



DATA NOTE

# A curated transcriptome dataset collection to investigate the blood transcriptional response to viral respiratory tract infection and vaccination. [version 1; peer review: 2 approved]

Salim Bougarn , Sabri Boughorbel , Damien Chaussabel , Nico Marr

Systems Biology and Immunology Department, Sidra Medicine, Doha, Qatar

**v1** First published: 13 Mar 2019, 8:284 (<https://doi.org/10.12688/f1000research.18533.1>)  
 Latest published: 13 Mar 2019, 8:284 (<https://doi.org/10.12688/f1000research.18533.1>)

**Abstract**


The human immune defense mechanisms and factors associated with good versus poor health outcomes following viral respiratory tract infections (VRTI), as well as correlates of protection following vaccination against respiratory viruses, remain incompletely understood. To shed further light into these mechanisms, a number of systems-scale studies have been conducted to measure transcriptional changes in blood leukocytes of either naturally or experimentally infected individuals, or in individual's post-vaccination. Here we are making available a public repository, for research investigators for interpretation, a collection of transcriptome datasets obtained from human whole blood and peripheral blood mononuclear cells (PBMC) to investigate the transcriptional responses following viral respiratory tract infection or vaccination against respiratory viruses. In total, Thirty one31 datasets, associated to viral respiratory tract infections and their related vaccination studies, were identified and retrieved from the NCBI Gene Expression Omnibus (GEO) and loaded in a custom web application designed for interactive query and visualization of integrated large-scale data. Quality control checks, using relevant biological markers, were performed. Multiple sample groupings and rank lists were created to facilitate dataset query and interpretation. Via this interface, users can generate web links to customized graphical views, which may be subsequently inserted into manuscripts to report novel findings. The GXB tool enables browsing of a single gene across projects, providing new perspectives on the role of a given molecule across biological systems in the diagnostic and prognostic following VRTI but also in identifying new correlates of protection. This dataset collection is available at: <http://vri1.gxbsidra.org/dm3/geneBrowser/list>.

**Keywords**

Transcriptomics, Bioinformatics, Software, Viral respiratory infection, Influenza viruses, Respiratory syncytial viruses (RSV), Rhinoviruses, Whole Blood, PBMC.

**Open Peer Review**

Reviewer Status  

|  | Invited Reviewers   |   |
|--|---|---|
|  | 1   | 2   |
| <b>version 1</b><br>published<br>13 Mar 2019 | <br>report | <br>report |

- 1 **Carlos Alberto Moreira-Filho**,  
University of São Paulo (USP), São Paulo, Brazil
- 2 **Benjamin M Tang**, Nepean Hospital,  
Kingswood, Australia

Any reports and responses or comments on the article can be found at the end of the article.



This article is included in the Sidra Medicine gateway.



This article is included in the **Data: Use and Reuse** collection.

**Corresponding author:** Salim Bougarn ([sbougarn@sidra.org](mailto:sbougarn@sidra.org))

**Author roles:** **Bougarn S:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Boughorbel S:** Data Curation, Resources, Software, Writing – Review & Editing; **Chaussabel D:** Conceptualization, Funding Acquisition, Project Administration, Supervision, Writing – Review & Editing; **Marr N:** Conceptualization, Project Administration, Supervision, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** All the authors listed on this publication received support from the Qatar Foundation. Support for this project was provided by the Qatar National Research Fund [NPRP10-0205-170348].

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2019 Bougarn S *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Bougarn S, Boughorbel S, Chaussabel D and Marr N. **A curated transcriptome dataset collection to investigate the blood transcriptional response to viral respiratory tract infection and vaccination. [version 1; peer review: 2 approved]** F1000Research 2019, **8**:284 (<https://doi.org/10.12688/f1000research.18533.1>)

**First published:** 13 Mar 2019, **8**:284 (<https://doi.org/10.12688/f1000research.18533.1>)

## Introduction

Viral respiratory tract infections (VRTI) are responsible for the majority hospitalizations among infants and the elderly. They are caused mainly by a heterogeneous group of viruses, including rhinoviruses, influenza viruses, parainfluenza viruses, respiratory syncytial virus (RSV), enteroviruses, coronaviruses, and certain strains of adenovirus<sup>1,2</sup>. Few antiviral therapies are currently approved and routinely used for VRTI. Most of these are specific inhibitors of influenza viruses<sup>3</sup>. Moreover, for most respiratory viruses, there is no licensed vaccine available<sup>4,5</sup>, with the exception of flu vaccines for which protection generally lasts only one flu season. Consequently, clinical management of individuals with VRTI is mostly restricted to supportive care<sup>5</sup>.

As clinical symptoms are often overlapping and are not specific for any of the viral species, it is difficult to establish a clinical diagnosis without laboratory testing<sup>4</sup>. Furthermore, clinical manifestations of VRTI are highly variable, ranging from asymptomatic infections or illness with mild symptoms (a common cold) to clinically severe disease with life-threatening complications, such as respiratory failure and in some cases may have a fatal outcome<sup>6</sup>. Infants, the elderly and patients with chronic lung or heart diseases in particular are at high risk<sup>7</sup>.

Thus, there is an evident need to better understand the molecular mechanisms underlying the disease pathogenesis, progression as well as severity of, and immunity against, VRTI among humans<sup>8</sup>. In this context, different large scale gene expression studies have been conducted using whole blood or peripheral blood mononuclear cells (PBMCs), to assess the human immune response to natural<sup>9-11</sup> and experimental viral respiratory infections<sup>12,13</sup>; in particular, to influenza and RSV infections, and also to vaccination<sup>14-17</sup>.

Here, we make available, through an interactive web application, a curated collection of datasets that were obtained from pediatric and adult patients with natural VRTI, volunteers with experimental exposition to respiratory viruses and also vaccinated volunteers. Transcriptomics datasets were obtained from whole blood and PBMCs.

A total of 31 datasets were retrieved and selected from the NCBI [Gene Expression Omnibus](#) (GEO), a public repository of transcriptome profiles. The identified datasets are particularly relevant to our interest in understanding the pathobiology of VRTI and vaccination. As described in recent publications<sup>18,19</sup>, these datasets were loaded into a custom interactive web application, the [Gene Expression Browser](#) (GXB), which enables easy access to large datasets and interactive visualization of our dataset collection related to VRTI and vaccination against respiratory viruses. It also provides access to demographic and clinical information. Importantly, the user can customize data plots by adding multiple layers of parameters (e.g. age, gender, sample type, type of infection, type of vaccine, sample collection time), modify the sample ordering and genes, and generate links (mini URL) that can be shared via e-mail or used in publications. Therefore, we are providing here a resource enabling browsing of datasets relevant to blood transcriptional responses to VRTI and vaccination that offers a unique opportunity to identify

host genes and their regulation that may be of diagnostic and/or prognostic value, or that may be tested as novel correlates of protection in subsequent studies. For example, a comparative approach of the transcriptional response signatures between experimentally infected and vaccinated individuals could be used to identify common mechanisms that define the poor health outcomes versus strong protection. The ability to pool, compare and analyze the immune responses to different infections and vaccines, in different individuals and at various age, offers a unique opportunity for a better understanding of the pathophysiology of VRTI.

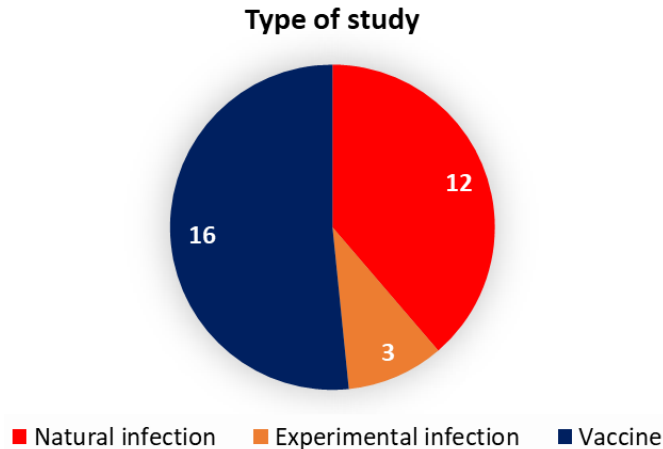
## Methods

A total of 120 datasets, potentially relevant to human immune responses to VRTI and vaccination, were identified in [GEO](#) using the following search query:

```
Homo sapiens[Organism] AND (“respiratory syncytial virus”[DESC] OR RSV[DESC] OR metapneumovirus[DESC] OR hMPV[DESC] OR influenza[DESC] OR parainfluenza[DESC] OR rhinovirus[DESC] OR rhinoviruses[DESC] OR adenovirus[DESC] OR adenoviruses[DESC] OR HAdV[DESC] OR coronaviruses[DESC] OR HCoV[DESC]) OR (vaccine OR vaccines OR vaccination) AND (blood[DESC] OR PBMC[DESC] OR PBMCs[DESC] OR lymphocyte[DESC] OR lymphocytes[DESC] OR “B cell”[DESC] OR “B cells”[DESC] OR “plasma cells”[DESC] OR “T cell”[DESC] OR “T cells”[DESC] OR Treg[DESC] OR Tregs[DESC] OR monocyte[DESC] OR monocytes[DESC] OR dendritic[DESC] OR DC[DESC] OR DCs[DESC] OR “natural killer”[DESC] OR NK[DESC] OR NKT[DESC] OR neutrophil[DESC] OR neutrophils[DESC]) AND (“Expression profiling by array”[gdsType] OR “Expression profiling by high throughput sequencing”[gdsType]).
```

Most of retrieved datasets were generated from human blood and human PBMC, using Illumina or Affymetrix commercial platforms or RNA-sequencing. All the entries that were returned with this query were manually curated. The process involved reading all the descriptions available of the datasets, the study design and the GEO-linked article in pubmed. Finally, only studies using human whole blood and human PBMCs, associated with natural or experimental VRTI, or vaccination against VRTI, were retained for our dataset collection. For the retained datasets, if the platform used to generate the transcriptome profiles was not supported by GXB or if from an *in vitro* study, they were excluded from our dataset collection. Based on these criteria, 31 datasets were retained. These include datasets that were generated from whole blood or PBMCs of individuals who were either naturally (12) or experimentally infected (3) (with influenza viruses, RSV, Rhinovirus, Rotavirus) as well as from healthy, uninfected (age-matched) volunteers. The remaining 16 datasets were generated from whole blood or PBMCs of individuals who had received flu vaccines ([Figure 1](#)). The datasets that comprise our collection are listed in [Table 1](#).

Once the final selection had been made, each dataset was downloaded from GEO by using the SOFT file format. Then, the datasets were uploaded on the Gene Expression Browser (GXB), an interactive web application hosted on the Amazon Web



**Figure 1. Break down of the dataset collection by category.** The pie chart indicates the type of studies carried out for the 31 datasets.

Services cloud<sup>20</sup>. Information about samples and study design were also uploaded. The available samples were put into groups based on relevant study variables and genes were ranked according to the different groups comparisons. A detailed description of the GXB software tool is available from recent publications<sup>19–21</sup>. This software interface allows user to easily navigate and filter the dataset collection. A web tutorial can be easily [accessed online](#). Annotation and functionality of the web software interface were described previously by our group<sup>18,19,21</sup>, and is reproduced here so that readers can use this article as a standalone resource. Briefly, datasets of interest can be quickly identified either by filtering on criteria from pre-defined sections on the left or by entering a query term in the search box at the top of the dataset navigation page. Clicking on one of the studies listed in the dataset navigation page opens a viewer designed to provide interactive browsing and graphic representations of large-scale data in an interpretable format. This interface is designed to present ranked gene lists and display expression results graphically in a context-rich environment. Selecting a gene from the rank ordered list on the left of the data-viewing interface will display its expression values graphically in the screen's central panel. Directly above the graphical display drop down menus give users the ability: a) To change how the gene list is ranked - this allows the user to change the method used to rank the genes, or to only include genes that are selected for specific biological interest; b) To change sample grouping (Group Set button) - in some datasets, a user can switch between groups based on cell type to groups based on disease type, for example; c) To sort individual samples within a group based on associated categorical or continuous variables (e.g. gender or age); d) To toggle between the bar chart view and a box plot view, with expression values represented as a single point for each sample. Samples are split into the same groups whether displayed as a bar chart or box plot; e) To provide a color legend for the sample groups; f) To select categorical information that is to be overlaid at the bottom of the graph - for example, the user can display gender or smoking status in this manner; g) To provide a color legend for the categorical information overlaid at the bottom of the graph; h) To download the graph as a portable network graphics (png) image. Measurements have no intrinsic utility in absence of contextual

information. It is this contextual information that makes the results of a study or experiment interpretable. It is therefore important to capture, integrate and display information that will give users the ability to interpret data and gain new insights from it. We have organized this information under different tabs directly above the graphical display. The tabs can be hidden to make more room for displaying the data plots, or revealed by clicking on the blue “show info panel” button on the top right corner of the display. Information about the gene selected from the list on the left side of the display is available under the “Gene” tab. Information about the study is available under the “Study” tab. Rolling the mouse cursor over a bar chart feature while displaying the “Sample” tab lists any clinical, demographic, or laboratory information available for the selected sample. Finally, the “Downloads” tab allows advanced users to retrieve the original dataset for analysis outside this tool. It also provides all available sample annotation data for use alongside the expression data in third party analysis software. Other functionalities are provided under the “Tools” drop-down menu located in the top right corner of the user interface. Some of the notable functionalities available through this menu include: a) Annotations, which provides access to all the ancillary information about the study, samples and dataset organized across different tabs; b) Cross-project view; which provides the ability for a given gene to browse through all available studies; c) Copy link, which generates a mini-URL encapsulating information about the display settings in use and that can be saved and shared with others (clicking on the envelope icon on the toolbar inserts the URL in an email message via the local email client); d) Chart options; which gives user the option to customize chart labels.

### Quality Control

Quality control checks can be performed on the datasets loaded on GXB, for example by examining concordance of the gender-specific expression of the *XIST* gene in those datasets for which gender information was available as metadata. The *XIST* gene is essential for imprinted and random X-chromosome inactivation<sup>22</sup> and therefore, expression is expected to be high in female and low in male samples. Respective hyperlinks are found

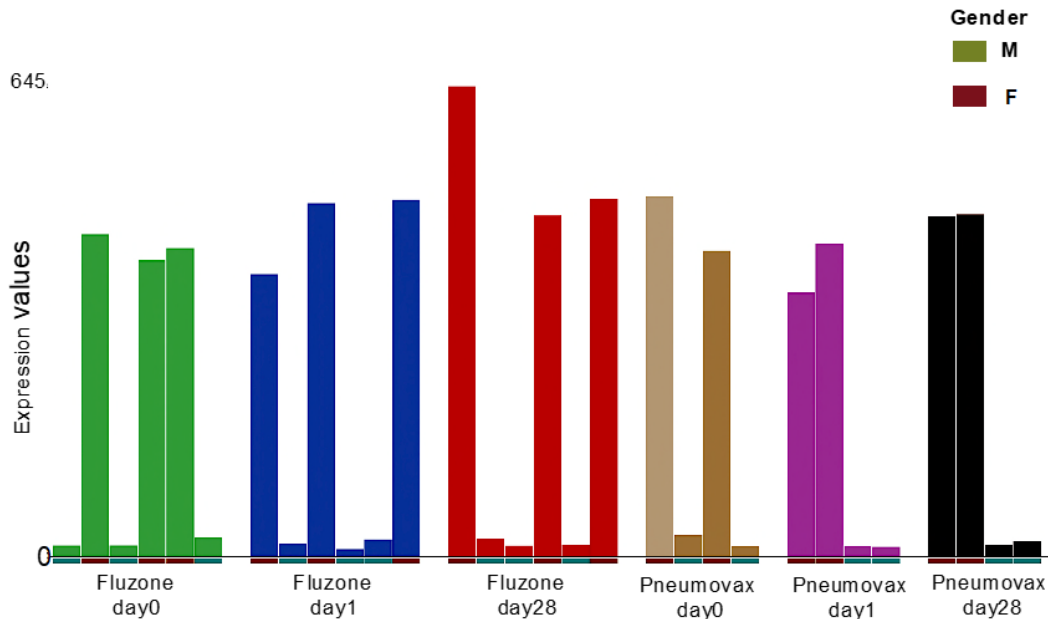
**Table 1. List of datasets constituting the collection.**

| Title  | Platforms used           | Response to                       | Virus/Vaccine                   | Cell type/<br>Tissues | Number of<br>samples | Selected<br>marker for<br>QC* | Citation<br># | GEO ID              |
|--|--------------------------|-----------------------------------|---------------------------------|-----------------------|----------------------|-------------------------------|---------------|---------------------|
| Blood transcriptome of human bacterial and influenza A pneumonia.  | Illumina HumanHT-12 v3   | Natural infection                 | Influenza H1N1                  | Whole blood           | 190                  | Xist                          | 9             | GSE40012            |
| Expression profiling of critically ill influenza and bacterial pneumonia patients, also influenza vaccination recipients.  | Illumina HumanHT-12 v3   | Natural infection and vaccination | Influenza H1N1                  | Whole blood           | 81                   | ND*                           | 14            | GSE20346            |
| FACS-sorted cells from Young Adults Vaccinated with Influenza TIV or LAIV Vaccines during 2009/09 Flu Season   | Affymetrix HG-U133A      | Vaccination                       | Influenza TIV vaccine           | PBMC                  | 84                   | ND*                           | 16            | GSE29618            |
| Gene expression analysis in children with complex seizures by influenza A (H1N1)pdm09 or rotavirus gastroenteritis.  | Affymetrix HG-U133_Plus2 | Natural infection                 | Influenza (H1N1) Rotavirus      | Whole blood           | 32                   | Xist                          | 10            | GSE50628            |
| Gene expression analysis of influenza vaccine response in Young and Old - Year 1.  | Illumina HumanHT-12 v4   | Vaccination                       | Influenza TIV vaccine           | PBMC                  | 72                   | ND*                           | 23            | GSE59635            |
| Gene expression analysis of influenza vaccine response in Young and Old - Year 2   | Illumina HumanHT-12 v4   | Vaccination                       | Influenza TIV vaccine           | PBMC                  | 156                  | ND*                           | 23            | GSE59654            |
| Gene expression signatures of symptomatic respiratory viral infection in adults.   | Affymetrix HG-U133A      | Experimental infection            | Influenza (H3N2) RSV Rhinovirus | Whole blood           | 113                  | ND*                           | 13            | GSE17156            |
| Genome-wide analysis of whole blood transcriptional response to Respiratory Syncytial Virus (RSV), Influenza and Rhinovirus lower respiratory tract infection (LRTI) in children.  | Illumina HumanHT-12 v4   | Natural infection                 | Influenza, RSV, rhinovirus      | Whole blood           | 36                   | Xist                          | 8             | GSE38900 (GPL10558) |
| Genome-wide analysis of whole blood transcriptional response to Respiratory Syncytial Virus, Influenza and Rhinovirus lower respiratory tract infection (LRTI) in children.  | Illumina HumanWG-6 v3    | Natural infection                 | Influenza, RSV, rhinovirus      | Whole blood           | 205                  | Xist                          | 8             | GSE38900 (GPL6884)  |
| Genome-wide profiling of whole blood from healthy adult volunteers before and after receiving non-live vaccines including seasonal influenza or pneumococcal vaccine or placebo (saline) injections I.                               | Illumina HumanHT-12 v3   | Vaccination                       | Influenza TIV vaccine           | Whole blood           | 72                   | Xist                          | 24            | GSE30059            |
| Genome-wide profiling of whole blood from healthy adult volunteers before and after receiving non-live vaccines including seasonal influenza or pneumococcal vaccine or placebo (saline) injections II. Finger prick-Training Set    | Illumina HumanHT-12 v3   | Vaccination                       | Influenza TIV vaccine           | Whole blood           | 185                  | Xist                          | 24            | GSE48762            |
| Genome-wide profiling of whole blood from healthy adult volunteers before and after receiving non-live vaccines including seasonal influenza or pneumococcal vaccine or placebo (saline) injections II. Venipuncture-Training Set.   | Illumina HumanHT-12 v3   | Vaccination                       | Influenza QIV vaccine           | Whole blood           | 214                  | Xist                          | 24            | GSE48762            |
| Genome-wide profiling of whole blood from healthy adult volunteers before and after receiving non-live vaccines including seasonal influenza or pneumococcal vaccine or placebo (saline) injections II. Venipuncture-Validation Set. | Illumina HumanHT-12 v3   | Vaccination                       | Influenza QIV vaccine           | Whole blood           | 30                   | Xist                          | 24            | GSE48762            |

| Title  | Platforms used                           | Response to               | Virus/Vaccine                      | Cell type/<br>Tissues | Number of<br>samples | Selected<br>marker for<br>QC* | Citation<br># | GEO ID                |
|--|--|---------------------------|------------------------------------|-----------------------|----------------------|-------------------------------|---------------|-----------------------|
| Global Analyses of Human Immune Variation Reveal Baseline Predictors of Postvaccination Responses.   | Affymetrix<br>HuGene 1.0 ST<br>v1        | Vaccination               | Influenza<br>QIV vaccine           | PBMC                  | 292                  | ND*                           | 17            | GSE47353              |
| Host gene expression signatures of influenza A H1N1 and H3N2 virus infection in adults.  | Affymetrix HG-<br>U133A                  | Experimental<br>infection | Influenza<br>H1N1, H3N2            | Whole<br>blood        | 649                  | ND*                           | 12            | GSE52428              |
| Host transcriptional response to influenza and other acute respiratory viral infections – a prospective cohort study.                            | Illumina<br>HumanHT-12 v4                | Natural<br>infection      | Influenza<br>H1N1                  | Whole<br>blood        | 880                  | Xist                          | 11            | GSE68310              |
| Hosts responses in critical disease caused by pandemic H1N1.   | Illumina<br>HumanWG-6 v2                 | Natural<br>infection      | Influenza<br>H1N1                  | Whole<br>blood        | 40                   | ND*                           | 25            | GSE21802              |
| Influenzavirus serotype association to global whole blood transcriptional changes.   | Illumina<br>HumanHT-12 v4                | Natural<br>infection      | Influenza<br>H1N1, H3N2            | Whole<br>blood        | 465                  | ND*                           | NA            | GSE29385              |
| Olfactomedin 4 serves as a marker for disease severity in pediatric Respiratory Syncytial Virus (RSV) infection.                                 | Affymetrix HG-<br>U133_Plus_2            | Natural<br>infection      | RSV                                | PBMC                  | 43                   | ND*                           | 26            | GSE69606              |
| Patient-based transcriptome-wide analysis identify interferon and ubiquination pathways as potential predictors of Influenza A disease severity. | Illumina<br>HumanHT-12 v4                | Natural<br>infection      | Influenza<br>H1N1, H3N2            | Whole<br>blood        | 402                  | ND*                           | 27            | GSE61821              |
| Peripheral blood cells expression data form 7 patients with severe pdm(H1N1) influenza and 7 gender and age matched healthy controls.            | Affymetrix<br>HuGene 1.0 ST<br>v1        | Natural<br>infection      | Influenza<br>H1N1                  | Whole<br>blood        | 21                   | ND*                           | 28            | GSE27131              |
| Systems biology of vaccination for seasonal influenza in humans.   | Affymetrix<br>HT_HG-U133_+<br>Plus_PM v1 | Vaccination               | Influenza<br>TIV, LAIV<br>vaccines | PBMC                  | 163                  | ND*                           | 16            | GSE29619              |
| Systems biology of vaccination for seasonal influenza in humans.   | Affymetrix HG-<br>U133A                  | Vaccination               | Influenza<br>TIV, LAIV<br>vaccine  | PBMC                  | 84                   | ND*                           | 16            | GSE29619<br>(GPL3921) |
| T cell responses to H1N1v and a longitudinal study of seasonal influenza vaccination - 2009.   | Illumina<br>HumanHT-12 v4                | Vaccination               | Influenza<br>QIV vaccine           | Whole<br>blood        | 72                   | ND*                           | NA            | GSE58943              |
| T cell responses to H1N1v and a longitudinal study of seasonal influenza vaccination - 2010.   | Illumina<br>HumanHT-12 v4                | Vaccination               | Influenza<br>TIV vaccine           | Whole<br>blood        | 75                   | Xist                          | NA            | GSE58898              |
| T cell responses to H1N1v and a longitudinal study of seasonal influenza vaccination - 2012.   | Affymetrix<br>PrimeView v1               | Vaccination               | Influenza<br>TIV vaccine           | Whole<br>blood        | 71                   | Xist                          | NA            | GSE64514              |
| Time Course of Adults Vaccinated with Influenza TIV Vaccine 2009-2012.   | Affymetrix<br>HT_HG-U133_+<br>Plus_PM v1 | Vaccination               | Influenza<br>TIV vaccine           | Whole<br>blood        | 621                  | ND*                           | 15            | GSE74817              |
| Time series of global gene expression after trivalent influenza vaccination in humans (female cohort).   | Illumina<br>HumanHT-12 v4                | Vaccination               | Influenza<br>TIV vaccine           | Whole<br>blood        | 417                  | Xist                          | 29            | GSE48023              |
| Time series of global gene expression after trivalent influenza vaccination in humans (male cohort).   | Illumina<br>HumanHT-12 v3                | Vaccination               | Influenza<br>TIV vaccine           | Whole<br>blood        | 431                  | Xist                          | 29            | GSE48018              |
| Transcriptional profile of PBMCs in patients with acute RSV or Influenza infection.  | Illumina<br>HumanHT-12 v4                | Natural<br>infection      | Influenza<br>(H1N1, H3N2),<br>RSV  | PBMC                  | 101                  | Xist                          | 30            | GSE34205              |

| Title  | Platforms used            | Response to               | Virus/Vaccine                      | Cell type/<br>Tissues | Number of<br>samples | Selected<br>marker for<br>QC* | Citation<br># | GEO ID   |
|--|---------------------------|---------------------------|------------------------------------|-----------------------|----------------------|-------------------------------|---------------|----------|
| Transcriptomic profiling facilitates classification of response to influenza challenge   | Illumina<br>HumanHT-12 v4 | Experimental<br>infection | Influenza<br>H3N2                  | Whole<br>blood        | 88                   | ND*                           | 31            | GSE61754 |
| Transcriptomic profiling in childhood H1N1/09 influenza reveals reduced expression of protein synthesis genes  | Illumina<br>HumanHT-12 v3 | Natural<br>infection      | Influenza<br>(H1N1), RSV           | Whole<br>blood        | 92                   | X(ist)                        | 32            | GSE42026 |
| Trivalent Inactivated Influenza Vaccine (TIV) and Live Attenuated Influenza Vaccine (LAIV) Induce Different B cell and Transcriptional Responses in Children | Illumina<br>HumanHT-12 v4 | Vaccination               | Influenza<br>TIV, LAIV<br>vaccines | Whole<br>blood        | 140                  | ND*                           | 30            | GSE52005 |
| Whole Blood Transcriptional Response to Pediatric Influenza Infection  | Illumina<br>HumanWG-6 v3  | Natural<br>infection      | Influenza                          | Whole<br>blood        | 31                   | X(ist)                        | 8             | GSE29366 |

\*Gender information and/or X(ist) probe was not available. QC – quality control, ND – no data



**Figure 2.** Shown are *XIST* gene expression levels and gender information from the venipuncture validation set of GSE48762. Gene expression data were from whole blood of healthy adult volunteers before and after receiving either placebo (saline) injections, seasonal influenza (Fluzone) or pneumococcal (Pneumovax) vaccination.

in Table 1 allow you to visualize the *XIST* expression based on the gender information provided with the GEO submission. Figure 2 shows *XIST* gene expression in a representative dataset, along with gender information available that was recorded and made available in GEO.

### Data availability

All datasets included in our curated collection are also available publically via the NCBI GEO website : <https://www.ncbi.nlm.nih.gov/gds/>; and are referenced throughout the manuscript by their GEO accession numbers (e.g. GSE17763). Signal files and sample description files can also be downloaded from the GXB tool under the “downloads” tab.

### Grant information

All the authors listed on this publication received support from the Qatar Foundation. Support for this project was provided by the Qatar National Research Fund [NPRP10-0205-170348].

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

### Acknowledgments

We would like to thank all the investigators who decided to make their datasets publically available by depositing them in GEO.

### References

- Hodinka RL: **Respiratory RNA Viruses.** *Microbiol Spectr.* 2016; 4(4).  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Tan WC: **Viruses in asthma exacerbations.** *Curr Opin Pulm Med.* 2005; 11(1): 21–6.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Ison MG: **Clinical use of approved influenza antivirals: therapy and prophylaxis.** *Influenza Other Respir Viruses.* 2013; 7 Suppl 1: 7–13.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Zhu T, Zhang C, Yu L, *et al.*: **The preventive effect of vaccine prophylaxis on severe respiratory syncytial virus infection: A meta-analysis.** *Virology.* 2015; 50(5): 371–8.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Caballero MT, Polack FP, Stein RT: **Viral bronchiolitis in young infants: new perspectives for management and treatment.** *J Pediatr (Rio J).* 2017; 93 Suppl 1: 75–83.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Jong VL, Ahout IM, van den Ham HJ, *et al.*: **Transcriptome assists prognosis of disease severity in respiratory syncytial virus infected infants.** *Sci Rep.* 2016; 6: 36603.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ackerson B, Tseng HF, Sy LS, *et al.*: **Severe Morbidity and Mortality Associated With Respiratory Syncytial Virus Versus Influenza Infection in Hospitalized Older Adults.** *Clin Infect Dis.* 2018.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Mejias A, Dimo B, Suarez NM, *et al.*: **Whole blood gene expression profiles to assess pathogenesis and disease severity in infants with respiratory syncytial virus infection.** *PLoS Med.* 2013; 10(11): e1001549.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Parnell GP, McLean AS, Booth DR, *et al.*: **A distinct influenza infection signature in the blood transcriptome of patients with severe community-acquired pneumonia.** *Crit Care.* 2012; 16(4): R157.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)



10. Tsuge M, Oka T, Yamashita N, *et al.*: **Gene expression analysis in children with complex seizures due to influenza A(H1N1)pdm09 or rotavirus gastroenteritis.** *J Neurovirol.* 2014; **20**(1): 73–84.  
[PubMed Abstract](#) | [Publisher Full Text](#)
11. Zhai Y, Franco LM, Atmar RL, *et al.*: **Host Transcriptional Response to Influenza and Other Acute Respiratory Viral Infections—A Prospective Cohort Study.** *PLoS Pathog.* 2015; **11**(6): e1004869.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
12. Woods CW, McClain MT, Chen M, *et al.*: **A host transcriptional signature for presymptomatic detection of infection in humans exposed to influenza H1N1 or H3N2.** *PLoS One.* 2013; **8**(1): e52198.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Zaas AK, Chen M, Varkey J, *et al.*: **Gene expression signatures diagnose influenza and other symptomatic respiratory viral infections in humans.** *Cell Host Microbe.* 2009; **6**(3): 207–17.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. Parnell G, McLean A, Booth D, *et al.*: **Aberrant cell cycle and apoptotic changes characterise severe influenza A infection—a meta-analysis of genomic signatures in circulating leukocytes.** *PLoS One.* 2011; **6**(3): e17186.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
15. Nakaya HI, Hagan T, Duraisingham SS, *et al.*: **Systems Analysis of Immunity to Influenza Vaccination across Multiple Years and in Diverse Populations Reveals Shared Molecular Signatures.** *Immunity.* 2015; **43**(6): 1186–98.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
16. Nakaya HI, Wrammert J, Lee EK, *et al.*: **Systems biology of vaccination for seasonal influenza in humans.** *Nat Immunol.* 2011; **12**(8): 786–95.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
17. Tsang JS, Schwartzberg PL, Kotliarov Y, *et al.*: **Global analyses of human immune variation reveal baseline predictors of postvaccination responses.** *Cell.* 2014; **157**(2): 499–513.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
18. Rinchai D, Boughorbel S, Presnell S, *et al.*: **A curated compendium of monocyte transcriptome datasets of relevance to human monocyte immunobiology research [version 2; referees: 2 approved].** *F1000Res.* 2016; **5**: 291.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. Rahman M, Boughorbel S, Presnell S, *et al.*: **A curated transcriptome dataset collection to investigate the functional programming of human hematopoietic cells in early life [version 1; referees: 2 approved].** *F1000Res.* 2016; **5**: 414.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Speake C, Presnell S, Domico K, *et al.*: **An interactive web application for the dissemination of human systems immunology data.** *J Transl Med.* 2015; **13**: 196.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
21. Marr AK, Boughorbel S, Presnell S, *et al.*: **A curated transcriptome dataset collection to investigate the development and differentiation of the human placenta and its associated pathologies [version 2; referees: 2 approved].** *F1000Res.* 2016; **5**: 305.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
22. Cerase A, Pintacuda G, Tattermusch A, *et al.*: **Xist localization and function: new insights from multiple levels.** *Genome Biol.* 2015; **16**: 166.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
23. Thakar J, Mohanty S, West AP, *et al.*: **Aging-dependent alterations in gene expression and a mitochondrial signature of responsiveness to human influenza vaccination.** *Aging (Albany NY).* 2015; **7**(1): 38–52.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
24. Obermoser G, Presnell S, Domico K, *et al.*: **Systems scale interactive exploration reveals quantitative and qualitative differences in response to influenza and pneumococcal vaccines.** *Immunity.* 2013; **38**(4): 831–44.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
25. Bermejo-Martin JF, Martin-Loeches I, Rello J, *et al.*: **Host adaptive immunity deficiency in severe pandemic influenza.** *Crit Care.* 2010; **14**(5): R167.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
26. Brand HK, Ahout IM, Ridder D, *et al.*: **Olfactomedin 4 Serves as a Marker for Disease Severity in Pediatric Respiratory Syncytial Virus (RSV) Infection.** *PLoS One.* 2015; **10**(7): e0131927.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Hoang LT, Tolfvenstam T, Ooi EE, *et al.*: **Patient-based transcriptome-wide analysis identify interferon and ubiquination pathways as potential predictors of influenza A disease severity.** *PLoS One.* 2014; **9**(11): e111640.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
28. Berdal JE, Mollnes TE, Wæhre T, *et al.*: **Excessive innate immune response and mutant D222G/N in severe A (H1N1) pandemic influenza.** *J Infect.* 2011; **63**(4): 308–16.  
[PubMed Abstract](#) | [Publisher Full Text](#)
29. Franco LM, Bucacas KL, Wells JM, *et al.*: **Integrative genomic analysis of the human immune response to influenza vaccination.** *eLife.* 2013; **2**: e00299.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
30. Cao RG, Suarez NM, Obermoser G, *et al.*: **Differences in antibody responses between trivalent inactivated influenza vaccine and live attenuated influenza vaccine correlate with the kinetics and magnitude of interferon signaling in children.** *J Infect Dis.* 2014; **210**(2): 224–33.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
31. Davenport EE, Antrobus RD, Lillie PJ, *et al.*: **Transcriptomic profiling facilitates classification of response to influenza challenge.** *J Mol Med (Berl).* 2015; **93**(1): 105–14.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
32. Herberg JA, Kaforou M, Gormley S, *et al.*: **Transcriptomic profiling in childhood H1N1/09 influenza reveals reduced expression of protein synthesis genes.** *J Infect Dis.* 2013; **208**(10): 1664–8.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

# Open Peer Review

Current Peer Review Status:  

---

## Version 1

Reviewer Report 13 June 2019

<https://doi.org/10.5256/f1000research.20283.r48849>

© 2019 Tang B. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



### Benjamin M Tang

Department of Intensive Care Medicine, Nepean Hospital, Kingswood, NSW, Australia

This paper represents an extension of the authors' extensive work in this area - which aims to provide an accessible web-based platform for researchers to perform multi-cohort analysis of pooled transcriptomics data.

The Introduction outlines the scientific rationale to support the development of such a platform, which I mostly agree. There are a number of reasons why such a platform can assist researchers, including (1) help researchers to leverage on publicly available dataset to address each researcher's own research question, (2) obtaining a large sample size for analysis, which is otherwise not feasible by a single research group, (3) provide access to a diverse range of patient population and varying degree of disease severity, thereby generating more biological insights, which are not achievable by a single-centre study. For these reasons, the platform provided by these author represents an important contribution to the research community.

To further place this into perspective, one can estimate the likely cost (in time and money) to obtain the same analysis if such a platform is not available. To collect 31 transcriptomics datasets of viral respiratory infection (as outlined by this paper), a researcher need to perform an extensive literature review and database search. After all the datasets are collected, the researcher then need to bring in a bioinformatic team to perform a pooled analysis, which required advanced expertise due to the huge variability in technical platform, assay methods and normalization approach among different datasets. This bioinformatic work is extensive - if such work is outsourced to an external bioinformatician/company, the cost is in excessive of US\$15,000 (our group has commissioned similar work in the past - hence our knowledge into the estimated cost). Here, the authors provides access to the analysis platform to the wider research community for free...an important step in democratising scientific data for the benefit of all researchers.

**Is the rationale for creating the dataset(s) clearly described?**

Yes

**Are the protocols appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and materials provided to allow replication by others?**

Yes

**Are the datasets clearly presented in a useable and accessible format?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Genomics, immunology, viral respiratory infection

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Reviewer Report 26 March 2019

<https://doi.org/10.5256/f1000research.20283.r45689>

© 2019 Moreira-Filho C. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Carlos Alberto Moreira-Filho**

Department of Pediatrics, Faculty of Medicine, University of São Paulo (USP), São Paulo, Brazil

The paper by Bougarn et al. describes the use of an interactive web application, GXB (Speake et al.(2015)<sup>1</sup>), for investigating the blood transcriptional response to viral respiratory tract infection and vaccination. A curated dataset collection was established for this purpose encompassing 31 datasets retrieved and selected from the NCBI Gene Expression Omnibus (GEO) public repository.

The rationale for creating the dataset collection and the methodology used in the paper are clearly described. The functionalities of the GXB tool are well explained to the reader, as well as the quality control procedures adopted by the authors. However, the GEO datasets do not always provide detailed information on virus genotypes. Distinct RSV genotypes induce very different PBMC transcriptional responses (see, for instance, Rodriguez-Fernandez et al. 2017<sup>2</sup>; and Vieira et al. 2019<sup>3</sup>). The authors could insert a comment on this issue in their paper.

**References**

1. Speake C, Presnell S, Domico K, Zeitner B, Bjork A, Anderson D, Mason MJ, Whalen E, Vargas O, Popov D, Rinchai D, Jourde-Chiche N, Chiche L, Quinn C, Chaussabel D: An interactive web application for the dissemination of human systems immunology data. *J Transl Med.* 2015; **13**: 196 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Rodriguez-Fernandez R, Tapia L, Yang C, Torres J, Chavez-Bueno S, Garcia C, Jaramillo L, Moore-Clingenpeel M, Jafri H, Peeples M, Piedra P, Ramilo O, Mejias A: Respiratory Syncytial Virus Genotypes, Host Immune Profiles, and Disease Severity in Young Children Hospitalized With

Bronchiolitis. *The Journal of Infectious Diseases*. 2018; **217** (1): 24-34 [Publisher Full Text](#)  
3. Vieira SE, Bando SY, de Paulis M, Oliveira DBL, Thomazelli LM, Durigon EL, Martinez MB, Moreira-Filho CA: Distinct transcriptional modules in the peripheral blood mononuclear cells response to human respiratory syncytial virus or to human rhinovirus in hospitalized infants with bronchiolitis. *PLoS One*. 2019; **14** (3): e0213501 [PubMed Abstract](#) | [Publisher Full Text](#)

**Is the rationale for creating the dataset(s) clearly described?**

Yes

**Are the protocols appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and materials provided to allow replication by others?**

Yes

**Are the datasets clearly presented in a useable and accessible format?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** genomics

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

---

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**