD81;HLA-B;FCGR2B
Cytokine Signaling in Immune system	1.1E-11	EIF4A1;STAT1;IRF1;HLA-B;PTPN6;ARIH1;CD81;TCF25;RPN1;GBF1;HLA-B;UBRA4;CYBA;PPP2R5D;FBXO32;FBXO11;ANKRD9;TLR1;PSMD8;CDH1;AKT2;PTPN6;ELOC;NA15;FCGR2B;SIP1;ARF5
Adaptive Immune System	4.1E-09	CCDC71L;DX3Y;EIF4A1;ASAH1;IL1RN;SERPINA1;TCF25;CD81;RPN1;RPLP0;UBRA4;TNFAIP3;CSF2RA;PLD2;PSMD8;ANKRD9;CDH1;AKT2;OLR1;ELOC;ARIH1;SERPINB2;TNFSF14;GSTO1;STAT1;GBF1;HLA-B;CYBA;PPP2R5D;FBXO32;FBXO11;FGR;CEACAM3;CLEC4A;TLR1;IRF1;TCP1;TNXIP;PTPN6;CYSTM1;NA15;FCGR2B;SIP1;BIRC2;ARF5;TRIM56
Immune System	1.3E-05	PINB2;TNFSF14;GSTO1;STAT1;GBF1;HLA-B;CYBA;PPP2R5D;FBXO32;FBXO11;FGR;CEACAM3;CLEC4A;TLR1;IRF1;TCP1;TNXIP;PTPN6;CYSTM1;NA15;FCGR2B;SIP1;BIRC2;ARF5;TRIM56
Gene and protein expression by JAK-STAT signaling after Interleukin-12 stimulation	3.2E-03	SERPINB2;GSTO1;TCP1;RPLP0;ARF5
Interleukin-12 family signaling	4.6E-03	SERPINB2;GSTO1;STAT1;TCP1;RPLP0;ARF5
Interleukin-12 signaling	8.0E-03	SERPINB2;GSTO1;TCP1;RPLP0;ARF5
Saliva DN: C1		
Pathway name	Entities FDR	Submitted entities found
Insulin-like Growth Factor-2 mRNA Binding Proteins (IGF2BPs/IMPs/VICKZs) bind RNA	0.003	CD44
Nuclear Events (kinase and transcription factor activation)	0.018	PPP2CB;TF;ID2;CHD4;FOS;DUSP6;DNM2
FOXO-mediated transcription of cell death genes	0.018	BCL2L11;BCL6;NFYC
Signaling by NTRKs	0.024	PPP2CB;RALA;TF;ID2;CLTA;FURIN;CHD4;FOS;DUSP6;DNM2
Signaling by NTRK1 (TRKA)	0.026	PPP2CB;RALA;TF;ID2;CLTA;CHD4;FOS;DUSP6;DNM2
Immune System	0.029	NAPA;RALA;CITA;AHCY1;RPN2;UNC93B1;ADE1;CLTA;BCL10;CFP;TANK;GNS;FCAR;STK10;PPP2CB;BCL2L11;TRIM29;ALOX5;NLRP3;FLNA;SIRPA;SLC12A6;L6;R;GBP4;RAP1GAP2;DDX17;CR1;WSB1;CISH;SH2D3C;KLHL21;FNDC3A;FOS;ILLRB3;MTOR;DUSP6;VEGFA;DNM2;TF;ZNF1X;NASP;BCL6;MAN2B1;TAC2;CD300C;CALM1;CD44;LGMN;RALA;CITA;CISH;RPN2;SH2D3C;FNDC3A;FOS;MTOR;DUSP6;VEGFA;PPP2CB;ZNF1X;BCL2L11;NASP;BCL6;TRIM29;ALOX5;FLNA;IL6R;GBP4;CD44
Cytokine Signaling in Immune system	0.036	DUSP6;VEGFA;PPP2CB;ZNF1X;BCL2L11;NASP;BCL6;TRIM29;ALOX5;FLNA;IL6R;GBP4;CD44
trans-Golgi Network Vesicle Budding	0.039	NAPA;CPD;CLTA;GNS;CLNT1;DNM2
NGF-stimulated transcription	0.039	TF;ID2;CHD4;FOS;DNM2

Table 1: Reactome pathway enrichment analysis. Statistically significant pathways are summarized for saliva DN community C0 and C1. In the full analysis, we omitted small communities with fewer than 8 genes [57], and 12 communities (C0 to C11) qualified for the pathway analysis.

B cell DN: C2		
Pathway name	Entities FDR	Submitted entities found
Peptide chain elongation	1.31462E-06	EEF1A1;RPL4;RPL7A;RPL27A;RPS6;RPL36;RPL14;RPS20;RPL15;FAU;UBA52;RPL28
Response of EIF2AK4 (GCN2) to amino acid deficiency	1.31462E-06	RPL4;RPL7A;RPL27A;RPS6;RPL36;RPL14;RPS20;FAU;RPL15;UBA52;RPL28;ATF3
Eukaryotic Translation Elongation	1.58667E-06	EEF1A1;RPL4;RPL7A;RPL27A;RPS6;RPL36;RPL14;RPS20;RPL15;FAU;UBA52;RPL28
GTP hydrolysis and joining of the 60S ribosomal subunit	5.88845E-06	RPL4;EIF4A1;RPL7A;RPL27A;RPS6;RPL36;RPL14;RPS20;RPL15;FAU;UBA52;RPL28
L13a-mediated translational silencing of Ceruloplasmin expression	5.88845E-06	EIF4A1;RPL7A;RPL27A;RPS6;RPL36;RPL14;RPS20;RPL15;FAU;UBA52;RPL28
Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC)	5.88845E-06	RPL4;RPL7A;RPL27A;RPS6;RPL36;RPL14;RPS20;RPL15;FAU;UBA52;RPL28
Formation of a pool of free 40S subunits	7.65112E-06	RPL4;RPL7A;RPL27A;RPS6;RPL36;RPL14;RPS20;RPL15;FAU;UBA52;RPL28
Eukaryotic Translation Termination	7.65112E-06	RPL4;RPL7A;RPL27A;RPS6;RPL36;RPL14;RPS20;RPL15;FAU;UBA52;RPL28
Cap-dependent Translation Initiation	9.21986E-06	EIF4A1;RPL7A;RPL27A;RPS6;RPL36;RPL14;RPS20;RPL15;FAU;UBA52;RPL28
Eukaryotic Translation Initiation	9.21986E-06	EIF4A1;RPL7A;RPL27A;RPS6;RPL36;RPL14;RPS20;RPL15;FAU;UBA52;RPL28
B cell DN: C4		
Pathway name	Entities FDR	Submitted entities found
Metabolism of RNA	0.014280864	SF3B4;MT-ND6;NUP205;UTP3;POP1;DDX23;CSTF2;PHAX;PLRG1;DIEXF;ZFP36L1;FTSJ3;CHERP;PSMD8;EFTUD2;PSMD9;PSMC4;PSME3;NUP35;SKIV2L2
Mitotic Anaphase	0.014280864	PSMD8;PSMD9;NUP205;CCNB1;SPAST;PSMC4;PSME3;NUP35;SMC1A;EMD;KPNB1
Mitotic Metaphase and Anaphase	0.014280864	PSMD8;PSMD9;NUP205;CCNB1;SPAST;PSMC4;PSME3;NUP35;SMC1A;EMD;KPNB1
FCER1 mediated NF- $\kappa$ B activation	0.014280864	IGLV2-11;PSMD8;PSMD9;IGKV2-29;IGKV1-16;PSMC4;PSME3;IGKV4-1
Signaling by the B Cell Receptor (BCR)	0.014280864	IGLV2-11;PSMD8;PSMD9;IGKV2-29;IGKV1-16;PSMC4;PSME3;IGKV4-1;PIK3AP1
Fc epsilon receptor (FCER1) signaling	0.014280864	IGLV2-11;PSMD8;PSMD9;IGKV2-29;IGKV1-16;PSMC4;PSME3;IGKV4-1
Host Interactions of HIV factors	0.014280864	PSMD8;PSMD9;NUP205;PSMC4;PSME3;NUP35;KPNB1
G1/S Transition	0.014280864	PSMD8;PSMD9;CCNB1;MCM7;PSMC4;PSME3;KPNB1
ABC-family proteins mediated transport	0.014280864	PSMD8;PSMD9;PSMC4;PSME3;CSTF2;EIF2S1
Assembly of the pre-replicative complex	0.014280864	PSMD8;PSMD9;MCM7;PSMC4;PSME3

Table 2: Reactome pathway enrichment analysis. Statistically significant pathways are summarized for primary B cell DN community C2 and C4. In the full analysis, we omitted small communities with fewer than 8 genes [57], and 12 communities (C0 to C11) qualified for the pathway analysis.

## 10. STAR Methods

### 420 RESOURCE AVAILABILITY

#### *Lead contact*

- Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Shuyue Xue (xueshuy1@msu.edu)

#### *Materials availability*

425

- This study did not generate new unique reagents.

#### *Data and code availability*

- This paper analyzes existing, publicly available data. These accession numbers for the datasets are listed in the key resources table.
- Mapped RNA-seq data have been deposited at Zenodo and are publicly available as of the date of publication. DOIs are listed in the key resources table.
- All original code has been deposited on Zenodo as of the date of publication. DOI is listed in the key resources table.
- All results files have been deposited on Zenodo as of the date of publication together with the code files. These files are referred to as Online Data Files (ODFs) in the manuscript. DOI is listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

430

435

## METHOD DETAILS

### 440 *Data acquisition*

Data for this investigation were obtained from Gene Expression Omnibus (GEO) for two time series studies using RNA-seq experiments, on Saliva (accession GSE108664) and Rituximab (GSE100441). Both sets of data are further described below. The raw RNA-seq data were mapped using Kallisto [58],

445 with bootstrap sample parameter,  $-b$ , was set to 100. GENCODE[36] v28 transcripts and genome built GRCh38.p12 were used for annotation. We used Sleuth [59](with DESeq[60] adjustment of Transcripts per Million) to compile results across timepoints.

450 The saliva dataset was obtained from our previous study of immune responses to the PPSV23 vaccine (GSE108664) [15]. In this study, hourly saliva samples were collected from a healthy individual over two 24 hour periods and profiled with RNA-seq every hour. The first 24 hour period provides a baseline RNA expression dataset, which we call *untreated* data. In the second 24 hour period, the same individual was monitored after receiving the PPSV23 455 vaccine. Saliva samples were again collected hourly over 24 hours and profiled by RNA-seq. This second step yielded the RNA expression dataset after the PPSV23 vaccination. We call these data the *treated* dataset. Both treated and untreated datasets have 24 time points of 84,647 possible expression signals using GENCODE annotation [36].

460 The perturbation in the primary B cell experiment was Rituximab, a monoclonal antibody drug used in the treatment of different types of lymphomas and leukemias. The experimental study (data from GSE100441) began by culturing in parallel primary B cells with and without Rituximab. During the 15 hours of Rituximab treatment, the treated and untreated primary B cells were both sampled 465 at the same 6 time points simultaneously and profiled by RNA-seq. The untreated group provided a baseline, which we call untreated data, whereas the treated experiment produced the treated dataset. Since this study included a replicated experiment, each of the first and second duplicates was processed to generate a separate network.

470 *Data Preprocessing*

For quality control, we pre-processed the experimental data and filtered sparse gene signals right after importing the published data files. We coded all the data analysis in Python in this study. Using Python's pandas package [61, 62], we checked for missing values for each gene's expression, removing

475 duplicate records and eliminating genes with constant values across all the 24 time points for the saliva dataset (6 time points for the B cell datasets).

We replaced missing signals with zero and also set values less than 1 to 1. Genes with zero variance in their time series were excluded in our analysis. Moreover, we considered a gene signal as sparse and removed it if its time series 480 had missing values for more than 1/8 of the time points. The same quality control procedure was used for both the saliva and primary B cell datasets.

#### *Gene Selection*

After quality control, we further processed the data to pre-screen and identify a pool of candidate genes that showed a significant response to the perturbations 485 (vaccine in the saliva and Rituximab in the B cell). We selected genes that are significantly expressed in both untreated and treated cases. For each of these genes, we calculated the the time averaged relative difference between treated and untreated normalized intensities,  $\Delta_{TU}$ :

$$\Delta_{TU} = \langle \frac{T_i - U_i}{T_i + U_i} \rangle_{time} \quad (1)$$

This calculation yielded a  $\Delta_{TU}$  distribution curve, from which we computed 490 the lower and upper quartiles. Genes were assumed to either be down-regulated or upregulated if their  $\Delta_{TUs}$  were within the bottom 25% or top 25% of the  $\Delta_{TU}$  distribution respectively. The Python Pandas package was used for all the above computations [61, 62].

#### *Correlation Networks Construction*

495 After gene selection, we calculated their pairwise Pearson correlation coefficients and built the co-expression networks. Genes were represented as nodes and were joined by edges if there was a non-zero correlation between them. We used the correlation coefficient as a weight for each edge. In the layout representation of the networks, the node-node distance reflects their correlation 500 coefficients. Two genes are nearby if they have a high positive correlation. They are far apart when they have a low positive correlation or remote if negatively

correlated. We used Python’s open source Networkx package [63] for network visualization and calculation of the network metrics.

To keep only the most significant signals, we kept edges only in the 99.5%  
505 quantile of the correlation distribution. This process removed many links on the network, detaching some nodes from others. We excluded those isolated nodes from the network. For the saliva data, we built one treated and one untreated network. Since we have data from two repeated experiments for B-cells, we built two networks for the Rituximab treatment and two networks for  
510 the untreated control. Then, we took the intersections between the two networks corresponding to the repeats to obtain a single Rituximab-treated network and one single control network.

#### *Differential Networks Construction*

We defined the DN as the control network subtracted from the treated networks both for the saliva and B cell case. The subtraction removed edges existing in both treated and untreated networks and preserved those only present in the treated network. This operation generates some isolated nodes, that were discarded.  
515

We analyzed the DN’s structure using modularity [25, 26], as complex biological networks usually display a high degree of modularity [29]. Specifically, we employed the Louvain community detection method [34], one of the best-known algorithms for its efficiency, to partition the entire DN into smaller clusters, also known as communities. The Louvain method has been integrated into a published Python package [64], and we used this existing implementation. This  
520 graph clustering algorithm is not deterministic, and can therefore result in different partitions for the same graph.  
525

#### *Enrichment Analysis and Heatmaps*

We conducted Reactome Enrichment Analysis [35] on each community to identify over-represented biological pathways within each community. Reactome  
530 Enrichment Analysis was performed using the Python package PyIomics [37].

As the majority of communities with low number of genes yield no enrichment, we focused on results for communities with 8 or more genes. Furthermore, we plotted the heatmaps for individual communities to visualize gene expression levels as time series, exploring the communities' collective behavior as a function 535 of time. On the heatmaps, we plotted differences in gene expression relative to time 0 (first time point). We computed these values by subtracting each gene's individual time points from time 0. Each time series in our heatmaps was normalized using the Euclidean norm. For dendrogram clustering of rows (genes), we applied the complete-linkage method (Farthest Point Algorithm) 540 [55, 56].

#### *Results formatting and visualization*

We stored the DN nodes and edges, communities, and pathway enrichment analyses into spreadsheets that are provided in the ODFs both for the saliva and B cell data. Using Mathematica [65], we visualized the saliva and B cell 545 DNs with their major connected components and communities.

## QUANTIFICATION AND STATISTICAL ANALYSIS

The data were quantified as discussed in the Method Details section above. Statistical considerations included: (i) selection of a correlation cutoff in network edge construction, based on a 0.995 quantile of the correlation distribution. (ii) 550 An FDR  $< 0.05$  cutoff was used to assess statistical significance of Reactome pathway enrichment analysis, generate through the use of the Reactome API in PyIOmica[35, 37].

## References

[1] U. Nagalakshmi, Z. Wang, K. Waern, C. Shou, D. Raha, M. Gerstein, 555 M. Snyder, The transcriptional landscape of the yeast genome defined by rna sequencing, *Science* 320 (5881) (2008) 1344–1349. doi:10.1126/science.1158441.

560 [2] J. C. Marioni, C. E. Mason, S. M. Mane, M. Stephens, Y. Gilad, RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays, *Genome research* 18 (9) (2008) 1509–1517. doi:10.1101/gr.079558.108.

565 [3] D. Spies, C. Ciaudo, Dynamics in transcriptomics: advancements in RNA-seq time course and downstream analysis, *Computational and structural biotechnology journal* 13 (2015) 469–477. doi:10.1016/j.csbj.2015.08.004.

570 [4] B. T. Wilhelm, J.-R. Landry, RNA-Seq—quantitative measurement of expression through massively parallel RNA-sequencing, *Methods* 48 (3) (2009) 249–257. doi:10.1016/j.ymeth.2009.03.016.

575 [5] Z. Bar-Joseph, A. Gitter, I. Simon, Studying and modelling dynamic biological processes using time-series gene expression data, *Nature Reviews Genetics* 13 (8) (2012) 552–564. doi:<https://doi.org/10.1038/nrg3244>.

580 [6] D. Spies, P. F. Renz, T. A. Beyer, C. Ciaudo, Comparative analysis of differential gene expression tools for rna sequencing time course data, *Briefings in bioinformatics* 20 (1) (2019) 288–298. doi:10.1093/bib/bbx115.

585 [7] S. Oh, S. Song, G. Grabowski, H. Zhao, J. P. Noonan, Time series expression analyses using RNA-seq: a statistical approach, *BioMed research international* 2013. doi:10.1155/2013/203681.

590 [8] S. Oh, S. Song, N. Dasgupta, G. Grabowski, The analytical landscape of static and temporal dynamics in transcriptome data, *Frontiers in genetics* 5 (2014) 35. doi:10.3389/fgene.2014.00035.

595 [9] T. Ideker, N. J. Krogan, Differential network biology, *Molecular systems biology* 8 (1) (2012) 565. doi:10.1038/msb.2011.99.

600 [10] M. J. Ha, V. Baladandayuthapani, K.-A. Do, Dingo: differential network analysis in genomics, *Bioinformatics* 31 (21) (2015) 3413–3420. doi:10.1093/bioinformatics/btv406.

[11] T.-H. Hsiao, Y.-C. Chiu, P.-Y. Hsu, T.-P. Lu, L.-C. Lai, M.-H. Tsai, T. H.-M. Huang, E. Y. Chuang, Y. Chen, Differential network analysis reveals the genome-wide landscape of estrogen receptor modulation in hormonal cancers, *Scientific reports* 6 (1) (2016) 1–16. doi:10.1038/srep23035.

590 [12] K. Mitra, A.-R. Carvunis, S. K. Ramesh, T. Ideker, Integrative approaches for finding modular structure in biological networks, *Nature Reviews Genetics* 14 (10) (2013) 719–732. doi:10.1038/nrg3552.

595 [13] L. Cabusora, E. Sutton, A. Fulmer, C. V. Forst, Differential network expression during drug and stress response, *Bioinformatics* 21 (12) (2005) 2898–2905.

[14] A. de la Fuente, From ‘differential expression’ to ‘differential networking’—identification of dysfunctional regulatory networks in diseases, *Trends in genetics* 26 (7) (2010) 326–333. doi:10.1016/j.tig.2010.05.001.

600 [15] G. I. Mias, V. V. Singh, L. R. Rogers, S. Xue, M. Zheng, S. Domanskyi, M. Kanada, C. Piermarocchi, J. He, Longitudinal saliva omics responses to immune perturbation: a case study, *Scientific Reports* 11 (1) (2021) 1–20. doi:10.1038/s41598-020-80605-6.

605 [16] J. C. Edwards, L. Szczepański, J. Szechiński, A. Filipowicz-Sosnowska, P. Emery, D. R. Close, R. M. Stevens, T. Shaw, Efficacy of b-cell-targeted therapy with rituximab in patients with rheumatoid arthritis, *New England Journal of Medicine* 350 (25) (2004) 2572–2581. doi:10.1056/NEJMoa032534.

[17] X. Bosch, M. Ramos-Casals, M. A. Khamashta, Drugs targeting B-cells in autoimmune diseases, Springer, 2014.

610 [18] World Health Organization, World Health Organization model list of essential medicines: 21st list 2019, Technical documents (2019).

[19] T. Shaw, J. Quan, M. Totoritis, B cell therapy for rheumatoid arthritis: the rituximab (anti-cd20) experience, *Annals of the rheumatic diseases* 62 (suppl 2) (2003) ii55–ii59. doi:10.1136/ard.62.suppl\_2.ii55.

615 [20] P. Johnson, M. Glennie, The mechanisms of action of rituximab in the elimination of tumor cells, in: *Seminars in oncology*, Vol. 30, Elsevier, 2003, pp. 3–8. doi:10.1053/sonc.2003.50025.

620 [21] R. A. Clynes, T. L. Towers, L. G. Presta, J. V. Ravetch, Inhibitory fc receptors modulate in vivo cytotoxicity against tumor targets, *Nature medicine* 6 (4) (2000) 443–446. doi:10.1038/74704.

625 [22] A. R. Jazirehi, X.-H. Gan, S. De Vos, C. Emmanouilides, B. Bonavida, Rituximab (anti-cd20) selectively modifies bcl-xl and apoptosis protease activating factor-1 (apaf-1) expression and sensitizes human non-hodgkin's lymphoma b cell lines to paclitaxel-induced apoptosis, *Molecular cancer therapeutics* 2 (11) (2003) 1183–1193.

URL <https://mct.aacrjournals.org/content/2/11/1183>

630 [23] A. R. Jazirehi, S. Huerta-Yepez, G. Cheng, B. Bonavida, Rituximab (chimeric anti-cd20 monoclonal antibody) inhibits the constitutive nuclear factor- $\kappa$ b signaling pathway in non-hodgkin's lymphoma b-cell lines: role in sensitization to chemotherapeutic drug-induced apoptosis, *Cancer research* 65 (1) (2005) 264–276.

URL <https://cancerres.aacrjournals.org/content/65/1/264>

635 [24] M. I. Vega, A. R. Jazirehi, S. Huerta-Yepez, B. Bonavida, Rituximab-induced inhibition of yy1 and bcl-xl expression in ramos non-hodgkin's lymphoma cell line via inhibition of nf- $\kappa$ b activity: role of yy1 and bcl-xl in fas resistance and chemoresistance, respectively, *The Journal of Immunology* 175 (4) (2005) 2174–2183. doi:10.4049/jimmunol.175.4.2174.

640 [25] M. Girvan, M. E. Newman, Community structure in social and biological networks, *Proceedings of the national academy of sciences* 99 (12) (2002) 7821–7826. doi:10.1073/pnas.122653799.

[26] M. E. J. Newman, Modularity and community structure in networks, *Proceedings of the National Academy of Sciences* 103 (23) (2006) 8577–8582. [doi:10.1073/pnas.0601602103](https://doi.org/10.1073/pnas.0601602103).

[27] S. Fortunato, Community detection in graphs, *Physics reports* 486 (3-5) (2010) 75–174. [doi:10.1016/j.physrep.2009.11.002](https://doi.org/10.1016/j.physrep.2009.11.002).

[28] M. A. Porter, J.-P. Onnela, P. J. Mucha, et al., Communities in networks, *Notices of the AMS* 56 (9) (2009) 1082–1097.  
URL <https://www.ams.org/journals/notices/200909/rtx090901082p.pdf>

[29] L. H. Hartwell, J. J. Hopfield, S. Leibler, A. W. Murray, From molecular to modular cell biology, *Nature* 402 (6761) (1999) C47–C52. [doi:10.1038/35011540](https://doi.org/10.1038/35011540).

[30] A. W. Rives, T. Galitski, Modular organization of cellular networks, *Proceedings of the national Academy of sciences* 100 (3) (2003) 1128–1133. [doi:10.1073/pnas.0237338100](https://doi.org/10.1073/pnas.0237338100).

[31] N. Gulbahce, S. Lehmann, The art of community detection, *BioEssays* 30 (10) (2008) 934–938. [doi:10.1002/bies.20820](https://doi.org/10.1002/bies.20820).

[32] J.-D. J. Han, Understanding biological functions through molecular networks, *Cell research* 18 (2) (2008) 224–237. [doi:10.1038/cr.2008.16](https://doi.org/10.1038/cr.2008.16).

[33] G. Calderer, M. L. Kuijjer, Community detection in large-scale bipartite biological networks, *Frontiers in Genetics* 12 (2021) 520. [doi:10.3389/fgene.2021.649440](https://doi.org/10.3389/fgene.2021.649440).

[34] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, *Journal of statistical mechanics: theory and experiment* 2008 (10) (2008) P10008. [doi:10.1088/1742-5468/2008/10/p10008](https://doi.org/10.1088/1742-5468/2008/10/p10008).

[35] D. Croft, G. O'Kelly, G. Wu, R. Haw, M. Gillespie, L. Matthews, M. Caudy, P. Garapati, G. Gopinath, B. Jassal, et al., Reactome: a database of reactions, pathways and biological processes, *Nucleic acids research* 39 (suppl\_1) (2010) D691–D697. doi:10.1093/nar/gkq1018.

[36] J. Harrow, A. Frankish, J. M. Gonzalez, E. Tapanari, M. Diekhans, F. Kokocinski, B. L. Aken, D. Barrell, A. Zadissa, S. Searle, et al., Gencode: the reference human genome annotation for the encode project, *Genome research* 22 (9) (2012) 1760–1774. doi:10.1101/gr.135350.111.

[37] S. Domanskyi, C. Piermarocchi, G. I. Mias, Pyiomica: longitudinal omics analysis and trend identification, *Bioinformatics* 36 (7) (2020) 2306–2307. doi:10.1093/bioinformatics/btz896.

[38] G. Joshi-Tope, M. Gillespie, I. Vastrik, P. D'Eustachio, E. Schmidt, B. de Bono, B. Jassal, G. Gopinath, G. Wu, L. Matthews, et al., Nucleic acids research 33 (suppl\_1) (2005) D428–D432. doi:10.1093/nar/gki072.

[39] L. Matthews, G. Gopinath, M. Gillespie, M. Caudy, D. Croft, B. de Bono, P. Garapati, J. Hemish, H. Hermjakob, B. Jassal, et al., Reactome knowledgebase of human biological pathways and processes, *Nucleic acids research* 37 (suppl\_1) (2009) D619–D622. doi:10.1093/nar/gkn863.

[40] W. T. Watford, M. Moriguchi, A. Morinobu, J. J. O'Shea, The biology of il-12: coordinating innate and adaptive immune responses, *Cytokine & growth factor reviews* 14 (5) (2003) 361–368. doi:10.1016/s1359-6101(03)00043-1.

[41] S. Klemm, J. Gutermuth, L. Hültner, T. Sparwasser, H. Behrendt, C. Peschel, T. W. Mak, T. Jakob, J. Ruland, The bcl10–malt1 complex segregates fc $\epsilon$ ri-mediated nuclear factor  $\kappa$ b activation and cytokine production from mast cell degranulation, *The Journal of experimental medicine* 203 (2) (2006) 337–347. doi:10.1084/jem.20051982.

695 [42] L. R. Brooks, G. I. Mias, Streptococcus pneumoniae's virulence and host  
immunity: aging, diagnostics, and prevention, *Frontiers in immunology* 9  
(2018) 1366. doi:10.3389/fimmu.2018.01366.

700 [43] US Food and Drug Administration, Pneumovax 23 Prescribing Information  
(Accessed Sept 18th, 2021).  
URL <https://www.fda.gov/downloads/BiologicsBloodVaccines/Vaccines/ApprovedProducts/UCM257088.pdf>

705 [44] J. C. Butler, R. F. Breiman, J. F. Campbell, H. B. Lipman, C. V. Broome,  
R. R. Facklam, Pneumococcal polysaccharide vaccine efficacy. an evaluation  
of current recommendations, *JAMA* 270 (15) (1993) 1826–31. doi:10.  
1001/jama.1993.03510150060030.

710 [45] P. Smit, D. Oberholzer, S. Hayden-Smith, H. J. Koornhof, M. R. Hille-  
man, Protective efficacy of pneumococcal polysaccharide vaccines, *JAMA*  
238 (24) (1977) 2613–6. doi:10.1001/jama.1977.03280250039019.

715 [46] M. W. Pletz, U. Maus, N. Krug, T. Welte, H. Lode, Pneumococcal vaccines:  
mechanism of action, impact on epidemiology and adaption of the species,  
International journal of antimicrobial agents 32 (3) (2008) 199–206. doi:  
10.1016/j.ijantimicag.2008.01.021.

720 [47] S. I. Gringhuis, T. M. Kaptein, B. A. Wevers, B. Theelen, M. Van  
Der Vlist, T. Boekhout, T. B. Geijtenbeek, Dectin-1 is an extracellular  
pathogen sensor for the induction and processing of il-1 $\beta$  via a noncanonical  
caspase-8 inflammasome, *Nature immunology* 13 (3) (2012) 246–254.  
doi:10.1038/ni.2222.

725 [48] S.-C. Cheng, F. L. van de Veerdonk, M. Lenardon, M. Stoffels, T. Plantinga,  
S. Smeekens, L. Rizzetto, L. Mukaremera, K. Preechasuth, D. Cavalieri,  
et al., The dectin-1/inflammasome pathway is responsible for the induction  
of protective t-helper 17 responses that discriminate between yeasts and  
hyphae of candida albicans, *Journal of leukocyte biology* 90 (2) (2011)  
357–366. doi:10.1189/jlb.1210702.

[49] A. K. Abbas, A. H. Lichtman, S. Pillai, Cellular and molecular immunology E-book, Elsevier Health Sciences, 2014.

725 [50] C. A. Bonjardim, P. C. Ferreira, E. G. Kroon, Interferons: signaling, antiviral and viral evasion, *Immunology letters* 122 (1) (2009) 1–11. doi: [10.1016/j.imlet.2008.11.002](https://doi.org/10.1016/j.imlet.2008.11.002).

730 [51] N. Seyfizadeh, N. Seyfizadeh, J. Hasenkamp, S. Huerta-Yepez, A molecular perspective on rituximab: a monoclonal antibody for b cell non hodgkin lymphoma and other affections, *Critical reviews in oncology/hematology* 97 (2016) 275–290. doi: [10.1016/j.critrevonc.2015.09.001](https://doi.org/10.1016/j.critrevonc.2015.09.001).

735 [52] A. R. Jazirehi, B. Bonavida, Cellular and molecular signal transduction pathways modulated by rituximab (rituxan, anti-cd20 mab) in non-hodgkin's lymphoma: implications in chemosensitization and therapeutic intervention, *Oncogene* 24 (13) (2005) 2121–2143. doi: [10.1038/sj.onc.1208349](https://doi.org/10.1038/sj.onc.1208349).

740 [53] B. Bonavida, Rituximab-induced inhibition of antiapoptotic cell survival pathways: implications in chemo/immuno-resistance, rituximab unresponsiveness, prognostic and novel therapeutic interventions, *Oncogene* 26 (25) (2007) 3629–3636.

745 [54] K. C. Yeung, D. W. Rose, A. S. Dhillon, D. Yaros, M. Gustafsson, D. Chatterjee, B. McFerran, J. Wyche, W. Kolch, J. M. Sedivy, Raf kinase inhibitor protein interacts with nf- $\kappa$ b-inducing kinase and tak1 and inhibits nf- $\kappa$ b activation, *Molecular and cellular biology* 21 (21) (2001) 7207–7217. doi: [10.1128/MCB.21.21.7207-7217.2001](https://doi.org/10.1128/MCB.21.21.7207-7217.2001).

[55] D. Defays, An efficient algorithm for a complete link method, *The Computer Journal* 20 (4) (1977) 364–366. doi: [10.1093/comjnl/20.4.364](https://doi.org/10.1093/comjnl/20.4.364).

[56] J. A. Hartigan, Statistical theory in clustering, *Journal of classification* 2 (1) (1985) 63–76. doi: [10.1007/BF01908064](https://doi.org/10.1007/BF01908064).

750 [57] J. Reimand, R. Isserlin, V. Voisin, M. Kucera, C. Tannus-Lopes, A. Ros-  
tamianfar, L. Wadi, M. Meyer, J. Wong, C. Xu, et al., Pathway enrichment  
analysis and visualization of omics data using g: Profiler, gsea,  
cytoscape and enrichmentmap, *Nature protocols* 14 (2) (2019) 482–517.  
doi:10.1038/nature09146.

755 [58] N. L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal probabilistic  
RNA-seq quantification, *Nature biotechnology* 34 (5) (2016) 525–527. doi:  
10.1038/nbt.3519.

760 [59] H. Pimentel, N. L. Bray, S. Puente, P. Melsted, L. Pachter, Differential  
analysis of rna-seq incorporating quantification uncertainty, *Nat Methods*  
14 (7) (2017) 687–690. doi:10.1038/nmeth.4324.  
URL <https://www.ncbi.nlm.nih.gov/pubmed/28581496>

765 [60] S. Anders, W. Huber, Differential expression analysis for sequence  
count data, *Genome Biol* 11 (10) (2010) R106. doi:10.1186/gb-2010-11-10-r106.  
URL <https://www.ncbi.nlm.nih.gov/pubmed/20979621>

[61] The Pandas development team, *pandas-dev/pandas: Pandas* (feb 2020).  
doi:10.5281/zenodo.3509134.

770 [62] Wes McKinney, Data Structures for Statistical Computing in Python, in: Stéfan van der Walt, Jarrod Millman (Eds.), *Proceedings of the 9th Python in Science Conference*, 2010, pp. 56 – 61. doi:10.25080/Majora-92bf1922-00a.

775 [63] A. A. Hagberg, D. A. Schult, P. J. Swart, Exploring network structure,  
dynamics, and function using networkx, in: G. Varoquaux, T. Vaught,  
J. Millman (Eds.), *Proceedings of the 7th Python in Science Conference*,  
Pasadena, CA USA, 2008, pp. 11 – 15.

[64] T. Bonald, N. de Lara, Q. Lutz, B. Charpentier, *Scikit-network: Graph*  
analysis in python, *Journal of Machine Learning Research* 21 (185) (2020)

1–6.

URL <http://jmlr.org/papers/v21/20-412.html>

<sup>780</sup> [65] Wolfram Research, Inc., Mathematica, Version 12.0.0, champaign, IL, 2021.

URL <https://www.wolfram.com/mathematica>