Title Page

Transcriptomic profiling in childhood H1N1/09 influenza reveals reduced expression of protein synthesis genes

Authors: Jethro A Herberg[1], Myrsini Kaforou[1], Stuart Gormley[1], Edward R Sumner[2], Sanjay Patel[1], Kelsey DJ Jones[1], Stéphane Paulus[1], Colin Fink[2], Federico Martinon-Torres[3], Giovanni Montana[4], Victoria J. Wright[1], Michael Levin[1]

[1]Section of Paediatrics and [4]Department of Mathematics, Imperial College, London W2 1PG, UK; [2]Micropathology Ltd, Warwick Science Park, Coventry CV4 7EZ, UK; [3]Department of Pediatrics, Hospital Clínico Universitario de Santiago, 15706 Santiago de Compostela, Spain.

Corresponding author:

Dr Jethro A Herberg

Section of Paediatrics

Imperial College London

Norfolk Place

London W2 1PG

UK

Tel: +44 (0)20 7594 3717

Fax: +44 (0)20 7594 3984

Email: j.herberg@imperial.ac.uk

Running head: Gene Expression Profile of Influenza H1N1

Abstract word count: 100

Main text word count: 2000

Footnote page

Conflict of interest: the authors do not have a commercial or other association that might pose a conflict of interest

Financial Support: This work was supported by the Imperial College Comprehensive Biomedical Research Centre and the Wellcome Trust Centre for Respiratory Infection at Imperial College (DMPED P26077 to JAH) and the Fondo de Investigación Sanitaria del plan nacional de I+D+I and 'fondos FEDER (FIS; PI070069 / PI1000540 to FMT). FMT represents the GENDRES research group (www.gendres.org).

Data presented at meetings: part of this work was presented in an oral presentation at the annual meeting of the European Society for Paediatric Infectious Diseases in Nice, France (June 2011)

Address for correspondence and reprint requests:

Dr Jethro A Herberg

Section of Paediatrics

Imperial College London

Norfolk Place

London W2 1PG

UK

Email: j.herberg@imperial.ac.uk

Current affiliations:

Sanjay Patel: Southampton University Hospitals NHS Trust, Southampton UK

Stéphane Paulus: Alder Hey Children's NHS Foundation Trust, Liverpool UK

Abstract

We compared the blood RNA transcriptome of children hospitalised with influenza A H1N1/09, RSV or bacterial infection, and healthy controls. Compared to controls, H1N1/09 patients showed increased expression of inflammatory pathway genes, and reduced expression of adaptive immune pathway genes. This was validated on an independent cohort. The most significant function distinguishing H1N1/09 patients from controls was protein synthesis, with reduced gene expression. Reduced expression of protein synthesis genes also characterised the H1N1/09 expression profile compared to children with RSV and bacterial infection, suggesting that this is a key component of the pathophysiological response in children hospitalised with H1N1/09 infection.

Key Words (MeSH): influenza; respiratory tract infection; gene expression profiling; RSV; Peptide Chain Initiation, Translational; eIF-2 Kinase; microarray analysis; pediatric

Introduction

Whilst children with H1N1/09 influenza A (H1N1/09) commonly had a mild illness, a proportion requiring intensive care developed cardiovascular shock, multi-organ failure and fatal disease [1]. UK paediatric mortality rates for H1N1/09 were 5–11 per million [2], the highest since 2004, when meningococcal serogroup C vaccine was introduced.

Whole blood RNA expression profiling is ideally suited to study emerging infections as it allows interrogation of the host response [3]. Influenza studies have included the immunopathogenesis of in vivo experimental challenge [4], in addition to in vitro infection and vaccination models. Published whole blood transcription data from H1N1/09-infected patients has focussed on adults, whose immune response is influenced by recall from previous influenza infection [5, 6]. We postulated that comparison of RNA expression in patients with H1N1/09 with other common childhood infections would reveal key immunopathogenic responses to the pandemic agent. We studied RNA expression profiles in whole blood of prospectively recruited children hospitalised with respiratory infections during 2009/10, as the H1N1/09 pandemic evolved in the UK.

Methods

Patient Cohorts

Between July 2009 and June 2010, we recruited 165 acutely ill febrile children (below 17 years) with respiratory infection of sufficient severity to warrant blood tests, at St Mary's Hospital London. Whole blood for RNA was collected in PAXgene® tubes, together with clinical samples, including a full blood count which was analysed in the hospital laboratory. Cases were recruited early in their hospital assessment, before diagnostic studies were completed, and were assigned to diagnostic categories once results were available (Supplementary Figure 1). The study had approval of the St Mary's Research Ethics Committee (REC 09/H0712/58). Written, informed consent was obtained.

Pathogen diagnosis

Viral diagnostic studies were undertaken on nasopharyngeal aspirates using immunofluorescence (RSV, adenovirus, parainfluenza virus, influenza A+B) and nested PCR (RSV, coronavirus, adenovirus, parainfluenza 1-4, influenza A+B, bocavirus, metapneumovirus, rhinovirus). Bacterial diagnostics included culture of blood and pleural fluid, and pneumococcal antigen detection in blood or urine where available.

RNA expression profiling**:**

PAXgene® tubes were extracted using PAXgene Blood RNA extraction kits (Qiagen) according to the manufacturer's instructions. After quantification and quality control, biotin-labelled cRNA was prepared from 330ng mRNA using Illumina TotalPrep RNA Amplification kits (Applied Biosystems). 750ng labelled cRNA was hybridised to Illumina HumanHT-12 v3 Expression BeadChips, and the microarrays scanned. Quality control parameters were assessed using Genome Studio software, visual inspection of the microarray images and Principal Component Analysis (PCA). PCA was used to identify significant influences on the data other than disease cohort, such as technical factors, comorbidities or concurrent medications. One RSV patient with poor quality array data was removed. Two children in the H1N1, RSV and bacterial cohorts

had oral or intravenous steroids prior to blood sampling, but none were found to be outliers from their cohorts using PCA or unsupervised clustering algorithms (data not shown).

Microarray Analysis

Expression data were analysed using '*R' Language and Environment for Statistical Computing 2.12.1* and GeneSpringGX 11.5 software (Agilent). Additional detail is provided online (Supplementary Methods). The data have been deposited in NCBI's Gene Expression Omnibus [7] and are accessible through Series accession number GSE42026 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE42026).

Pathway analysis

Data were analysed using IPA (Ingenuity® Systems, www.ingenuity.com). We filtered the datasets, ranked by *P*-value, to a limit of 800 network-associated molecules, to identify the strongest biological signals. We identified the pathways from the IPA library of canonical pathways that were most significantly represented in our dataset.

Statistical analysis of cohorts

Differences between cohorts in age, proportion of white cell type and days from illness onset to presentation were calculated using Kruskall-Wallis tests, with Dunn's post-test comparisons of each cohort (Graphpad Prism). Differences in gender, severity of illness and deaths were calculated using chi-squared tests of independence (Supplementary Table 1).

Results

We recruited febrile children hospitalised for respiratory illness, including 25 patients with H1N1/09 infection (6 co-infected with other pathogens), 34 with respiratory syncytial virus (RSV) infection (11 co-infected), and 18 with gram-positive bacterial infection (Supplementary Figure 1). Thirty-three healthy control children were recruited from the outpatients department. Patient demographic and clinical data are summarised in Supplementary Table 1.

RNA expression

RNA expression profiling was undertaken on venous blood. Comparing children with H1N1/09 infection and controls, we found 1,267 significantly differentially expressed (SDE) transcripts (P<0.001) (Supplementary Data S1, S2a). Unsupervised clustering analysis separated the subjects into distinct highly concordant groups (Figure 1a). When the transcript set was tested using Support Vector Machine (SVM) with leave-one-out validation, strong class prediction sensitivity was achieved (96%) (Supplementary Data S1).

When the two other infection cohorts (RSV and bacterial infection) were compared to controls, we identified 1,172 and 1,869 SDE transcripts ($P<0.001$) respectively (Supplementary Data S1, S2b,c). The validity of these transcript sets in distinguishing each infection from controls was analysed by SVM giving a sensitivity of 95% and 98% for RSV- and bacterial-infected patients respectively) (Supplementary Data S1). As the RSV cohort was, as expected, significantly younger than the control or H1N1 cohorts, we investigated the effect of age on gene expression by comparing the RSV patients to a subgroup of the youngest controls (n=10, median age 9 months). There were 2,411 SDE genes (corrected $P<0.05$). These showed strong concordance with the SDE genes in the RSV vs. all controls comparison: 772 of the top 800 SDE transcripts between RSV and youngest controls had the same direction of expression in both datasets; 28 transcripts were not represented in the comparator dataset.

Identification of biological processes in H1N1/09 and comparator infections

The biological significance of the transcript sets distinguishing H1N1/09, RSV, and bacterial infections from controls was investigated with IPA (Ingenuity® Systems, www.ingenuity.com), which assigns SDE transcripts to known biological functions and pathways. Biological functions matched by the SDE genes for all 3 comparisons were enriched for infectious disease, respiratory disease and inflammatory response genes (Supplementary Data S3).

In the H1N1/09 vs. controls dataset, the functional category "protein synthesis" was significantly overrepresented (P=7.93x10^{-26}). Furthermore, the 3 most significant canonical pathways in this dataset were related to the initiation of protein translation: eukaryotic initiation factor 2 (eIF2 - P=3.16x10^{-25}), eukaryotic initiation factor 4 (eIF4 - P=3.24x10^{-88}) and mammalian target of rapamycin (mTOR - P=1.91x10^{-6}) (Supplementary Data S4).

When transcripts with increased or decreased abundance were examined separately, those with increased abundance were enriched for immune response pathways and functions, including signalling from pattern recognition receptors (top increased genes *OAS3, C1QB, C1QC*), interferon signalling (*IFI35, IFIT1, IFT3*), and antigen presentation through MHC class I (*TAP1, MR1, HLA-B*). Transcripts with reduced abundance were enriched for pathways including antigen presentation through MHC class II (most decreased: *HLA-DQA, HLA-DPB, HLA-DOA*), T cell signalling and protein synthesis - including those involved in translation, elongation and ribosomal transport (Supplementary Data S4). There was reduced expression in 63 of 69 SDE transcripts in the eIF2 pathway, P<0.0001. Protein synthesis-related transcripts with increased expression included genes with known inhibitory functions in protein synthesis, including *EIF2AK2, IFIT1/ISG56* and *IFIT2/ISG54* (fold change (FC) of 15.3, 1013 and 16.2 respectively in H1N1/09 relative to controls) [8, 9].

Comparison to an independent validation dataset

In order to validate our findings, we used a publically available adult dataset for H1N1/09 and controls [6]. There was complete segregation of H1N1/09 patients and controls when SDE transcripts from our data were used for clustering (Supplementary Figure 2a). We used the SVM prediction model trained on our H1N1/09 vs. control data and applied this to the independent dataset, there was 100% sensitivity and specificity of disease and control status. We investigated whether the 69 represented transcripts in the top canonical pathway, EIF2 signalling, were significantly differentially expressed in the validation dataset. 32 were SDE in the adult study, including 31 with the same direction of regulation (Supplementary Figure 2b).

Comparison of the H1N1/09 host response with RSV and bacterial infection

Having identified the expression profile distinguishing patients with each pathogen infection from healthy children, we compared the response to H1N1/09 with RSV and bacterial infection in order to identify unique aspects of the transcription profile of H1N1/09 infection. Transcripts that were SDE between H1N1/09 and RSV (n=601 of 7,295 transcripts, corrected P<*0.01*) and between H1N1/09 and bacterial infection (n=734 of 6,808 transcripts, corrected P<*0.01*) were used for unsupervised clustering and for classification, H1N1/09 cases were distinguished from RSV cases (Figure 1b) and bacterial cases (Figure 1c) with accuracy of 93% for H1N1/09 vs. RSV and 92% for H1N1/09 vs. bacteria (Supplementary Data S1, S2d,e).

'Protein synthesis' was the most significant functional category in both the H1N1/09 vs. RSV, and H1N1/09 vs. bacterial infection datasets (corrected $P=2.73\text{x}10^{-13}$ and $1.6\text{x}10^{-11}$ respectively). The most significant canonical pathways related to protein synthesis - the eIF2 pathway ($P=2.0 \text{ x } 10^{-20}$ and $9.0 \text{ x } 10^{-10}$); eIF4 ($P=4.7 \text{ x } 10^{-8\text{-x}}$ and 0.006) and mTOR ($P=8.1 \text{ x } 10^{-7}$ and 0.016) for H1N1/09 vs. RSV, and H1N1/09 vs. bacterial infection respectively (Figure 2 and Supplementary Data S3, S4).

Discussion

Whole blood RNA expression profiling demonstrated marked differences in transcriptional profiles in H1N1/09-infected patients compared to controls or to RSV and gram-positive bacterial infection. In relation to controls, H1N1/09 patients showed increased expression in well-established pathways of the innate antiviral immune response, including pattern recognition receptor, interferon signalling and myeloid cell activation pathways. Many interferon-induced genes with antiviral roles showed increased differential expression, including *IFI27* (P=5.8x10^{-11}, FC 5.8x10^{21}), and genes implicated in repression of protein synthesis including *EIF2AK2*, and also *IFIT1* and *IFIT2*, whose antiviral roles may be mediated through effects on translation [10, 11]. Recent work in humans and mice has implicated the interferon-induced *IFITM3* gene in restricting influenza A infection [12]. Our finding of strongly increased *IFITM3* expression (P=3x10^{-3}, FC 76.2) supports its role in H1N1/09 infection.

Differences in whole blood gene expression between cohorts can reflect altered proportions of cell types and/or altered transcript abundance within cell types. We used concurrent clinical full blood count data to compare lymphocyte proportions in each cohort (Supplementary Table 1). Lymphocyte proportion differed significantly only between H1N1/09 and controls (ratio 2.1:1). By comparison, the fold change between cohorts was higher in all 800 transcripts used for Ingenuity analysis (range 2.3-5x10^{21}, median 4.5). This indicates that differences in lymphocyte proportion are not the major determinants of gene expression level. This does not exclude a role for lymphocyte subset populations.

Whilst H1N1/09 patients showed strongly increased expression of innate immune genes, there was reduced expression in many adaptive immune response pathways, including T cell activation through NFAT, B cell activation and MHC class II antigen presentation. Reduced expression of adaptive T and B cell pathways may represent a host negative regulatory feedback mechanism [5], or viral-induced immune subversion [13], analogous to viral inhibition of antigen presentation by inhibition of the MHC class II transactivator.

We found that the 3 most significant pathways in our dataset were involved in protein synthesis and translation initiation (eIF2, eIF4 and mTOR), and that within these the majority of genes showed reduced expression. In contrast, inhibitors of translation including *EIF2AK2* (which encodes PKR) were upregulated. PKR, when bound to viral double-stranded RNA, phosphorylates eukaryotic initiation factor 2 alpha (eIF2a) to cause the arrest of translation initiation [14]. This suggests that protein synthesis in H1N1/09 patients may be impaired by repressed transcription of genes required for the translation apparatus, and by increased transcription of inhibitors of translation. The reduced expression of genes involved in protein synthesis was highly significant between H1N1/09 and controls (P=7.9x10^{-26}) and also between H1N1/09 and RSV or bacterial infection (P=2.73x10^{-13} and 1.6x10^{-11} respectively) (Figure 2). In an influenza A challenge study of healthy adult volunteers, differential expression of translation initiation pathway genes was identified in symptomatic but not subclinical infections [4], supporting a role for altered protein synthesis as a key component of the host response to H1N1/09 infection.

Increased expression of protein synthesis-inhibiting genes with reduced expression of translation initiation genes is likely to impair protein synthesis. Influenza virus has been observed to shut-off protein synthesis and decrease mRNA levels in infected cells [15], but the phenomenon we have observed in peripheral blood suggests a widespread alteration affecting cells distant from the site of infection in the airway. We therefore hypothesize that suppression of protein synthesis may be a distinctive feature of H1N1/09 infection, and may play a role in the prolonged debility, fatigue and delayed recovery seen in severely affected patients. Many of the fatal cases of respiratory failure in H1N1/09 infection had an unusual illness with prolonged Acute Respiratory Distress syndrome, requiring prolonged ventilation [1], and late pulmonary haemorrhage or persistent respiratory failure causing death weeks after the acute illness. Supressed protein synthesis might result in poor tissue recovery, and may provide an explanation for the prolonged illness and poor outcome in some patients.

In a neonatal pig model of critical care myopathy, impairment of protein synthesis in association with sepsis was mediated through inhibition of mTOR and eIF2 pathways [16]. In vitro infection studies using H5N1 influenza virus indicate that impaired mTOR signalling leads to autophagic cell death [17]. Further studies are required to confirm that the changes we observed in peripheral blood RNA expression for these pathways do represent inhibition of protein synthesis, and to confirm the role of this phenomenon in severe illness.

It is unclear whether the changes in expression of genes involved in protein synthesis we have observed in peripheral blood are triggered by the virus, or by host mediators released during the immune response. Influenza infection is largely confined to epithelial surfaces in the lung, and viraemia is infrequent, though more common in severe infection. Changes in the peripheral blood may reflect a response to entrance of virus or viral components into the blood, or an indirect effect of local cellular activation by host inflammatory molecules. Potentially, there are advantages to both the virus and the host in switching off protein synthesis – for instance viral immune subversion by impairing leukocyte function, or limitation of production of viral proteins.

Limitations

Limitations in this study include the younger age of the RSV cohort. The highly varied ethnicity of children attending our hospital meant it was not possible to match ethnicities between groups. We used a whole-blood gene expression approach: changes in expression during disease result from both shifting cell type proportions and upregulation or downregulation of genes within cells. The association of influenza with bacterial infection is well known, and occult bacterial infection may have contributed to disease in some patients.

In conclusion, our comparative study of RNA expression in children with H1N1/09 infection has provided new evidence on the host response to the pandemic strain. The finding that changes in expression of genes involved in protein synthesis are a key feature distinguishing H1N1/09

from other common childhood infections provides a new avenue for investigating the immunopathogenesis of influenza.

Funding:

This work was supported by the Imperial College Comprehensive Biomedical Research Centre and the Wellcome Trust Centre for Respiratory Infection at Imperial College (DMPED P26077 to JAH) and the Fondo de Investigación Sanitaria del plan nacional de I+D+I and 'fondos FEDER (FIS; PI070069 / PI1000540 to FMT). FMT represents the GENDRES research group (www.gendres.org).

Acknowledgements

We thank the clinicians and patients for taking part, Dr Jayne Dennis who gave invaluable microarray advice and Mary Cross who gave clinical research support.

References

1. Herberg JA, Jones KD, Paulus S, et al. Comparison of pandemic and seasonal influenza reveals higher mortality and increased prevalence of shock in children with severe h1n1/09 infection. Pediatr Infect Dis J **2011**; 30:438-40.

2. Sachedina N, Donaldson LJ. Paediatric mortality related to pandemic influenza A H1N1 infection in England: an observational population-based study. Lancet **2010**; 376:1846-52.

3. Ramilo O, Allman W, Chung W, et al. Gene expression patterns in blood leukocytes discriminate patients with acute infections. Blood **2007**; 109:2066-77.

4. 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.

5. Bermejo-Martin JF, Martin-Loeches I, Rello J, et al. Host adaptive immunity deficiency in severe pandemic influenza. Critical care (London, England) **2010**; 14:R167.

6. Parnell GP, McLean AS, Booth DR, et al. A distinct influenza infection signature in the blood transcriptome of patients with severe community-acquired pneumonia. Critical care (London, England) **2012**; 16:R157.

7. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic acids research **2002**; 30:207-10.

8. Guo J, Hui DJ, Merrick WC, Sen GC. A new pathway of translational regulation mediated by eukaryotic initiation factor 3. EMBO J **2000**; 19:6891-9.

9. Terenzi F, Hui DJ, Merrick WC, Sen GC. Distinct induction patterns and functions of two closely related interferon-inducible human genes, ISG54 and ISG56. J Biol Chem **2006**; 281:34064-71.

10. Andrejeva J, Norsted H, Habjan M, Thiel V, Goodbourn S, Randall R. ISG56/IFIT1 is primarily responsible for IFN-induced changes in patterns of PIV5 transcription and protein synthesis. J Gen Virol **2012**.

11. Daffis S, Szretter KJ, Schriewer J, et al. 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature **2010**; 468:452-6.

12. Everitt AR, Clare S, Pertel T, et al. IFITM3 restricts the morbidity and mortality associated with influenza. Nature **2012**.

13. Mukherjee S, Vipat VC, Mishra AC, Pawar SD, Chakrabarti AK. Pandemic (H1N1) 2009 influenza virus induces weaker host immune responses in vitro: a possible mechanism of high transmissibility. Virology journal **2011**; 8:140.

14. Raven JF, Koromilas AE. PERK and PKR: old kinases learn new tricks. Cell Cycle **2008**; 7:1146-50.

15. Inglis SC. Inhibition of host protein synthesis and degradation of cellular mRNAs during infection by influenza and herpes simplex virus. Mol Cell Biol **1982**; 2:1644-8.

16. Orellana RA, Wilson FA, Gazzaneo MC, Suryawan A, Davis TA, Nguyen HV. Sepsis and development impede muscle protein synthesis in neonatal pigs by different ribosomal mechanisms. Pediatric research **2011**; 69:473-8.

17. Ma J, Sun Q, Mi R, Zhang H. Avian influenza A virus H5N1 causes autophagy-mediated cell death through suppression of mTOR signaling. Journal of genetics and genomics **2011**; 38:533-7.

Figures

Figure 1. Unsupervised clustering of the top SDE transcripts for a) H1N1/09 and controls, b) H1N1/09 and RSV and c) H1N1/09 and bacterial infection

Heatmaps show separation of H1N1/09 and comparator groups by unsupervised Manhattan clustering, based on top SDE transcripts. Each row represents one transcript; each column represents one patient, with a red bar below indicating H1N1/09 and a blue bar a control (a), RSV (b), or bacterial infection (c). Curtailed transcript lists of highly significant probes were used for clarity (*p* value thresholds of 10^{-5} (n=90 transcripts), *0.002* (n=97) and *0.001* (n=156) respectively for control, RSV and bacterial comparisons). Expression intensity is indicated by colour (increased abundance in red, decreased in blue, intermediate in yellow).

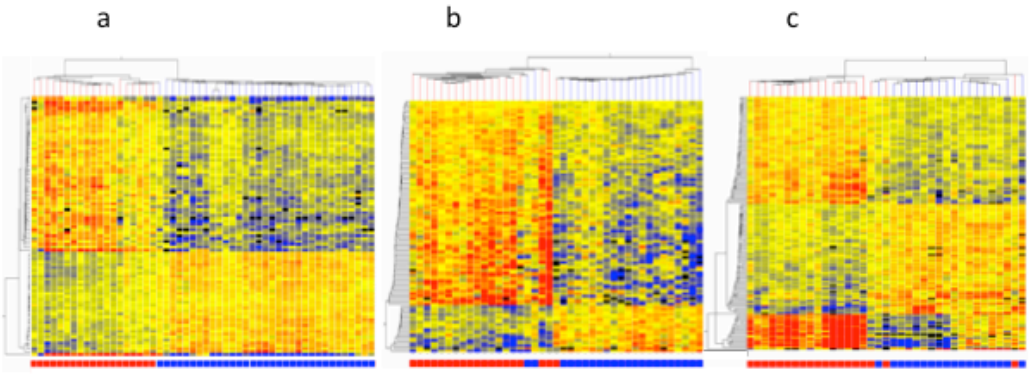
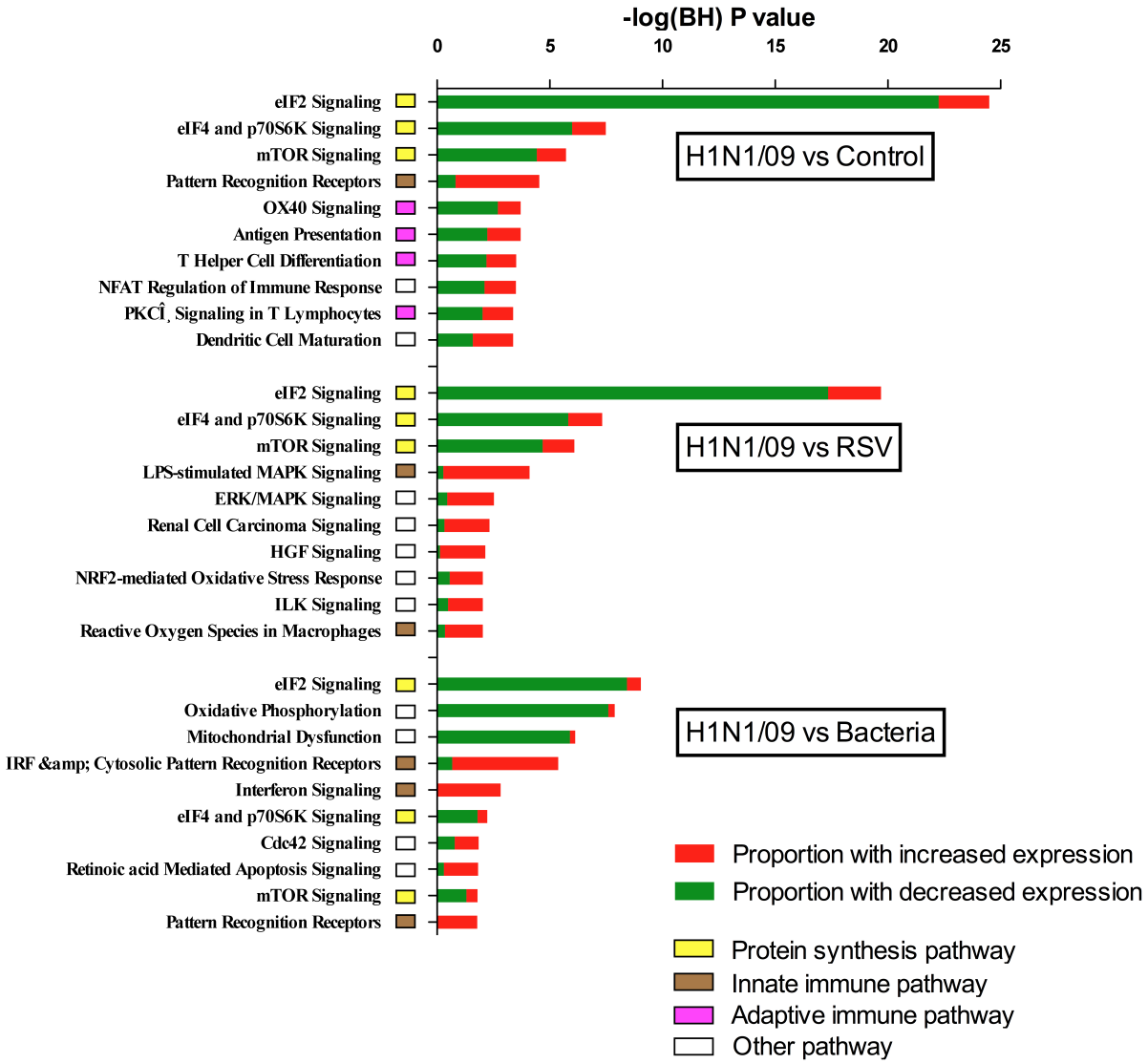


Figure 2. Top Canonical pathways differing between H1N1/09 and controls, RSV and bacterial infection

Each bar is coloured in proportion to the number of SDE H1N1/09 transcripts increased (red) or decreased (green) in abundance relative to the comparator cohort. The total bar length is proportional to *P* value. Coloured blocks next to each pathway are coded according to biological function. Protein synthesis pathways (yellow) were the most significant in all 3 comparisons, with predominant decreased expression in H1N1/09 patients relative to the comparator group. Innate immune pathway transcripts (brown) were increased in H1N1/09 patients, whilst adaptive immune transcripts (pink) were reduced relative to controls.



Supplementary Table 1 - Demographic and clinical data of recruited subjects

	H1N1/09	RSV	Bacterial	Controls	*P* value
Number of patients	25 (+2[a])	34 (+2[a])	18	33	N/A
No co-infection	19	23	13[b]	N/A	
Co-infection	6[c]	11[c]	5		
Bocavirus	5	5	0		
Rhinovirus	2	4	0		
Adenovirus	0	2	1		
Seasonal flu	0	1	1		
parainfluenza	0	0	1		
Metapneumovirus	0	0	2		
RSV	(+1 [a])	N/A	N/A		
H1N1/09	N/A	0	N/A		
S.pneumoniae	(+1 [a])	(+2 [a])	12		
S.pyogenes	0	0	4		
S.aureus	0	0	2		
Sex M:F (% male)	12:13 (48)	21:13 (62)	8:13 (38)	18:15 (55)	*NS*
Age (years): median (IQR)	4.0 (1.6-7.5)	0.4 (0.1-1.4)	1.9 (1.0-4.4)	3.4 (1.5-6.9)	*P<0.0001*
Ethnicity- Caucasian (%)	15 (53)	19 (52)	10 (55)	16 (48)	*NS*
No. additional ethnic groups	7	8	7	8	*NS*
Days from symptoms to recruitment: median (IQR)	5 (3.0-7.0)	4 (2.0-6.3)	3.5 (2.0-10.5)	N/A	*NS*
Intensive care required	12 of 25 (48%)	20 of 34 (59%)	15 of 18 (83%)	n/a	*NS*
Deaths	5	0	1	N/A	*NS*
Pathogen cohort for arrays	19 (without co-infection)	22[d] (without co-infection)	18 (excludes H1N1, RSV)	33	N/A
Lymphocyte proportion (array patients): median (IQR)	0.21 (0.10-0.32)	0.39 (0.28-0.49)	0.17 (0.08-0.25)	0.45 (0.38-0.56)	*P<0.001* for HvsC
Neutrophil proportion (array patients): median (IQR)	0.69 (0.52-0.84)	0.47 (0.40-0.64)	0.74 (0.64-0.87)	0.45 (0.35-0.51)	*P<0.001* for HvsC
Monocyte proportion (array patients): median (IQR)	0.04 (0.01-0.08)	0.09 (0.03-0.15)	0.03 (0.0-0.08)	0.07 (0.06-0.09)	*NS*

NS - not significant (corrected *P<0.05*); IQR – interquartile range; N/A - not applicable.

[a] Two patients each in the H1N1/09 and RSV cohorts with confounding coinfections (RSV or bacterial) were excluded from array analysis and from demographic calculations.

[b]After excluding patients with H1N1/09 or RSV, patients with confirmed gram-positive bacterial infection were analysed irrespective of other viral coinfection - no virological investigations were available for 9 bacterial infection patients recruited outside the pandemic period.

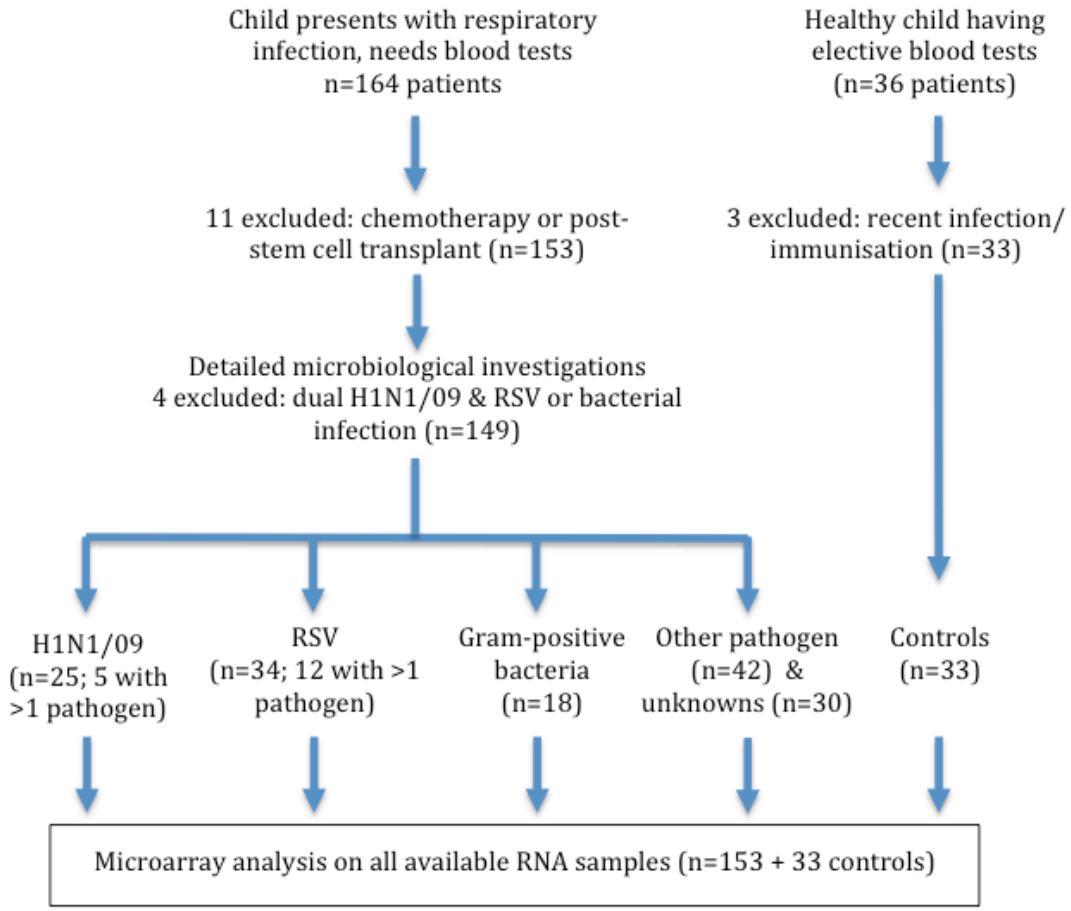
[c]Note that one patient each with RSV and H1N1/09 had two viral coinfections

[d]One RSV patient was excluded as the array data did not pass QC on PCA plot

The gender distribution between cohorts was not different. The ages of the H1N1/09, bacterial and control cohorts were not significantly different. The RSV cohort was younger, as expected for RSV bronchiolitis admissions. Days from symptom onset to recruitment, and deaths in each cohort were not significantly different. Lymphocyte proportion was lower, and neutrophil proportion higher (denominator total leucocytes) in H1N1/09 patients than controls, but was not significantly different when compared to the RSV or bacterial groups.

Supplementary Figure 1. Study overview showing patient numbers and analysis

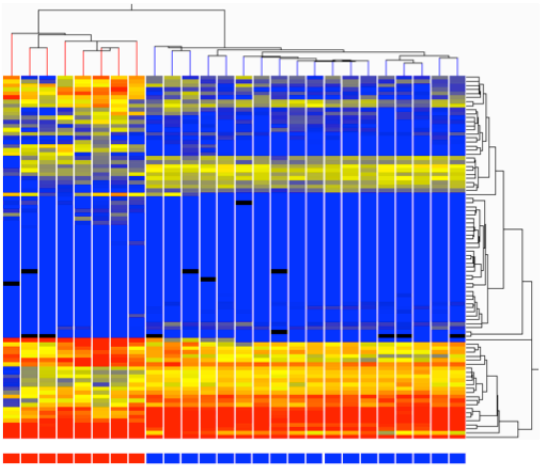
Children with acute infection, sick enough to warrant blood tests, were recruited as early as possible at or after presentation. Respiratory or nasopharyngeal secretions were screened for viruses as part of routine clinical care, before we used the samples for additional viral detection. Patients were assigned to pathogen cohorts on the basis of microbiology results and a compatible clinical syndrome. Controls were recruited in out-patients. They had no recent (previous two weeks) or current infectious symptoms (for instance coryza), nor had any identified or suspected chronic infectious or inflammatory conditions, nor recent immunisation. Children with co-morbidities likely to have a profound impact on gene expression were excluded, including 11 children with either prior bone marrow transplant or on chemotherapy. In the pathogen cohorts we included children with or without other comorbidities, with neurodisability the most common (2 in RSV and bacterial; 4 in H1N1).


Child presents with respiratory infection, needs blood tests n=164 patients
Healthy child having elective blood tests (n=36 patients)
11 excluded: chemotherapy or post-stem cell transplant (n=153)
3 excluded: recent infection/ immunisation (n=33)
Detailed microbiological investigations
4 excluded: dual H1N1/09 & RSV or bacterial infection (n=149)
H1N1/09 (n=25; 5 with >1 pathogen)
RSV (n=34; 12 with >1 pathogen)
Gram-positive bacteria (n=18)
Other pathogen (n=42) & unknowns (n=30)
Controls (n=33)
Microarray analysis on all available RNA samples (n=153 + 33 controls)

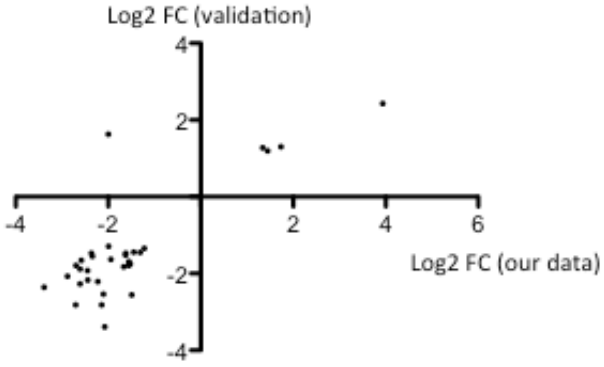
Supplementary Figure 2. Validation of H1N1/09 vs control gene expression on an independent cohort

Data from our cohort were compared with the published adult dataset of Parnell et al [6]. (a): Heatmap showing full segregation of adult H1N1/09 patients and controls by unsupervised Manhattan clustering in the validation dataset, using the top 90 strongest classifiers derived from the 1,267 transcript set defining H1N1/09 vs controls in our own cohort. Each row represents one transcript; each column represents one patient, with a red bar below indicating H1N1/09 and a blue bar a control. Expression intensity is indicated by colour (increased abundance in red, decreased in blue, intermediate in yellow). There was also fully concordant separation when the full list of 1,267 transcripts was used (data not shown). (b): 32 transcripts in the EIF2 signalling canonical pathway were significantly differentially expressed in the validation dataset, with strongly correlating fold change and direction of change from controls (Spearman r=0.63, *P=0.0001*). Upregulated transcripts are plotted in the positive quadrant; negative regulation is plotted in the negative quadrant. One transcript was not concordant (*RPL24*).

a



b



Transcriptomic profiling in childhood H1N1/09 influenza reveals reduced expression of protein synthesis genes

Authors: Jethro A. Herberg[1], Myrsini Kaforou[1], Stuart Gormley[2], Edward R Sumner[3], Sanjay Patel[1], Kelsey DJ Jones[1,2], Stéphane Paulus[1], Colin Fink[3], Federico Martinon-Torres[4], Giovanni Montana[5], Victoria J. Wright[1], Michael Levin[1]

Online data supplement: supplementary methods

Microarray Analysis

Expression data were analysed using '*R' Language and Environment for Statistical Computing 2.12.1* [18] and GeneSpringGX 11.5 software (Agilent).

Mean raw intensity values for each probe were corrected for local background intensities, and quantile normalised. The dataset was filtered to exclude probes that were flagged as 'present' on less than 90% of the arrays in at least one group of interest. Expression values were transformed to a $\log_2$ scale. Principal component analysis was used to review the quality of all arrays (one RSV outlier was removed), and to identify significant influences on the data other than disease cohort, such as technical factors, comorbidities or concurrent medications.

In order to ensure that analyses carried out using the GeneSpring platform were valid, and due to infection type, data were first analysed in R in order to determine the influence of age, gender and technical batch on the expression data. These were considered as independent factors in a linear regression model. After normalization, batch, age and gender were found to have a weak influence on the expression values. Technical batch effects (particularly array chip) were detected during unsupervised clustering of unnormalised data, but the experimental design mitigated these effects through the randomization of samples at each stage of sample processing - RNA extraction, labelling, hybridisation. The raw background-subtracted data were therefore normalized for further analysis without further adjustment.

The hypothesis that the expression level for each probe differed between comparator patient groups was assessed using Welch's moderated t-test [19]. *P* values were adjusted using Benjamini and Hochberg's method to control for the false discovery rate [20]. For each comparison of interest the most significant probes were selected, based on *P* value and absolute fold-change >2.

We compared each infection cohort to controls to derive a list of significantly DE transcripts for each comparison with $P<0.05$ and $\log_2 FC>1$ (Supplementary Data S1). When comparing the transcriptional response of two infection cohorts, we included the union of DE transcripts between healthy controls and either pathogen.

The Support Vector Machine (SVM) method for supervised learning was used to classify patients into groups, based on our pre-defined signatures. We applied a linear SVM to define a hyperplane in a high-dimensional transformed feature space that maximally discriminated two patient groups. We used leave-one-out cross-validation to calculate the classification accuracy.

References

1. Herberg JA, Jones KD, Paulus S, et al. Comparison of pandemic and seasonal influenza reveals higher mortality and increased prevalence of shock in children with severe h1n1/09 infection. Pediatr Infect Dis J **2011**; 30:438-40.
2. Sachedina N, Donaldson LJ. Paediatric mortality related to pandemic influenza A H1N1 infection in England: an observational population-based study. Lancet **2010**; 376:1846-52.
3. Ramilo O, Allman W, Chung W, et al. Gene expression patterns in blood leukocytes discriminate patients with acute infections. Blood **2007**; 109:2066-77.
4. 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.
5. Bermejo-Martin JF, Martin-Loeches I, Rello J, et al. Host adaptive immunity deficiency in severe pandemic influenza. Critical care (London, England) **2010**; 14:R167.
6. Parnell GP, McLean AS, Booth DR, et al. A distinct influenza infection signature in the blood transcriptome of patients with severe community-acquired pneumonia. Critical care (London, England) **2012**; 16:R157.
7. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic acids research **2002**; 30:207-10.
8. Guo J, Hui DJ, Merrick WC, Sen GC. A new pathway of translational regulation mediated by eukaryotic initiation factor 3. EMBO J **2000**; 19:6891-9.

9. Terenzi F, Hui DJ, Merrick WC, Sen GC. Distinct induction patterns and functions of two closely related interferon-inducible human genes, ISG54 and ISG56. J Biol Chem **2006**; 281:34064-71.
10. Andrejeva J, Norsted H, Habjan M, Thiel V, Goodbourn S, Randall R. ISG56/IFIT1 is primarily responsible for IFN-induced changes in patterns of PIV5 transcription and protein synthesis. J Gen Virol **2012**.
11. Daffis S, Szretter KJ, Schriewer J, et al. 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature **2010**; 468:452-6.
12. Everitt AR, Clare S, Pertel T, et al. IFITM3 restricts the morbidity and mortality associated with influenza. Nature **2012**.
13. Mukherjee S, Vipat VC, Mishra AC, Pawar SD, Chakrabarti AK. Pandemic (H1N1) 2009 influenza virus induces weaker host immune responses in vitro: a possible mechanism of high transmissibility. Virology journal **2011**; 8:140.
14. Raven JF, Koromilas AE. PERK and PKR: old kinases learn new tricks. Cell Cycle **2008**; 7:1146-50.
15. Inglis SC. Inhibition of host protein synthesis and degradation of cellular mRNAs during infection by influenza and herpes simplex virus. Mol Cell Biol **1982**; 2:1644-8.
16. Orellana RA, Wilson FA, Gazzaneo MC, Suryawan A, Davis TA, Nguyen HV. Sepsis and development impede muscle protein synthesis in neonatal pigs by different ribosomal mechanisms. Pediatric research **2011**; 69:473-8.
17. Ma J, Sun Q, Mi R, Zhang H. Avian influenza A virus H5N1 causes autophagy-mediated cell death through suppression of mTOR signaling. Journal of genetics and genomics = Yi chuan xue bao **2011**; 38:533-7.
18. R Development Core Team. R: A language and environment for statistical computing. Vienna, **2012**.
19. Welch BL. The generalization of 'Students' problem when several different population variances are involved. Biometrika **1947**; 34:28-35.
20. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Ann Statist **2001**; 29:1165-88.