



# A novel immune biomarker *IFI27* discriminates between influenza and bacteria in patients with suspected respiratory infection

Benjamin M. Tang<sup>1,2,3,21</sup>, Maryam Shojaei<sup>1,2,21</sup>, Grant P. Parnell<sup>2</sup>, Stephen Huang<sup>1</sup>, Marek Nalos<sup>1</sup>, Sally Teoh<sup>1</sup>, Kate O'Connor<sup>2</sup>, Stephen Schibeci<sup>2</sup>, Amy L. Phu<sup>1</sup>, Anand Kumar<sup>4</sup>, John Ho<sup>5</sup>, Adrienne F. A. Meyers<sup>5</sup>, Yoav Keynan<sup>6,7,8,9</sup>, Terry Ball<sup>5,9</sup>, Amarnath Pisipati<sup>9</sup>, Aseem Kumar<sup>10</sup>, Elizabeth Moore<sup>11</sup>, Damon Eisen<sup>12</sup>, Kevin Lai<sup>13</sup>, Mark Gillett<sup>14</sup>, Robert Geffers<sup>15</sup>, Hao Luo<sup>16</sup>, Fahad Gul<sup>16</sup>, Jens Schreiber<sup>17</sup>, Sandra Riedel<sup>17</sup>, David Booth<sup>2</sup>, Anthony McLean<sup>1</sup> and Klaus Schughart<sup>18,19,20</sup>

@ERSpublications  
***IFI27* could discriminate between influenza and bacterial infection in suspected respiratory tract infection cases** <http://ow.ly/VEaY30bmlS3>

**Cite this article as:** Tang BM, Shojaei M, Parnell GP, *et al.* A novel immune biomarker *IFI27* discriminates between influenza and bacteria in patients with suspected respiratory infection. *Eur Respir J* 2017; 49: 1602098 [<https://doi.org/10.1183/13993003.02098-2016>].

**ABSTRACT** Host response biomarkers can accurately distinguish between influenza and bacterial infection. However, published biomarkers require the measurement of many genes, thereby making it difficult to implement them in clinical practice. This study aims to identify a single-gene biomarker with a high diagnostic accuracy equivalent to multi-gene biomarkers.

In this study, we combined an integrated genomic analysis of 1071 individuals with *in vitro* experiments using well-established infection models.

We identified a single-gene biomarker, *IFI27*, which had a high prediction accuracy (91%) equivalent to that obtained by multi-gene biomarkers. *In vitro* studies showed that *IFI27* was upregulated by TLR7 in plasmacytoid dendritic cells, antigen-presenting cells that responded to influenza virus rather than bacteria. *In vivo* studies confirmed that *IFI27* was expressed in influenza patients but not in bacterial infection, as demonstrated in multiple patient cohorts (n=521). In a large prospective study (n=439) of patients presented with undifferentiated respiratory illness (aetiologies included viral, bacterial and non-infectious conditions), *IFI27* displayed 88% diagnostic accuracy (AUC) and 90% specificity in discriminating between influenza and bacterial infections.

*IFI27* represents a significant step forward in overcoming a translational barrier in applying genomic assay in clinical setting; its implementation may improve the diagnosis and management of respiratory infection.

---

This article has supplementary material available from [erj.ersjournals.com](http://erj.ersjournals.com)

Received: June 21 2015 | Accepted after revision: March 15 2017

**Support statement:** This work was supported by Nepean Institute of Critical Care Education and Research (Kingswood, Australia). The mouse study was funded by an intra mural grant from the Helmholtz-Association (Program Infection and Immunity) and research grant FluResearchNet (no. 01KI07137) from the German Ministry of Education and Research to Klaus Schughart. Benjamin M. Tang, Grant P. Parnell, Maryam Shojaei and Anthony McLean had filed a patent application (PCT) on *IFI27*. The Nepean Institute of Critical Care Education and Research, and the public health agency of Canada, had no role in: the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript. The corresponding author had full access to all the data and the final responsibility to submit for publication. Funding information for this article has been deposited with the Crossref Funder Registry.

**Conflict of interest:** Disclosures can be found alongside this article at [erj.ersjournals.com](http://erj.ersjournals.com)

Copyright ©ERS 2017

**Affiliations:** <sup>1</sup>Dept of Intensive Care Medicine, Nepean Hospital, Sydney, Australia. <sup>2</sup>Westmead Institute for Medical Research, Centre for Immunology and Allergy Research, Sydney, Australia. <sup>3</sup>Respiratory Virus Infection Research, Marie Bashir Institute for Infectious Diseases and Biosecurity, Sydney, Australia. <sup>4</sup>Section of Critical Care Medicine and Section of Infectious Diseases, Dept of Medicine, Medical Microbiology and Pharmacology, University of Manitoba, Winnipeg, Canada. <sup>5</sup>National Laboratory for HIV Immunology, JC Wilt Infectious Disease Research Centre, Public Health Agency of Canada, Winnipeg, Canada. <sup>6</sup>Dept of Internal Medicine, University of Manitoba, Winnipeg, Canada. <sup>7</sup>Dept of Medical Microbiology, University of Manitoba, Winnipeg, Canada. <sup>8</sup>Dept of Community Health Sciences, University of Manitoba, Winnipeg, Canada. <sup>9</sup>Dept of Medical Microbiology and Infectious Diseases, University of Manitoba, Winnipeg, Canada. <sup>10</sup>Dept of Chemistry and Biochemistry, Laurentian University, Sudbury, Canada. <sup>11</sup>Transfusion Research Unit, Dept of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia. <sup>12</sup>Dept of Infectious Diseases, Townsville Hospital, Townsville, Australia. <sup>13</sup>Dept of Emergency Medicine, Westmead Hospital, Sydney, Australia. <sup>14</sup>Dept of Emergency Medicine, Royal North Shore Hospital, Sydney, Australia. <sup>15</sup>Genome Analytics, Helmholtz Centre for Infection Research, Braunschweig, Germany. <sup>16</sup>Nepean Clinical School, Sydney Medical School, University of Sydney, Sydney, Australia. <sup>17</sup>Clinic of Pneumology, Otto-von-Guerike University of Magdeburg, Magdeburg, Germany. <sup>18</sup>Dept of Infection Genetics, Helmholtz Centre for Infection Research, Braunschweig, Germany. <sup>19</sup>University of Veterinary Medicine, Hannover, Germany. <sup>20</sup>University of Tennessee Health Science Center, Memphis, TN, USA. <sup>21</sup>These authors contributed equally to the study.

**Correspondence:** Benjamin M. Tang, Nepean Genomic Research Group, Dept of Intensive Care Medicine, Nepean Hospital & Nepean Clinical School, University of Sydney, Sydney, Australia.  
E-mail: benjamin.tang@sydney.edu.au

## Introduction

Acute respiratory tract infections are the most common causes of infectious disease morbidity and mortality in the world (WHO - 2014). Viruses and bacteria are the main causes of respiratory infections. Among the viruses, the influenza virus causes the highest levels of morbidity and mortality. Accurately differentiating between the influenza virus and bacteria is the first key step in managing respiratory infections, because initial treatment (antiviral *versus* antibiotic) is guided mainly by the aetiology. However, differentiating between influenza and bacterial infections on clinical grounds is often difficult, because both infections share similar clinical features (*e.g.* cough, dyspnoea and fever).

Pathogen testing (*e.g.* virus detection by PCR) is an important step towards clarifying the diagnosis. Although important, virus detection alone is insufficient to guide patient management. Like most respiratory tract viruses, the influenza virus produces a wide clinical spectrum. In some patients, the virus acts as a bystander and its presence is unrelated to the presenting illness [1]. In others, the virus is directly responsible for the patient's symptoms. Clinical decision-making requires consideration of whether the virus is a bystander or the cause of presenting symptoms, because anti-viral therapy (*e.g.* Oseltamivir) should be administered only to individuals in whom symptoms are severe and directly caused by the virus [2–4]. Currently, no laboratory test exists that can distinguish between virus detection and “active infection”.

Host response biomarkers represent an alternative approach to circumvent this problem. Host gene expression biomarkers could accurately identify the immune response underpinning influenza infection, and thus assist clinicians in distinguishing between influenza and bacterial infections [5–8]. These biomarkers represent a promising strategy to complement current diagnostic methods. However, published biomarker panels contain a large number of genes, making it difficult to implement them in clinical practice. To overcome this translational barrier, we conducted an integrated genomic analysis of 1071 individuals, to search for a single-gene biomarker with high diagnostic accuracy equivalent to that of multi-gene biomarkers in distinguishing between influenza and bacterial infections.

## Materials and methods

### Study design

This study consisted of a discovery phase and a validation phase (figure 1). First, a candidate biomarker was identified in discovery cohorts (n=111). Next, the identified biomarker underwent a two-phase validation. In phase 1 validation, we evaluated the biomarker in four independent, external data sets (n=521) (figure 1). In addition, we performed *in vitro* experiments to understand the mechanism by which immune cells produced the biomarker, and we performed experiments with mice to validate the biomarker *in vivo* (see supplementary material for full details of all experiments). In phase 2 validation, we undertook a large prospective study (n=439) to evaluate real-world performance of the biomarker in patients with suspected respiratory infection.

### Participants

Participants were recruited in discovery cohorts and a prospective validation cohort. The discovery cohorts consisted of three case-control studies. Each study included patients with well-defined phenotypes

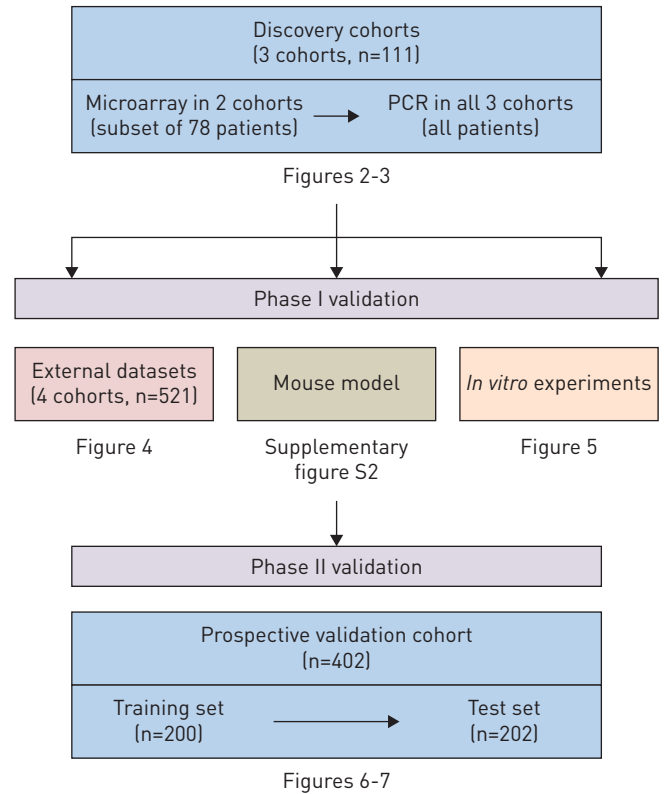


FIGURE 1 Schematic of the study.

(influenza infection, bacterial infection and non-infectious illnesses) and healthy controls. In phase 2 validation, we performed a large, multi-centre study (“prospective validation study”) in patients presenting with “flu-like” illness to emergency departments. The study protocol was approved by the institutional review board of each participating institution and all study participants gave informed consent. Demographic and clinical characteristics of all study participants are summarised in table 1. Detailed information on all cohorts (inclusion criteria, recruitment process, sample collection and follow-up) is presented in the supplementary material.

### Statistical analysis

To evaluate the diagnostic performance of the biomarker, sensitivity, specificity, diagnostic threshold and area under the curve (AUC) of the receiver operating characteristic curve were calculated using established methods (see supplementary material for full details). Comparisons between two groups were calculated using the unpaired two-tailed *t*-test or the non-parametric Mann–Whitney U-test where appropriate. Comparisons between multiple groups were calculated using a one-way ANOVA or the Kruskal–Wallis test, where appropriate. Statistical analysis was performed using the NCSS (LLC, Kaysville, UT, USA) and PRISM (GraphPad Software, La Jolla, CA, USA) software packages.

## Results

### Discovery cohorts

To screen for influenza biomarkers, we analysed the whole blood of influenza patients in three discovery cohorts (table 1 and supplementary table S1). Influenza patients were defined as those with a positive virus result detected on viral nucleic acid testing, who presented with severe illness compatible with influenza that required admission to a hospital or intensive care unit. We first performed microarray analysis to identify differentially expressed genes (between influenza patients and healthy controls) (figure 1). This analysis revealed that among differentially expressed genes, interferon alpha inducible protein 27 (*IFI27*) performed significantly better than all other genes (figure 2). This finding was consistent in two separate discovery cohorts (figure 2) and on different microarray platforms (supplementary figure S1). We subsequently performed PCR validation in all three cohorts, which confirmed that *IFI27* was upregulated in influenza-positive patients (supplementary table S1). Altogether, findings from all discovery cohorts were consistent with each other and collectively suggested that *IFI27* is a potential biomarker of influenza infection.

TABLE 1 Overview of all cohorts

	Mean age years	Males/females n	Clinical setting	Infection severity	Sample size n <sup>#</sup>	Case mix in each cohort n (%)			
						Viral	Bacterial	Non-infectious	Healthy controls
<b>Discovery cohorts</b>									
Australia	51	28/26	Adult ICU	Severe	54	8 (22) <sup>¶</sup>	16 (44) <sup>¶</sup>	12 (33) <sup>¶</sup>	18
Canada	48	9/24	Adult ICU	Severe	33	33 (100)	0	0	0
Germany	61	14/10	Adult ICU and ward	Moderate/severe	24	24 (100)	0	0	0
<b>External cohorts</b>									
GSE6269 [9]	3	81/57	Paediatric ward	Moderate	138	36 (26)	95 (69)	0	7 (5)
GSE42026 [10]	2.8	59/33	Paediatric ICU/ward	Moderate/severe	92	41 (44)	18 (20)	0	33 (36)
GSE60244 [11]	61	89/69	Adult ward	Moderate	158*	71 (45)	22 (14)	0	40 (25)
GSE68310 [12]	NA	NA	Healthy adults	Mild	133	109 (82)	NA	24 (18)	0
<b>Prospective validation cohort</b>	53	180/220	Adult outpatient, ER and ICU	Mild, moderate, severe	439*	162 (40) <sup>¶</sup>	48 (12) <sup>¶</sup>	177 (44) <sup>¶</sup>	37

ICU: intensive care unit; GSE: reference number for publicly available microarray datasets in the National Centre for Biotechnology Information Gene Expression Omnibus (reference for published paper is provided beside each GSE number); NA: not available; ER: emergency room. #: total sample 1071; ¶: percentage calculated based on number of patients, healthy controls excluded; \*: includes additional co-infection cases (not shown in table).

### Host response

To confirm that *IFI27* reflected “active infection” in patients, we examined *IFI27*'s correlation with immune response genes. This analysis showed that *IFI27* was co-expressed with many anti-viral immune response genes. It also showed that *IFI27* was highly correlated (coefficient of correlation >0.6) with 69% of the differentially expressed genes (24 of 35) in the immune system pathway, and 90% of the differentially expressed genes (19 of 21) in the interferon signalling pathway (figure 3). Many of these *IFI27*-correlated genes have anti-viral functions, such as *MX1* (anti-influenza), *ISG15* (inhibits influenza virus replication), *OAS1*, *OAS2* and *OAS3* (degrades viral RNA), *EIF2AK2* (inhibits viral replication), *HERC5* (positive regulator of the anti-viral response) and *IFIT2* (inhibits expression of viral mRNA). These findings confirm that *IFI27* was part of the host response against respiratory viruses.

### External validation

In phase 1 validation, we validated our findings in both mouse and human datasets. Mouse data suggested that *ifi27* was upregulated in influenza infections of varying severity (supplementary figure S2). In human datasets, we analysed microarray datasets of four independent external cohorts (GSE6269, GSE42026, GSE68310 and GSE60244) (table 1). Two cohorts consisted of paediatric patients: GSE6269 (n=131) and

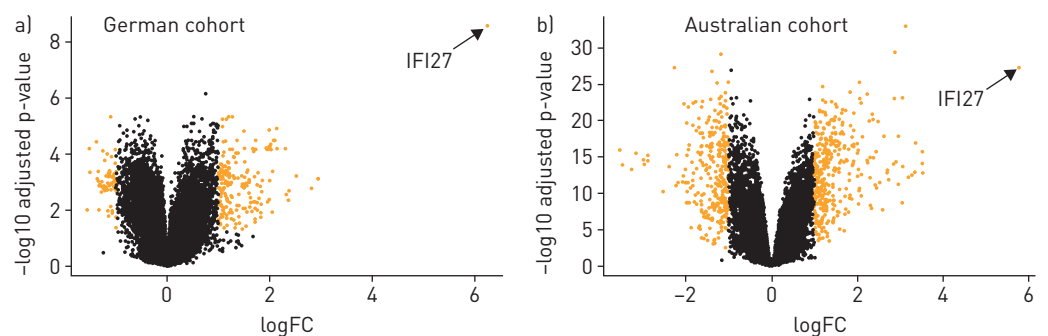


FIGURE 2 *IFI27* performs better than other genes as a biomarker for influenza infection. Microarray data are presented as volcano plots, which arrange genes along dimensions of biological significance (horizontal axis) and statistical significance (vertical axis). The horizontal axis denotes gene expression levels (fold changes) on a log<sub>2</sub> scale between two groups (influenza versus healthy controls). The vertical axis denotes the adjusted p-value on a negative log<sub>2</sub> scale between the groups. The genes with higher expression levels (>2 fold changes) and adjusted p-values (<0.05) are highlighted in yellow.

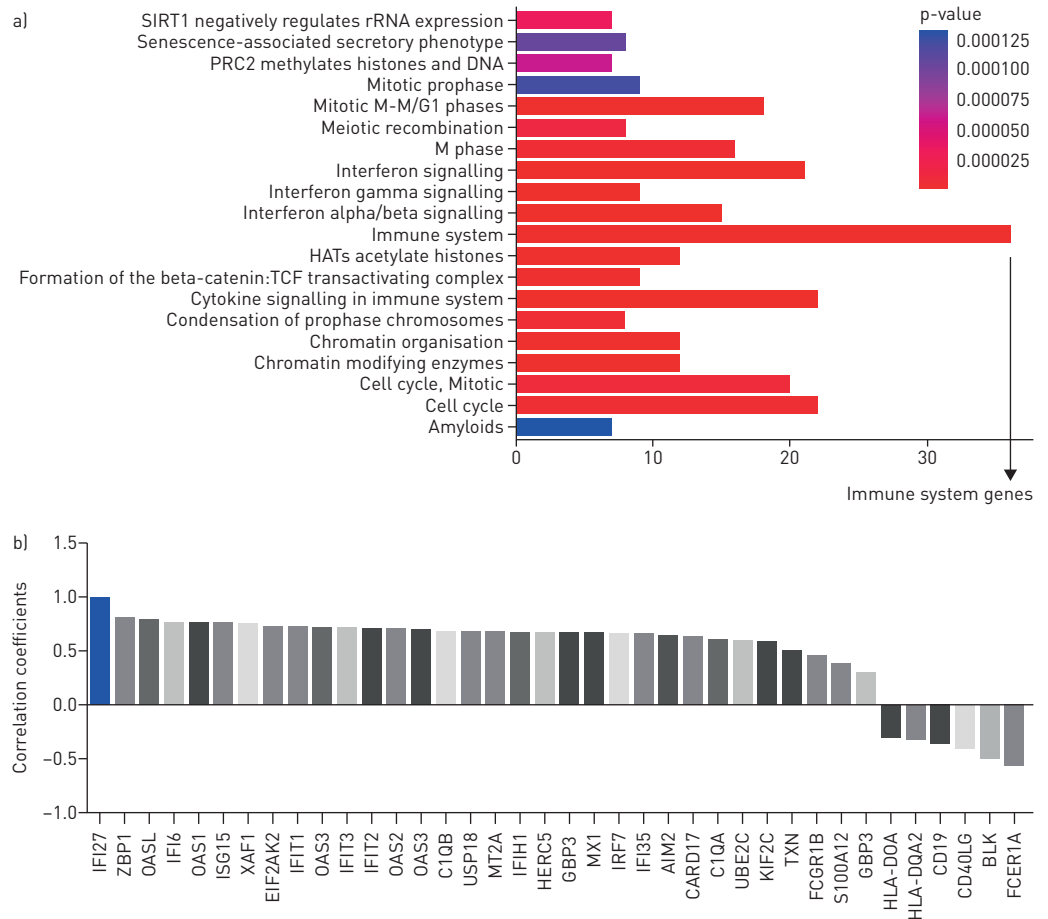


FIGURE 3 a) The immune system pathway was the most over-represented pathway. Pathway enrichment analysis was performed using differentially expressed genes in influenza-positive patients. Pathway terms are indicated on the y-axis and the number of differentially expressed genes in the respective pathway category is indicated on the x-axis. The p-value for the probability that the observed distribution of expression occurred by chance is represented by the various colours of the bars. b) *IFI27* correlated with the immune system pathway. Correlation analysis was performed between *IFI27* and individual genes in the immune system pathway. Representative genes are presented here along with their correlation coefficients indicated on the y-axis. Only statistically significant genes ( $p < 0.05$  using Pearson correlation analysis) are shown. *IFI27* is provided here as a reference gene [correlation coefficient=1].

GSE42026 (n=92) [9, 10]. Two other cohorts consisted of adult patients: GSE60244 (n=158) and GSE68310 (n=133) [11, 12]. We found that *IFI27* was among the signature genes for influenza infection in all four cohorts. In three cohorts (GSE42026, GSE60244 and GSE68310), *IFI27* was the most highly upregulated gene in infected patients (figure 4). In one cohort (GSE6269), *IFI27* was the second most highly upregulated gene in infected patients. These findings are consistent with those of the discovery cohorts, confirming that *IFI27* is an important signature gene for influenza infection.

#### *In vitro model*

To confirm further whether *IFI27* gene expression was directly activated by the influenza virus, we co-cultured peripheral blood samples obtained from healthy volunteers with the live influenza virus. These experiments revealed a dose-response relationship of *IFI27* gene expression in response to viral load (figure 5a), confirming that *IFI27* was activated by the influenza virus. Additional experiments using different influenza A virus strains (H3N2 and H1N1) and the influenza B virus strain showed similar findings (data not shown).

#### *IFI27 is produced by plasmacytoid dendritic cells that recognise the influenza virus*

To identify the immune cells that expressed *IFI27* in peripheral blood, we studied *IFI27* gene expression in eight immune cell subsets, including neutrophils, B cells, CD4, CD8, monocytes, natural killer (NK) cells, myeloid dendritic cells and plasmacytoid dendritic cells (pDCs). We found that in response to

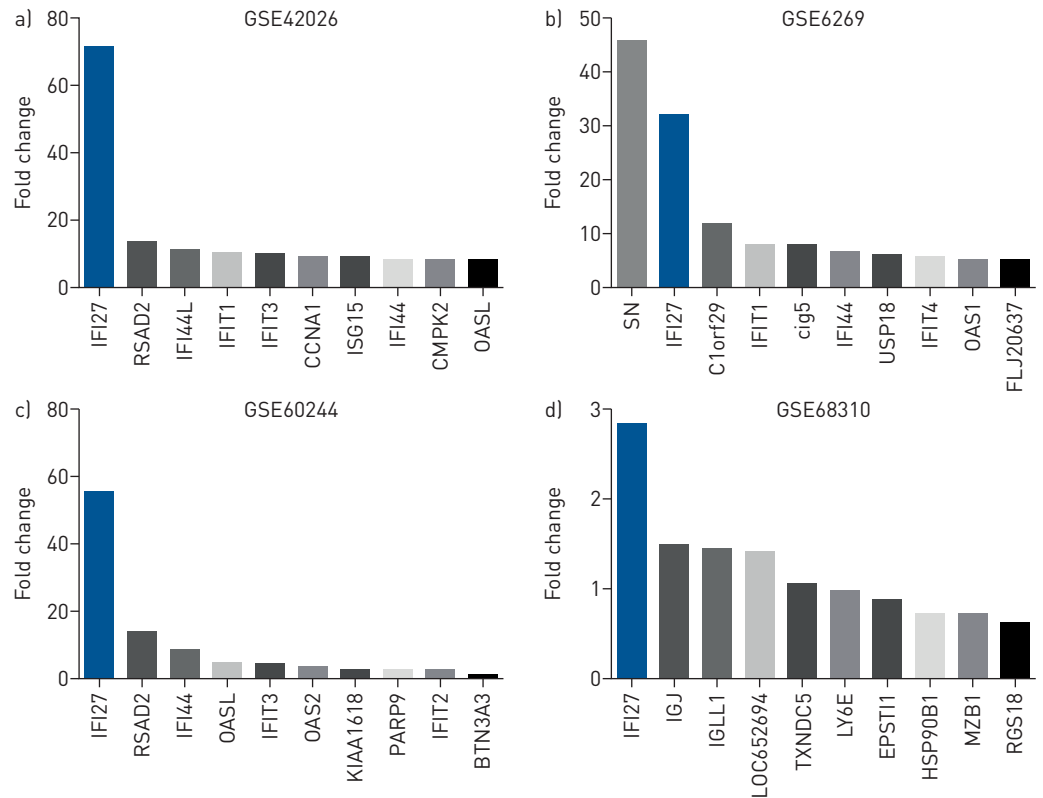


FIGURE 4 Top 10 most upregulated genes in independent external cohorts. Signature genes upregulated by influenza infection are shown in microarray data sets a) GSE42026 [10], b) GSE6269 [9], c) GSE60244 [11] and d) GSE68310 [12]. Gene expression levels are presented here as per original values reported by study authors. Each study used a different microarray platform; therefore, signal intensities were not directly comparable between studies. In data set GSE68310, day 5 data (after infection onset) are presented here for consistency with discovery cohorts (onset=5–10 days).

stimulation by the influenza antigen, pDCs displayed the highest level of *IFI27* gene expression relative to other cell subsets (figure 5b). Notably, neutrophils did not upregulate *IFI27*, despite their known effects on the interferon pathway (figure 5b and d). We also found that the *IFI27* response in peripheral blood was predominantly driven by TLR7 activation (figure 5c), and this TLR7-driven *IFI27* response was mainly observed in pDCs and NK cells (figure 5d). We noted that the *IFI27* response was particularly strong in pDCs (its expression levels were significantly higher in pDCs than in NK cells) (figure 5e). Based on these findings, we concluded that pDCs (and to a much lesser extent, NK cells) were most likely the predominant sources of *IFI27* in peripheral blood. Because the pDC-TLR7 pathway specifically recognises respiratory viruses (e.g. the influenza virus) but does not recognise bacteria, these findings suggest that *IFI27* could be a useful biomarker to discriminate between influenza and bacterial infection.

#### *IFI27* discriminates between bacterial and influenza infection

Consistent with the aforementioned findings, we found that *IFI27* expression in peripheral blood could discriminate influenza pneumonitis from bacterial pneumonia in the discovery cohorts (figure 6a). This was further confirmed by our *in vitro* experiments (figure 6b). In these experiments, influenza viruses (H1N1 and H3N2) and bacterial antigen (lipopolysaccharide) were co-cultured with peripheral blood or purified pDCs. In response to lipopolysaccharide stimulation, *IFI27* upregulation was negligible (in both peripheral blood and pDCs). In contrast, *IFI27* upregulation in response to influenza viruses (both H1N1 and H3N2) was very high (up to 777 fold change in peripheral blood and 1475 fold change in pDCs) (figure 6b).

#### Prospective validation cohort

In phase 2 validation, we prospectively recruited 402 patients with suspected respiratory tract infections and 37 healthy volunteers across five teaching hospitals in Australia from 2012 to 2016 (table 1, supplementary table S2–S3). In this prospective cohort, the diagnostic threshold of *IFI27* was tested on a real-time PCR

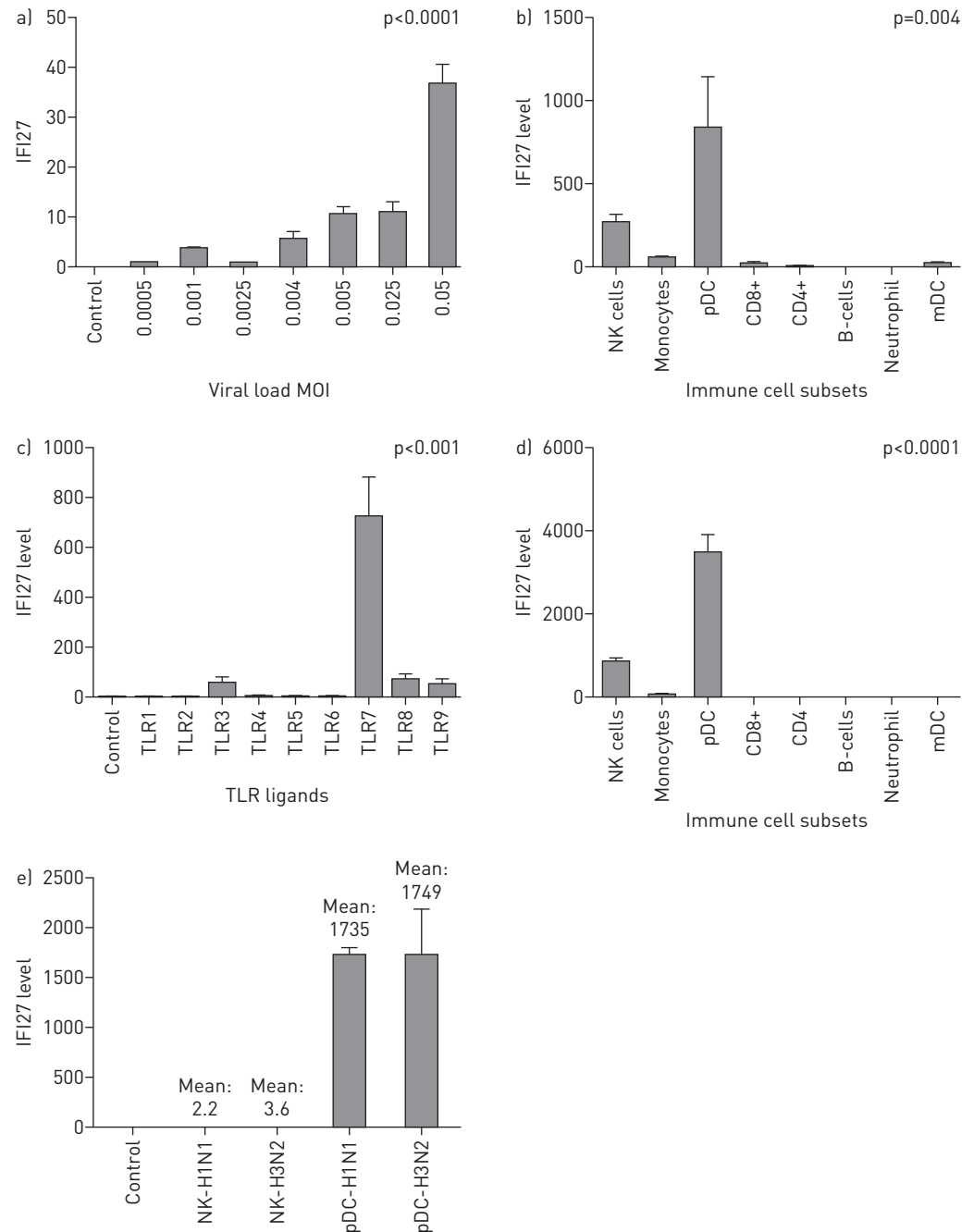


FIGURE 5 *In vitro* studies. Experiments were performed in peripheral blood samples and purified immune cells were obtained from healthy volunteers. All experiments were performed using independent samples from at least three different subjects. *IFI27* gene expression was measured by quantitative real-time PCR in all experiments. *IFI27* levels (on y-axis) were expressed as fold change (relative to glyceraldehyde 3-phosphate [GAPDH]). Column height depicts mean fold change $\pm$ SD. Unless specified otherwise, p-values were calculated using the Kruskal-Wallis test for comparison of multiple groups. a) *IFI27* response to influenza virus. The x-axis represents the range of viral load (MOI)=the ratio of virus to total number of cells. The p-value was calculated using a linear regression model with *IFI27* expression as the dependent variable and viral load as the predictor variable. b) *IFI27* gene expression in immune cell subsets. Influenza viral antigen haemagglutinin (0.135  $\mu\text{g}\cdot\text{mL}^{-1}$ ) was co-cultured with different subsets of immune cells. c) Toll-like receptor (TLR) activation of *IFI27*. Peripheral blood samples were co-cultured with TLR1-9 ligands. d) TLR7 activation in immune cell subsets. Gardiquimod (TLR7 ligand) (1  $\mu\text{g}\cdot\text{mL}^{-1}$ ) was used to stimulate different subsets of immune cells. e) *IFI27* in natural killer (NK) cells and plasmacytoid dendritic cells (pDC). Influenza viruses (H1N1 and H3N2) were co-cultured with NK cells and pDC. *IFI27* gene expression was negligible in NK cells but it was highly upregulated in pDCs. mDC: myeloid dendritic cells.



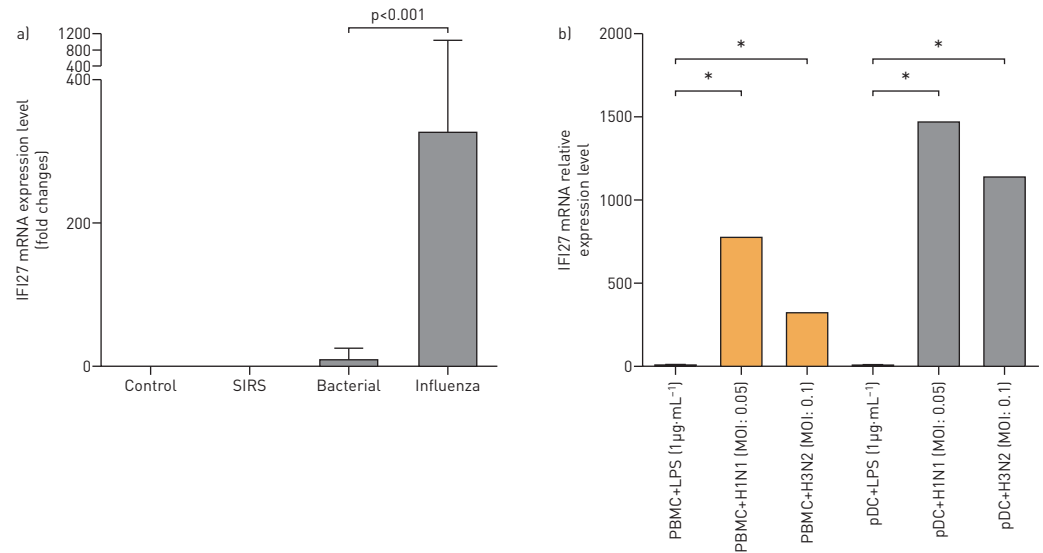


FIGURE 6 a) *IFI27* levels distinguish bacterial from influenza infection in patients. *IFI27* gene expression in peripheral blood samples of patients with influenza virus pneumonitis (n=8), bacterial pneumonia (n=16), systemic inflammatory response syndrome (SIRS) (n=12) and healthy controls (n=18). Patients with SIRS had non-infectious phenotypes including surgery, trauma, pancreatitis, cardiac and pulmonary conditions. *IFI27* levels were measured by quantitative real-time PCR. b) *In vitro* *IFI27* response to bacterial antigen and influenza viruses. Peripheral blood mononuclear cells (PBMC) and plasmacytoid dendritic cells (pDC) were co-cultured with lipopolysaccharide (LPS) or live influenza viruses (H3N2 and H1N1). Multiplicity of infection (MOI) is the ratio of virus to total number of cells. \*: statistically significant difference (p<0.05) between the two groups using the non-parametric Mann–Whitney U-test.

platform. We divided the cohort into a training set (n=200) and a test set (n=202). The diagnostic threshold was first optimised in the training set and was subsequently confirmed in the test set. The calculated AUC of the receiver operating characteristic curve, specificity and sensitivity were nearly identical in both train and test sets. Henceforth, we present the summary findings with both datasets combined.

We found that a threshold of 74 fold change (based on PCR) provided the most optimal level of *IFI27* to differentiate between influenza and bacterial infections (for full details on how this threshold was determined, see supplementary material). At this cut-off value, *IFI27* had a sensitivity of 0.80 and specificity of 0.90. Furthermore, *IFI27* had an AUC of 91% (figure 7), which was within the same range of the multi-gene biomarkers (87%–95%) (table 3). In addition, *IFI27* had an AUC of 83% (sensitivity 0.80 and specificity 0.75) in differentiating influenza virus from other respiratory viruses. This performance is modest compared to its ability to distinguish influenza from bacterial infections. However, we noted that none of the previously published multi-gene biomarkers could distinguish influenza from other respiratory viruses (table 3).

To evaluate further the clinical utility of *IFI27*, we sought to determine whether it could identify influenza infections among all patients in the cohort (figure 7). This additional evaluation represents a more vigorous evaluation of its performance, because *IFI27* would need to identify influenza infection against multiple alternative aetiologies (bacterial infections, other viral infections and all non-infectious conditions), rather than just two aetiologies (influenza *versus* bacteria). Only one previous multi-gene biomarker study evaluated biomarkers in this manner (GSE63990, see table 3) [13]. In our evaluation, *IFI27* showed a high AUC (88%) in identifying influenza infection against multiple aetiologies (sensitivity 0.80 and specificity 0.80) (figure 7). This performance was similar to a previous multi-gene biomarker study (87%) (GSE63990) [13]. Collectively, these findings show that *IFI27* has high diagnostic accuracy, equivalent to those obtained by previously published multi-gene biomarkers.

## Discussion

This study addressed a major challenge in translating genomic science into clinical practice. Over the last decade, considerable advances have been made in discovering the gene expression signatures of respiratory viral infection. These gene expression signatures identify the immune response associated with the infection, thereby significantly improving the clinician's ability to make an accurate diagnosis. However, all published gene expression signatures consist of multi-gene lists, ranging from dozens of genes in some studies to hundreds of genes in others. Measuring expression levels of a large number of genes in clinical



TABLE 2 Detection methods and viruses detected amongst the cohorts

	Virus detection methods				Assay for other viruses <sup>#</sup>	Viruses detected			
	Viral serology	Nucleic acid PCR	Immuno-fluorescence	Viral culture		Influenza A	Influenza B	Other viruses <sup>#</sup>	Co-infection <sup>¶</sup>
<b>Discovery cohorts</b>									
Australia	No	Yes	Yes	No	Yes	Yes (H1N1) <sup>*</sup>	ND	ND	ND
Canada	No	Yes	Yes	Yes	Yes	Yes (H1N1) <sup>*</sup>	ND	ND	ND
Germany	No	Yes	No	No	Yes	Yes (H1N1, H3N2)	Yes	ND	ND
<b>External cohorts</b>									
GSE6269 [9]	No	No	Yes	Yes	Unknown	Yes (subtype unknown)	Yes	ND	ND
GSE42026 [10]	No	Yes	No	No	Yes	Yes (H1N1)	ND	RSV	Yes
GSE60244 [11]	Yes	Yes	Yes	Yes	Yes	Yes (subtype unknown)	Yes	RSV, HMPV	Yes
GSE68310 [12]	Yes	Yes	No	No	Yes	Yes (H1N1, H3N2)	Yes	RSV, RV, coronavirus	Yes
<b>Prospective validation cohort</b>									
	No	Yes	No	No	Yes	Yes (H1N1, H3N2)	Yes	RSV, PI, RV, HMPV, adenovirus	Yes

ND: not detected; GSE: reference number for publicly available microarray datasets in the National Centre for Biotechnology Information Gene Expression Omnibus [reference for published paper is provided beside each GSE number]; RSV: respiratory syncytial virus; HMPV: human metapneumovirus; RV: rhinovirus; PI: parainfluenza virus. <sup>#</sup>: non-influenza respiratory viruses including RSV, HMPV, RV, PI, adenovirus and bocavirus; <sup>¶</sup>: includes virus–virus and virus–bacteria co-infections; <sup>\*</sup>: Australian and Canadian cohorts consisted of exclusively pandemic H1N1/09 influenza virus strains, other cohorts contained seasonal H1N1 and H3N2 influenza virus strains.

practice is problematic. Current real-time PCR technology measures only a small number of genes. Furthermore, determining a diagnostic threshold for many genes is computationally challenging. Most studies have used very complex machine learning algorithms to determine such a threshold. This is impractical in clinical practice, because most institutions do not have access to specialist bioinformatics expertise. In this paper, we demonstrated that a carefully chosen single-gene biomarker might offer an alternative approach. Using evidence from *in vitro* experiments, a mouse model and multiple patient cohorts, we demonstrated that as a single-gene biomarker, *IFI27* could consistently identify influenza infection in different patient populations, in individuals with varying extents of disease severity and in patients with multiple comorbidities. Furthermore, our findings suggest that this biomarker could achieve high prediction accuracy in distinguishing between influenza and bacterial infections, equivalent to those obtained by multi-gene biomarkers.

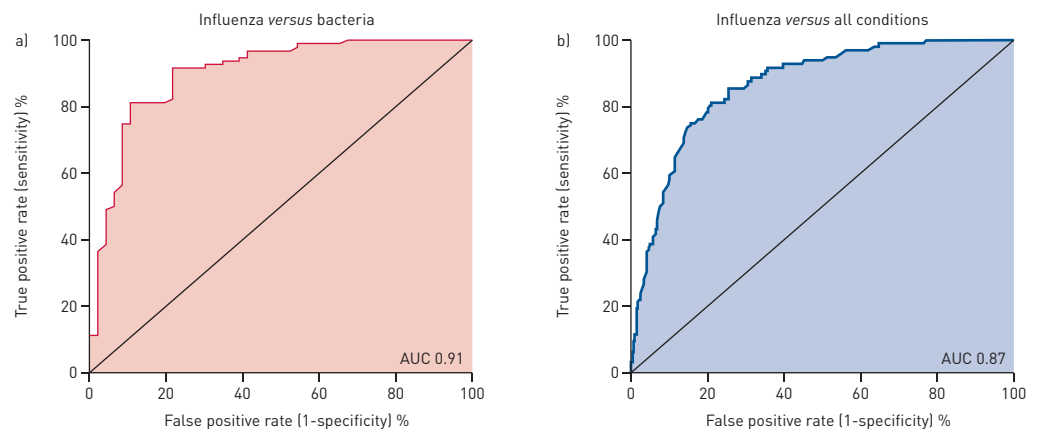


FIGURE 7 Diagnostic performance of *IFI27* on a real-time PCR platform. a) Influenza infection (n=96) versus bacterial infection (n=48). Area under curve [AUC] is 0.91 [95% CI 0.83–0.95,  $p < 0.001$ ]. b) Influenza infection (n=96) versus all conditions (bacterial infection, viral infection and non-infectious “flu-like” illness) (n=293). AUC is 0.87 [95% CI 0.82–0.90,  $p < 0.001$ ]. AUC is calculated using the full range of *IFI27* cut-off values (fold change 10–6000).

TABLE 3 Diagnostic performance

	Aims	Influenza specific	Overall diagnostic accuracy % <sup>#</sup>	Genes included in signature n	Subjects in validation cohort n	Inclusion of non-infectious phenotypes in validation cohort
<b>Multi-gene signatures</b>						
GSE60244 [11]	Discriminate viral <i>versus</i> bacterial infections	No	96	10	46	Yes
GSE63990 [13]	Discriminate viral infection from all conditions <sup>¶</sup>	No	87	33	273	Yes
GSE6269 [9]	Discriminate influenza <i>versus</i> bacterial infections	Yes	87–95	35	75	No
GSE42026 [10]	Discriminate influenza <i>versus</i> bacterial infections	Yes	92	734	43	No
<b>Single-gene signature</b>						
<i>IFI27</i>	Discriminate influenza infection from all conditions <sup>¶</sup> or from bacterial infection	Yes	87–91	1	402	Yes

GSE: reference number for publicly available gene expression dataset in the National Centre for Biotechnology Information Gene Expression Omnibus (reference for published paper is provided beside each GSE number). <sup>#</sup>: the most common parameter reported by all studies. Other parameters, such as specificity and sensitivity, were not reported in most studies. In one study (GSE6269), the gene expression signature was validated in several cohorts, each yielding its own prediction accuracy; therefore, a range of prediction accuracy is presented here. One study (GSE68310) did not assess diagnostic performance; therefore, it is not included in the table. <sup>¶</sup>: includes bacterial and viral infections and non-infectious “flu-like” illness.

Our study is by far the largest biomarker study undertaken in influenza virus infection. It consisted of eight cohorts (n=1071), with patients drawn from heterogeneous populations in different geographical locations, settings (outpatients, emergency departments and intensive care units) and different levels of infection severity (mild, moderate to severe). Such a broad clinical case-mix makes it likely that our findings are generalisable to other patient populations. Our findings are also consistent with a recently published meta-analysis of respiratory viral infection. That meta-analysis included five influenza cohorts (n=292) (two cohorts originated from the present study) [14]. It identified 127 multi-gene signatures associated with the influenza infection. Of these 127 genes, *IFI27* was the most highly upregulated gene.

Our *in vitro* data are important, because it provides a biologically plausible explanation for the ability of *IFI27* to differentiate influenza from bacterial infections. Our findings revealed that *IFI27* was produced by plasmacytoid dendritic cells (pDC), a pivotal immune cell subset that is known to recognise influenza virus, but not bacteria [15, 16]. Our data also revealed that the TLR7 pathway mediates *IFI27* upregulation. This is consistent with established literature, which shows that pDCs use TLR7 to recognise single-stranded RNA viruses, such as the influenza virus [17, 18]. We also found that bacteria-recognizing TLRs (TLR1, TLR2, TLR4, TLR5 and TLR6) did not induce *IFI27* upregulation, whereas TLR7 ligand did. These findings, together with previous studies [17–20], collectively support the notion that the TLR7-pDC pathway represents a virus-specific detection system (for single-stranded RNA viruses). This mechanistic insight helps explain the ability of the *IFI27* biomarker to differentiate between influenza and bacterial infections.

The prospective validation cohort in the present study had two design features (large sample size and prospective patient recruitment), which made it possible to address an important limitation in previously published biomarker studies. In previous studies, the validation cohort typically used patient phenotypes that were clearly “infectious” or phenotypes that were clearly “non-infectious” (e.g. healthy volunteers). A recent study by TSALIK *et al.* [13] demonstrated that using such clearly defined phenotypes resulted in an over-estimation of biomarker performance. A methodologically more vigorous approach should include less distinctive phenotypes, such as patients with non-infectious “flu-like” illness [13]. Our prospective validation cohort had a significantly greater number of non-infectious patients compared to previous studies (table 1). Another strength of the present study was the recruitment of patients prospectively. Patients presented in an undifferentiated manner, whereby both infectious and non-infectious aetiologies were possible, thus simulating a realistic clinical scenario commonly faced by clinicians. Furthermore, many of our patients had respiratory comorbidities (asthma 17% and chronic lung diseases 22%), which could also mimic respiratory infections (e.g. with similar symptoms of cough and shortness of breath).

Our cohort therefore recapitulated many of the real-world challenges and embodied the most stringent conditions under which the performance of *IFI27* could be tested.

A single-gene biomarker, such as *IFI27* circumvents several methodological problems faced by multi-gene biomarkers. Multi-gene biomarkers are vulnerable to over-fitting (using too many genes to develop the disease signature). This results in poor reproducibility across external patient populations. In the present study, we minimised over-fitting by focusing on a small number of biomarkers, from which we identified a single biomarker that was highly reproducible across multiple patient cohorts. Another common problem with multi-gene biomarkers is that there are many non-overlapping genes among published studies. This creates an immense translational problem, because clinicians cannot determine which gene list to use with their patients. The simplicity of using a single-gene biomarker such as *IFI27* represents a major step forward in overcoming this translational roadblock. A third problem with multi-gene lists is that most studies have been derived from microarray studies. Here, we validated *IFI27* on a real-time PCR platform, which is widely available in most hospitals. This opens the door for other clinicians to perform independent validation, an essential next step towards translation into clinical practice.

It is worth noting that existing multi-gene biomarkers do have an advantage over a single biomarker. Multi-gene biomarkers broadly reflect host response, by capturing a constellation of genomic signals from different parts of the immune signalling pathway. In comparison to multi-gene biomarkers, a single biomarker might underestimate both the range and depth of the host response. However, our findings indicate that *IFI27* levels did provide a reasonable assessment of the immune signalling pathway. This was evidenced by a high level of correlation between *IFI27* and many genes in the interferon pathway (90%) and immune system pathway (69%). Altogether, these findings suggest that *IFI27*, as a single-gene biomarker, is reflective of the immune response to a similar extent as multi-gene biomarkers.

*IFI27* differs from previously established influenza biomarkers (e.g. *OAS* and *PKR*) in three important aspects. First, *IFI27* performed significantly better than other biomarkers, as demonstrated in our microarray analysis. Second, *IFI27* is produced predominantly by plasmacytoid dendritic cells (pDCs), immune cells that recognise respiratory viruses. In contrast, most biomarkers are produced by immune cells (e.g. neutrophils and macrophages) that can be activated by bacteria. Third, *IFI27* has considerable discriminatory power in differentiating between influenza and bacterial infections. Our data show that as a single-gene biomarker, *IFI27* alone provides equivalent diagnostic capability comparable to that of multi-gene biomarkers. Previous biomarkers do not have comparable capabilities.

Nevertheless, the present study has limitations. Firstly, the specificity of *IFI27* in differentiating between influenza virus and other respiratory viruses is moderate (AUC 83%). Second, we had insufficient data on patients with DNA viral infections (e.g. cytomegalovirus) or virus-bacteria co-infection; our cohort had only thirteen such cases, which is too small a number for in-depth analysis. In our previously published study, we found that a longitudinal study design was needed to investigate patients with co-infection (in order to capture the transition from influenza infection to co-infection) [21]. Unfortunately, the current study is limited by the fact that patient samples were collected only at one time point. Third, our study did not assess important factors that might affect clinical decision-making, such as prior probability or the cost of missed diagnoses and its impact on patient/economic outcomes. These factors are best addressed in future cost-benefit studies or clinical trials.

In conclusion, our findings suggest that *IFI27* is a promising immune biomarker; however, further study is needed to evaluate its clinical utility in patients with viral/bacterial co-infections.

### Acknowledgements

We would like to thank Kenneth McPhie and Ian Barr (WHO Collaborating Centre for Reference and Research on Influenza, Melbourne, Australia) for supply of the influenza virus stock for our *in vitro* experiments, Ya Wang (Dept of Intensive Care Medicine, Nepean Hospital, Kingswood, Australia) and Nicole Fewings (Centre for Immunology and Allergy Research, Westmead Institute for Medical Research, Westmead, Australia) for assistance in analysing flow cytometry data and Ian Barr for helpful review on an earlier version of our manuscript. We would also like to thank several organisations for assisting our study, including: the Nepean Institute of Critical Care Education and Research (Kingswood, Australia) for providing funding to support our study in Australia, the Czech Republic and Hong Kong; Public Health Agency of Canada (PHAC) (Ottawa, Canada) for providing funding to support the Canadian arm of our study; Roche Pharmaceuticals (Ottawa, Canada) for funding H1N1 sample collection in Canada; and the Canadian Severe Influenza Study group (see below) for contributing patient samples to the study.

The collaborating members of the Canadian Severe Influenza Study group are as follows: Anand Kumar (Section of Critical Care Medicine and Section of Infectious Diseases, Dept of Medicine, Medical Microbiology and Pharmacology, University of Manitoba, Winnipeg); Gordon Wood (Royal Jubilee Hospital and Victoria General Hospital, Victoria); Steve Reynolds (Vancouver General Hospital, University of British Columbia, Vancouver); Brent Winston (Foothills Medical Centre and Peter Lougheed Hospital, University of Calgary, Calgary); Sean Bagshaw (University of Alberta Hospital, University of Alberta, Edmonton); Jim Kutsogiannis (Royal Alexandra Hospital, University of Alberta, Edmonton); Michael Silverman (Rouge Valley Hospital, Ajax, Canada and Lakeridge Hospital, Oshawa); Margaret Herridge (Toronto General Hospital and Toronto Western Hospital, University of Toronto, Toronto); Rob Fowler

(Sunnybrook Hospital, University of Toronto, Toronto); Alison McGeer (Mount Sinai Hospital, University of Toronto, Toronto); Mary-Anne Aarts (St Michael's Hospital, University of Toronto, Toronto); Deborah Cook (St Joseph's Healthcare, McMaster University, Hamilton); Lauralyn McIntyre (Ottawa Hospital General Campus and Civic Campus, University of Ottawa, Ottawa); Kosar Khwaja (Montreal General Hospital and Royal Victoria Hospital, McGill University, Montreal); Natalie Bandrauk (Health Sciences Centre and St Claire's Mercy Hospital, Memorial University, St. John's); Richard Hall (Queen Elizabeth Health Sciences Centre – Halifax Infirmary and Victoria General Site, Dalhousie University, Halifax).

The authors contributed to the study in the following ways. Benjamin M. Tang conceived the study. Maryam Shojaei conducted the *in vitro* study. Klaus Schughart conducted the mouse study. Klaus Schughart and Grant P. Parnell conducted the microarray analysis. Maryam Shojaei, Grant P. Parnell, Anand Kumar, Aseem Kumar, John Ho, Yoav Keynan, Adrienne F.A. Meyers, Terry Ball, Amarnath Pispipati, Amy L. Phu, Stephen Huang, Marek Nalos, Kate O'Connor, Stephen Schibeci, Kevin Lai, Hao Luo, Fahad Gul, Mark Gillett, David Booth, Elizabeth Moore, Robert Geffers, Jens Schreiber, Damon Eisen and Benjamin M. Tang participated in the study and critical review of the manuscript; B. Tang drafted the manuscript.

## References

- Jansen RR, Wieringa J, Koekkoek SM, *et al.* Frequent detection of respiratory viruses without symptoms: toward defining clinically relevant cutoff values. *J Clin Microbiol* 2011; 49: 2631–2636.
- Harper SA, Bradley JS, Englund JA, *et al.* Seasonal influenza in adults and children: diagnosis, treatment, chemoprophylaxis, and institutional outbreak management: clinical practice guidelines of the infectious diseases society of America. *Clin Infect Dis* 2009; 48: 1003–1032.
- Zarychanski R, Stuart TL, Kumar A, *et al.* Correlates of severe disease in patients with 2009 pandemic influenza (H1N1) virus infection. *CMAJ* 2010; 182: 257–264.
- Kumar A. Early *versus* late oseltamivir treatment in severely ill patients with 2009 pandemic influenza A (H1N1): speed is life. *J Antimicrob Chemother* 2011; 66: 959–963.
- Zaas AK, Garner BH, Tsalik EL, *et al.* The current epidemiology and clinical decisions surrounding acute respiratory infections. *Trends Mol Med* 2014; 20: 579–588.
- Jenner RG, Young RA. Insights into host responses against pathogens from transcriptional profiling. *Nat Rev Microbiol* 2005; 3: 281–294.
- Herberg JA, Kaforou M, Wright VJ, *et al.* Diagnostic test accuracy of a 2-transcript host RNA signature for discriminating bacterial vs viral infection in febrile children. *JAMA* 2016; 316: 835–845.
- Mahajan P, Kuppermann N, Mejias A, *et al.* Association of RNA biosignatures with bacterial infections in febrile infants aged 60 days or younger. *JAMA* 2016; 316: 846–857.
- Ramilo O, Allman W, Chung W, *et al.* Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood* 2007; 109: 2066–2077.
- 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: 1664–1668.
- Suarez NM, Bunsow E, Falsey AR, *et al.* Superiority of transcriptional profiling over procalcitonin for distinguishing bacterial from viral lower respiratory tract infections in hospitalized adults. *J Infect Dis* 2015; 212: 213–222.
- 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: e1004869.
- Tsalik EL, Henao R, Nichols M, *et al.* Host gene expression classifiers diagnose acute respiratory illness etiology. *Sci Transl Med* 2016; 8: 322ra11.
- Andres-Terre M, McGuire HM, Pouliot Y, *et al.* Integrated, multi-cohort analysis identifies conserved transcriptional signatures across multiple respiratory viruses. *Immunity* 2015; 43: 1199–1211.
- Barchet W, Cella M, Colonna M. Plasmacytoid dendritic cells – virus experts of innate immunity. *Semin Immunol* 2005; 17: 253–261.
- Reizis B, Bunin A, Ghosh HS, *et al.* Plasmacytoid dendritic cells: recent progress and open questions. *Annu Rev Immunol* 2011; 29: 163–183.
- Guiducci C, Coffman RL, Barrat FJ. Signalling pathways leading to IFN- $\alpha$  production in human plasmacytoid dendritic cell and the possible use of agonists or antagonists of TLR7 and TLR9 in clinical indications. *J Intern Med* 2009; 265: 43–57.
- Diebold SS, Kaisho T, Hemmi H, *et al.* Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* 2004; 303: 1529–1531.
- Lund JM, Alexopoulou L, Sato A, *et al.* Recognition of single-stranded RNA viruses by Toll-like receptor 7. *Proc Natl Acad Sci USA* 2004; 101: 5598–5603.
- Huang Y, Zaas AK, Rao A, *et al.* Temporal dynamics of host molecular responses differentiate symptomatic and asymptomatic influenza A infection. *PLoS Genet* 2011; 7: e1002234.
- 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: R157.