Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Network analysis of estradiol-treated, influenza infected epithelial cells identifies hormone-regulated host replication factors

Pariksheet Nanda
University of Pittsburgh
Lauren M. Nichols
University of Pittsburgh
Tatum S. Liparulo
University of Pittsburgh
John F. Alcorn
UPMC Children's Hospital of Pittsburgh
Rudiyanto Gunawan
University at Buffalo
Jason E. Shoemaker
jason.shoemaker@pitt.edu
University of Pittsburgh

Article

Keywords:

Posted Date: November 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11570726/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

NETWORK ANALYSIS OF ESTRADIOL-TREATED, INFLUENZA INFECTED EPITHELIAL CELLS IDENTIFIES HORMONE-REGULATED HOST REPLICATION FACTORS

A PREPRINT

Pariksheet Nanda
Department of Chemical Engineering
University of Pittsburgh
Pittsburgh, 15261
pan79@pitt.edu

Lauren M. Nichols
Department of Chemical Engineering
University of Pittsburgh
Pittsburgh, 15261
lan120@pitt.edu

Tatum S. Liparulo
Department of Chemical Engineering
University of Pittsburgh
Pittsburgh, 15261
tam193@pitt.edu

John F. Alcorn
Department of Pediatrics
UPMC Children's Hospital of Pittsburgh
Pittsburgh, 15224
john.alcorn@chp.edu

Rudiyanto Gunawan
Department of Chemical and Biological Engineering
University at Buffalo
Buffalo, 14260
rgunawan@buffalo.edu

Jason E. Shoemaker
Department of Chemical Engineering
University of Pittsburgh
Pittsburgh, 15261
Department of Computational and Systems Biology
University of Pittsburgh
Pittsburgh, 15260
McGowan Institute for Regenerative Medicine
University of Pittsburgh
Pittsburgh, 15219
jason.shoemaker@pitt.edu

October 31, 2026

ABSTRACT

Females of reproductive age (15–49 years of age) experience higher influenza morbidity and mortality than males, yet the contribution of sex hormones to influenza A virus (IAV) infection has been little studied, and network-based approaches have not been applied. Here, we performed a meta-analysis of gene expression data from primary human nasal epithelial cells derived from female donors that were treated with 17β-estradiol (E2) and then challenged with IAV infection. We applied differential expression analysis and a network algorithm, DeMAND, to identify genes and pathways regulated by E2 during infection. DeMAND uses protein interaction networks and the variation among replicates to identify dysregulated gene pairs and mechanism of action (MoA) genes. The two approaches identified largely distinct gene sets (5–11% overlap). Genes identified by DeMAND, but not by differential expression, were significantly enriched for known IAV–host interaction and replication factors. Genes identified by DeMAND were enriched for estrogen-related pathways at 48 hours, including ESR1, AGO1, HSP90AB1—a steroid receptor chaperone that was previously proposed as a host-directed antiviral target. Both approaches detect type I interferon and cytokine signaling. Network analysis also prioritized eight candidate transcription factors, most with no prior link to influenza, and drug enrichment analysis found the estrogen receptor antagonist fulvestrant. Our results add to the growing evidence in literature that network analysis can complement differential expression for identifying mechanisms regulating virus replication. This work identifies new candidate host replication factors that are E2-regulated, suggesting novel pathways linking circulating hormones to IAV replication that can be experimentally examined in future work.

Keywords sex differences • protein-protein interaction • network analysis • enrichment analysis • gene expression • microarray

version-230-g3099b76

1 Introduction

Globally, influenza-associated respiratory deaths number 291,000–645,000 annually[1] and in the United States influenza remains a major cause of morbidity and mortality[2] with approximately 8% of the US population each season experiencing symptomatic influenza infection[3]. Despite significant epidemiological evidence of higher influenza morbidity and mortality in women of reproductive age (15–49 years), there has been limited study of hormone regulation of influenza infection[4,5] and, to our knowledge, no one has attempted to use a network-based approach to identify mechanisms linking hormones to influenza virus replication or the associated immune response[6–11]. Here, we conduct a network-based analysis using microarray data of differentiated human nasal epithelial cell (hNEC) cultures from female donors[12] to identify new potential pathways and mechanisms of action linking estradiol to influenza infection in women.

Estrogens have been shown to modify the immune response[13–15] and to possibly impact influenza virus replication[12,16–18]. There are four principal endogenous estrogens (E1–E4)[14]. Estrogen receptors are of two types: nuclear receptors—including ER-α and ER-β—and membrane-initiated steroid signaling (MISS) receptors—including G-protein-coupled estrogen receptor (GPER) and GPR30[19,20]. Nuclear receptors have been identified in innate immune cell types (macrophages, monocytes, neutrophils, natural killer cells, dendritic cells) and adaptive immune cell types (T cells, B cells)[21,22]. ER signaling has effects on cell proliferation, growth, cytokine activity (interferons, IL-6, IL-1, VEGF, amphiregulin, TGF-β), signaling pathways including the proinflammatory response (NF-κB, STAT, TGF-β, TNF), and transcription factors and their coregulators (c-Fos, c-Myc, Myb, JunB)[15,23–25]. Estrogen is involved in several inflammatory disorders—dermatitis, rheumatoid arthritis, and systemic lupus erythematosus—and it reduces wound inflammation, but it also has protective effects, such as with atherosclerosis, and has promise in the treatment of osteoarthritis in women over 65 years of age and in the treatment of multiple sclerosis[21]. Together, several potential, context-specific inflammatory and immune response pathways merit consideration when assessing the impact of estrogen on influenza infection replication and the induced immune response.

Estradiol (E2) is the primary estrogen for women of reproductive age, has bipotential effects on immune regulation, and may alter influenza replication. High concentrations of E2 increase the activity of Type 2 T helper cells (Th2), which produce anti-inflammatory cytokines such as IL-4 and IL-5 that cause increased antibody production and eosinophil activation[14]. Low concentrations of E2 induce human monocytes and macrophages to produce IL-6 and TNFα, and generally induce a pro-inflammatory type 1 T helper cell response (Th1)[13,14]. Human females also have greater type I interferon (IFN) activity than males, likely due to E2 promotion of monocytes into inflammatory dendritic cells (DCs) that produce IFN and other pro-inflammatory cytokines[26,27]. Human female-derived natural killer (NK) cells show greater activation levels after IFN stimulation compared to males[28]; however, males have higher NK cell counts[29]. Human female-derived neutrophils have significantly higher expression of interferon-stimulated genes (ISGs) compared

to males, which is likely to be sex hormone driven[30]. ISGs are critical regulators of antiviral responses[27]. In regard to influenza virus replication, Peretz et al. isolated nasal epithelial cells from female and male donors and determined that pre-treating cells from female donors (but not male) with E2 led to significantly lower numbers of virions[12]. This leads to two possibilities: E2 induces antiviral pathways prior to infection in female-derived cells, or E2 differentially regulates many aspects of the immune response, directly impacting virus replication. These possibilities suggest that a systems-level approach that captures gene regulatory interactions is warranted and may reveal novel mechanisms linking E2 to virus replication or immunity.

Network modeling uses molecular interactome data and gene expression data to detect possible regulatory mechanisms, and has been successful in the context of viral infection at identifying new host–viral interactions[31,32], predicting antiviral drug targets[33–35], revealing new regulatory pathways, and providing new insights into infection progression. Network analysis is a topology-based approach to biological pathway analysis[36] and has been shown to outperform differential expression analysis in identifying relevant pathways[36], signaling mechanisms and effects of perturbation[37], and targets for drug mechanism of action (MoA)[38]. Unlike dynamical mathematical models of biological pathways, network analysis broadly considers cellular interactions and does not require estimates of kinetic parameters, flux measurements, microscopy images, or other training data[39,40]. Among network modeling metrics, we have highlighted the predictive power of the betweenness network topology measure and published the TREAP algorithm[41] that uses both betweenness and gene expression to quickly and accurately identify candidate drug targets. Similarly, we have shown that using controllability analysis of host protein interaction networks or creating host-virus interaction subnetworks are powerful approaches for identifying immune signaling pathways induced by influenza infection and host–influenza replication factors (i.e., host proteins that are needed by the virus to replicate)[42,43].

Here, we investigate published microarray gene expression data[12] using network modeling to determine how E2 modifies the host response in nasal epithelial cells infected by influenza A virus (IAV) and immunomodulates inflammation as well as identify any relationships to epithelial tissue repair for recovery from infection. We apply the DeMAND network modeling algorithm[38] in the context of IAV infection and estradiol treatment primarily at 48 hours post-infection. We show that results of network analysis are largely divergent from differential analysis; network analysis enriches for cell cycle and estrogen signaling pathways whereas differential gene expression enriches for methylation, programmed cell death, and meiosis pathways. The network analysis significantly identifies relevant host factors in viral replication screens. Half the transcription factors we uncover were previously unstudied and only three have been implicated in any viral infection. Thus, we reveal the potential roles of E2 in IAV infection by identifying candidate host replication proteins, virus replication pathways, and immune pathways that are regulated by E2.

2 Materials and methods

All computations beginning with the input data and resulting in figures of this paper are reproducible using the freely available code in the Data availability section.

2.1 Gene expression data and analysis

We used the processed normalized log2 microarray intensity datasets uploaded to the Gene Expression Omnibus (GEO) accession GSE75208 to maintain consistency with the original analysis[12]. This microarray data only consists of primary female human nasal epithelial cell (hNEC) cultures because female-derived cells responded to estradiol (E2) exposure by suppressing virus replication whereas virus replication was unchanged in equivalent male cell cultures exposed to either E2 or testosterone[12].

To minimize annotation system conversion losses, wherever possible we use NIH Entrez gene identifiers. While the Bioconductor hugene20sttranscriptcluster.db annotation package[44] maps Entrez gene identifiers to the Affymetrix Human Gene (HuGene) 2.0 ST array probe identifiers, there were instances where (i) probesets mapped to no Entrez gene IDs, (ii) probesets mapped to multiple Entrez gene IDs, and (iii) multiple probesets mapped to the same Entrez gene ID. From the microarray 53,617 probesets, we removed probesets mapping to multiple Entrez gene IDs (46% loss, n = 24,921) and, among the remaining probesets, aggregated multiple probesets mapping to the same Entrez gene ID (9% aggregation, n = 2,675) using $\log_2\left(\sum 2^x\right)$ where x is a vector of normalized probeset log2 expression mapping to the same Entrez gene ID. Thus, normalized expression from 53,617 microarray probesets was mapped to 26,021 Entrez genes.

We applied a limma version 3.62.2[45] gene expression linear model with the design matrix encoding all 8 possible comparisons of treatment and infection conditions listed in Table 1 to obtain corresponding gene log2-fold-change values. These limma fold changes are only used for the subsequent differential expression analysis and not for performing network analysis.

Table 1: Regression contrasts applied to gene expression data throughout this study. "Treatment" always refers to the use of estradiol and "infection" always refers to the use of influenza A virus. Combinations of treatment and infection are collectively called the sample "condition". The contrast portion of the condition is indicated in the "Varied" column of the table, and the common portion of the condition is indicated in the "Fixed" column. Samples were collected at 24 and 48 hours. The regression contrasts were applied using the limma Bioconductor R package and are provided in the supplementary code. Not all results had significant hits for each of these contrasts and such instances are discussed in detail in Figures S1 and S2. See Figure 1 for the corresponding experimental system recapitulated from Peretz et al. 2016.

Symbol	Time (hours)	Fixed	Varied
$t_{24}inf^{-}trt^{\pm}$	24	Uninfected	Treated (estradiol) vs. Untreated
$t_{24}trt^{-}inf^{\pm}$	24	Untreated	Infected (virus) vs. Uninfected
$t_{24}inf^{+}trt^{\pm}$	24	Infected (virus)	Treated (estradiol) vs. Untreated
$t_{24}trt^{+}inf^{\pm}$	24	Treated (estradiol)	Infected (virus) vs. Uninfected
$t_{48}inf^{-}trt^{\pm}$	48	Uninfected	Treated (estradiol) vs. Untreated
$t_{48}trt^{-}inf^{\pm}$	48	Untreated	Infected (virus) vs. Uninfected
$t_{48}inf^{+}trt^{\pm}$	48	Infected (virus)	Treated (estradiol) vs. Untreated
$t_{48}trt^{+}inf^{\pm}$	48	Treated (estradiol)	Infected (virus) vs. Uninfected

2.2 DeMAND network analysis

We built a protein interaction network from all Entrez genes uniquely matching microarray probes, regardless of expression level, using STRING version 12.0[46] thresholded by a confidence score of 0.8. There was a 31% loss converting Entrez gene identifiers to STRING Ensembl identifiers resulting in 17,954 gene-protein pairs and 308,499 protein-protein interactions. There will be further loss when reconciling these with the microarray probes.

Using the interaction network, we ran the Detecting Mechanism of Action by Network Dysregulation (DeMAND) algorithm[38] that enriches for interactions between gene pairs using Kullback–Leibler divergence (KLD) and then uses those gene pairs to predict mechanism of action (MoA) genes. DeMAND does not conceptually distinguish between genes and proteins in an interaction network. In short, the DeMAND algorithm detects dysregulation of gene networks by comparing the Gaussian kernel smoothed probability density of a gene pair and compares how this calculated density changes between treatment and the control. The change of gene pair replicates is measured by the KLD and the statistical significance of the KLD is measured by gene shuffling. Finally, the KLD significance values are propagated into single MoA genes using Brown's method[38, see methods]. In our analysis of this microarray dataset[12], we find about 50% and 42% of the genes are common between the KLD and MoA network methods at the 24-hour and 48-hour timepoints, respectively. Thus, while the MoA genes are the primary output of DeMAND, the secondary output of KLD gene pairs directly measures dysregulation between the two gene networks to help understand treatment effects.

2.3 Non-parametric gene-set selection across enrichment methods

The magnitudes and spread of p-values differ significantly among the three methods applied here of differential expression, KLD, and MoA. Choosing a set of significant genes using a typical distribution-dependent approach, such as a p-value cut-off of 0.001, yields very different numbers of 0–704 significant genes across the methods shown in Figure S2. Therefore, we instead use the top 1 percentile of genes for distribution-agnostic gene-set selection. Due to the different 26,021 input genes to differential expression and 13,386 input genes to DeMAND, the 1 percentile of genes obtained therefore also differ, at 260 and 133, respectively, in Figure 2.

2.4 Significance testing with previously collected viral–host protein datasets

We used previously collected datasets of host proteins involved in viral interactions[47, Table S5 (n = 1,449)], [11, Table S1A (n = 1,292)], and viral replication[11, Table S2A (n = 330)]; these datasets were subset to the 26,021 genes measured by the microarray resulting in a total of 1,283, 1,223, and 321 known host proteins, respectively. Fisher's Exact test was performed in R using the greater alternative hypothesis to test for positive associations against the complete set of 26,021 genes. All 26,021 genes were used regardless of method because subsetting to the protein-protein interaction network is part of the network enrichment method workflow.

2.5 Enrichment using ToppGene and Bioconductor interface

To run comprehensive enrichment analyses on our gene sets identified by DeMAND network analysis, we ran Sugeno fuzzy-measure-based similarity (FMS) against several curated databases using ToppGene[48]. ToppGene uses NIH Entrez gene identifiers; therefore, as explained earlier, we primarily used Entrez annotations to avoid annotation system conversion losses. We developed a Bioconductor package[49] to communicate with the ToppGene web server via its OpenAPI 3.0 specification to minimize code complexity and support reproducible research.

Although ToppCluster is intended to compare multiple gene set enrichments using the underlying FMS mathematics of ToppGene, ToppCluster requires that the gene set sizes must be of similar cardinality[50]. In our case, the significant DeMAND KLD and MoA gene set sizes differed considerably owing to their different numbers of input genes, therefore using ToppGene was more appropriate than ToppCluster.

We chose input genes from differential expression and network analysis methods using top percentiles / ranks as explained in Section 2.3. This is due to incomparable significance calculations of the two methods that, thus, rule out absolute significance cut-offs.

We standardized the drug identifiers from the various databases to PubChem compound identifiers (CIDs) using the PubChem Identifier Exchange Service[51]. The standardized PubChem CIDs allow aggregating drug identifiers from the various databases.

The reproducibility of any data analysis that relies on external databases is likely to change over time as those databases change. Wherever possible we use versioned databases, however ToppGene uses the latest databases and therefore the significance, rankings, and hits may diverge from those presented in this work when the queried databases are updated.

3 Results

We analyzed a previously collected influenza A virus (IAV) infection dataset of patient-derived epithelial cell cultures sampled at two timepoints and included treatment with a sex hormone, estradiol[12]. To help contextualize the data and analysis, a graphical overview of our workflow including procedures done in previous work[11] is shown in Figure 1.

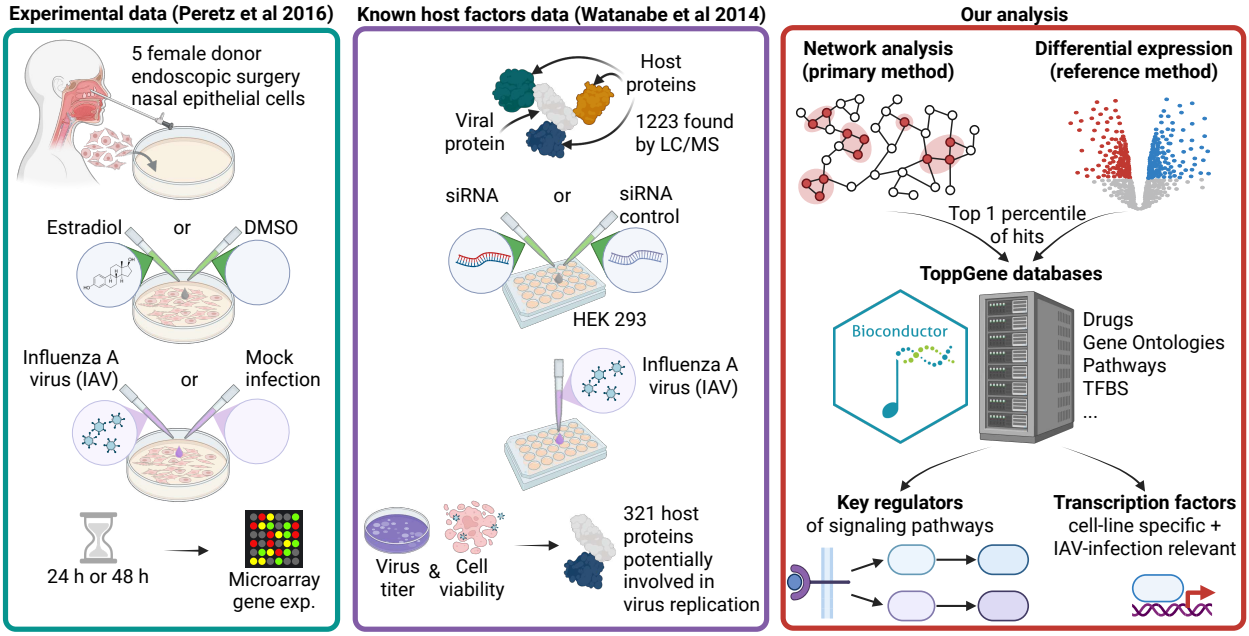


Figure 1: Graphical abstract detailing procedures of previously collected data and our analysis workflow. We use previously collected experimental data and known host factor data for our study. The experimental data (Peretz et al. 2016) consists of clinically derived female nasal epithelial cells that were cultured and treated with estradiol (E2) or a dimethyl sulfoxide (DMSO) control and then IAV or mock infection. The gene expression of 5 replicates of each of those combinations of samples was assayed by microarray at 24 and 48 hours. Host factors involved in IAV replication (Watanabe et al. 2014) were identified by liquid chromatography–mass spectrometry (LC/MS) and then subjected to an RNAi screen of host proteins in HEK 293 cells. The samples in the RNAi screen were then IAV infected and assayed for viral titer and cell viability to identify host factors involved in IAV replication. While not illustrated here, we also used a larger list of IAV interactions across six studies, previously compiled in Ackerman et al. 2018. Our analysis compares the DeMAND network analysis method against differential expression using the same experimental and known host factor data. We performed gene set enrichment analysis (GSEA) and used the top 1 percentile of significant hits from both methods were used to query the ToppGene web service with our `toppgene` Bioconductor R package. We used ToppGene to query many databases for candidate drugs, gene ontologies, signaling pathways, and transcription factor binding sites. We explore these hits to uncover key signaling regulators and enriched transcription factors for our experimental context. Created in BioRender. Nanda, P. (2026) https://BioRender.com/rz73hyx.

3.1 Network and differential expression analyses identify largely non-overlapping gene sets

The top 1 percentile significant gene sets in the UpSet plot of Figure 2 show little overlap between DeMAND and differential expression. Only 5%–11% of genes overlap with over half being unique to differential expression. 20%–54% of genes found significant by differential expression are not present in the protein interaction network. Therefore, even when considering fewer input genes, network analysis uncovers a very different set of genes than differential gene expression.

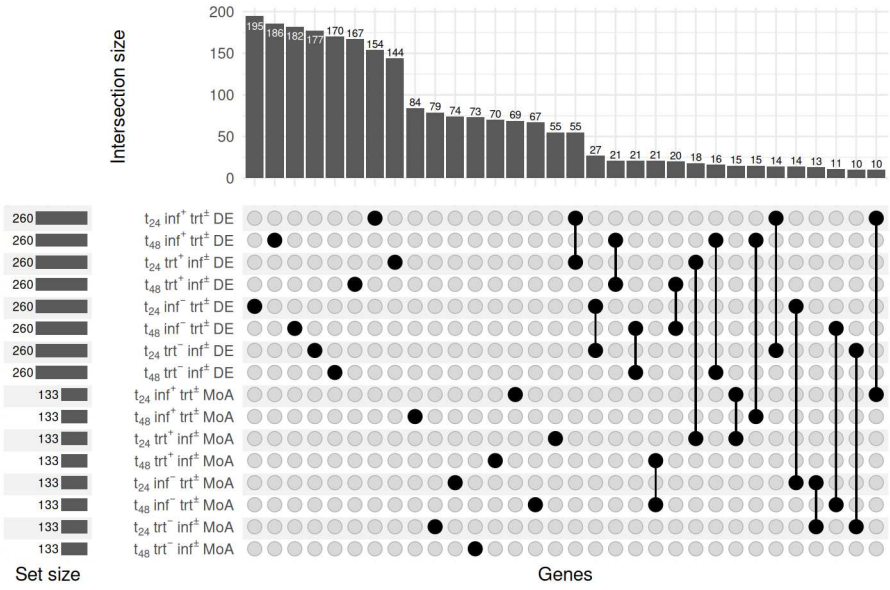


Figure 2: Among the top 1 percentile of hits, DeMAND identifies distinct genes compared to DE. It is necessary to use an UpSet plot instead of a Venn diagram to represent gene set intersections among the 16 combinations of timepoints (t_{24}, t_{48}), infection conditions (inf^+, inf^-, $inf^\pm$), treatment conditions (trt^+, trt^-, $trt^\pm$), and inference methods of differential expression (DE) and DeMAND mechanism of action (MoA) using the sample naming scheme described in Table 1. The UpSet plot consists of 3 parts: each row represents a sample set of genes, each column represents a set intersection, and the row–column dots and lines represent the relationship of intersections between sets being performed; no intersection line between dots indicates that portion of the set is sample exclusive. Each row consists of the top 1 percentile of genes (DE n = 260 / 26,021, MoA n = 133 / 13,386). Intersections of size 9 or lower are omitted for brevity and because only 5%–11% of genes identified by these methods overlap at all.

DeMAND network analysis only computes significance and not fold-change: therefore, the plots in Figure S1 preserve the fold-change values from DE and show how network analysis reprioritizes genes that would not otherwise be considered by DE. Therefore, we see genes ignored at the bottom of the DE volcano plots gain higher significance using network analysis.

3.2 Network analysis suggests multiple host replication factors are regulated by E2

E2 has been hypothesized to impact IAV replication[12]; therefore, we next determined whether the proteins identified as important by DeMAND may be IAV replication host factors. Replication host factors are host proteins that play either a direct or indirect role in supporting the virus replication process. We assessed how well the significant genes identified host interaction factors (proteins that co-immunoprecipitated with IAV proteins) and host replication factors. We considered host interaction and host replication factors identified in a screen single[11] and we considered a compendium of IAV host replication studies we collected previously[47]. Our comparison in Figure 3 shows that DeMAND better captures these known host factors than differential expression. Individual genes of known host factors found by both methods are shown in Table S1 and individual genes within pathways found by both methods are shown in Figure S3. In Figure S3, both methods identified immune-related cytokine signaling genes using translation initiation factors (EIF2AK2, EIF4E), but network analysis was able to additionally contextualize cellular responses to viral infection. Altogether, our known host factor analysis found higher enrichment of IAV–host interaction and replication factors using network analysis.

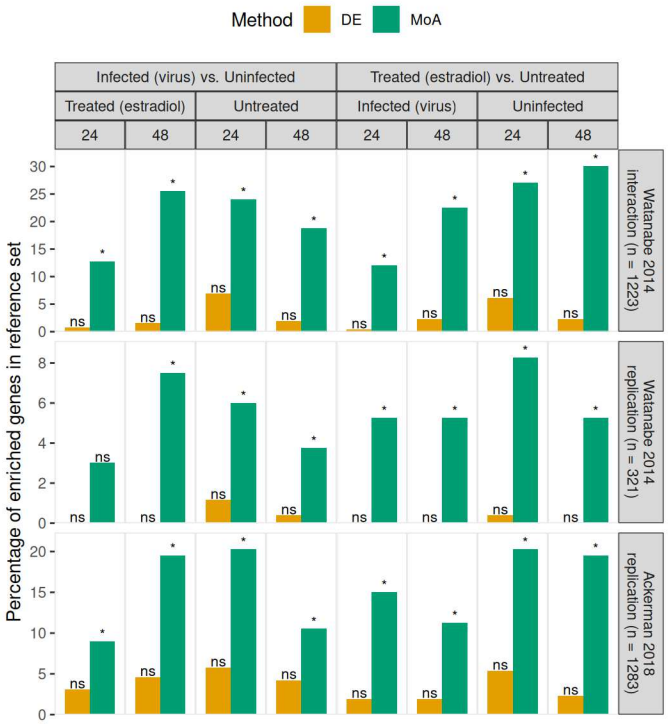


Figure 3: DeMAND significantly identifies IAV–host protein interactions and IAV-replication host factors. Rows represent the comparison set of known host factors for influenza A virus interaction or replication. Percentages represent the total number of method-identified genes present in the set of known host factors. In no case does differential gene analysis identify a significant number of genes (more than expected by chance using Fisher's Exact test). In all cases, except for the 24-hour estradiol treated sample, DeMAND identifies a significant number of genes. Virus replication host factors exclusively identified as significant by MoA are further investigated in Figure 4. Significance was measured using Fisher's Exact test and is represented as follows: $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), otherwise not significant (ns).

When investigating the top host–IAV replication genes uniquely uncovered by DeMAND (Figure 4), we find six genes identified in multiple conditions (RPS7, FUS, EIF3A, DHX13, GNAI2, HSP90AB1) but no gene is significantly present at both 24 and 48 hours, indicating strong time dependence. We find HSP90AB1 (heat shock protein 90 alpha family class B1) is the only estrogen receptor signaling protein enriched when contrasting infection both with and without estrogen at the 48-hour timepoint; HSP90AB1 was the most enriched host factor that enhanced porcine deltacoronavirus (PDCoV) infection in epithelial LLC-PK cells[52] and HSP90AB1 was one of three therapeutic targets in the estrogen signaling pathway in lung cells identified for anti-inflammatory treatment of SARS-CoV-2 infection[53]. Ribosomal protein RPS7 is present in all but the estrogen untreated condition at 24 hours and has previously been identified as an upregulated hub gene during COVID-19 infection[54]. Therefore, we suspect the E2-regulated HSP90AB1 estrogen receptor signaling protein alters influenza infection at the 48-hour post-infection timepoint.

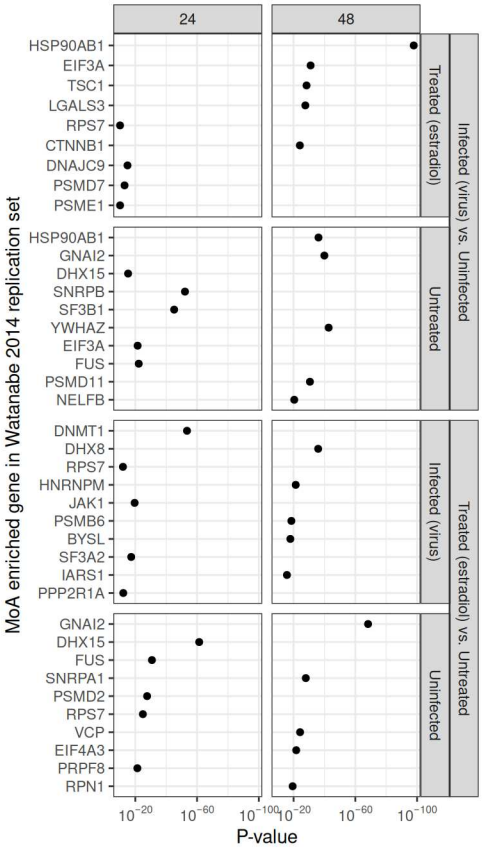


Figure 4: DeMAND identifies distinct sets of top 5 genes at 24 and 48 hours among known IAV host replication factors. We intersect the significant genes found only by DeMAND MoA with the known host factors for replication (Watanabe et al. 2014). The top 5 genes found at 24 and 48 hours for each of the conditions do not intersect, but six genes (RPS7, FUS, E1F3A, DHX13, GNAI2, HSP90AB1) appear in at least two sets of conditions.

3.3 Estrogen receptor signaling is a highly ranked pathway found by network analysis

Our pathway analysis in Figure 5 found that both differential expression and network analysis methods were able to detect type I interferon and cytokine signaling in the top 20 hits, but only network analysis was able to uncover relevant estrogen signaling and responses and the influenza infection pathway. The uncurated list of top 20 pathways hits is shown in Figure S4, which includes fewer notable pathways. There were no pathway hits at the 24-hour timepoint even using a more relaxed corrected false discovery rate (FDR) cut-off of 0.05; therefore, we limit our pathway analysis to the 48-hour timepoint. We see modulation of cytoskeleton (Rho GTPases activate PKNs) by IAV infection consistent with previous studies[55] and that estradiol is associated with less disruption of the cytoskeleton, although the implications for viral infection are unclear[55]. When contrasting the effect of infection, differential expression detects the IL-35 anti-inflammatory pathway and network analysis detects the IL-7 pro-inflammatory pathway. There are overlapping hits when contrasting the effect of estradiol and contrasting the effect of infection that aid interpretation of common pathways; for example, we see higher significance of ESR mediated signaling when contrasting the effect of estrogen. Therefore, network analysis is able to better identify contextual pathways.

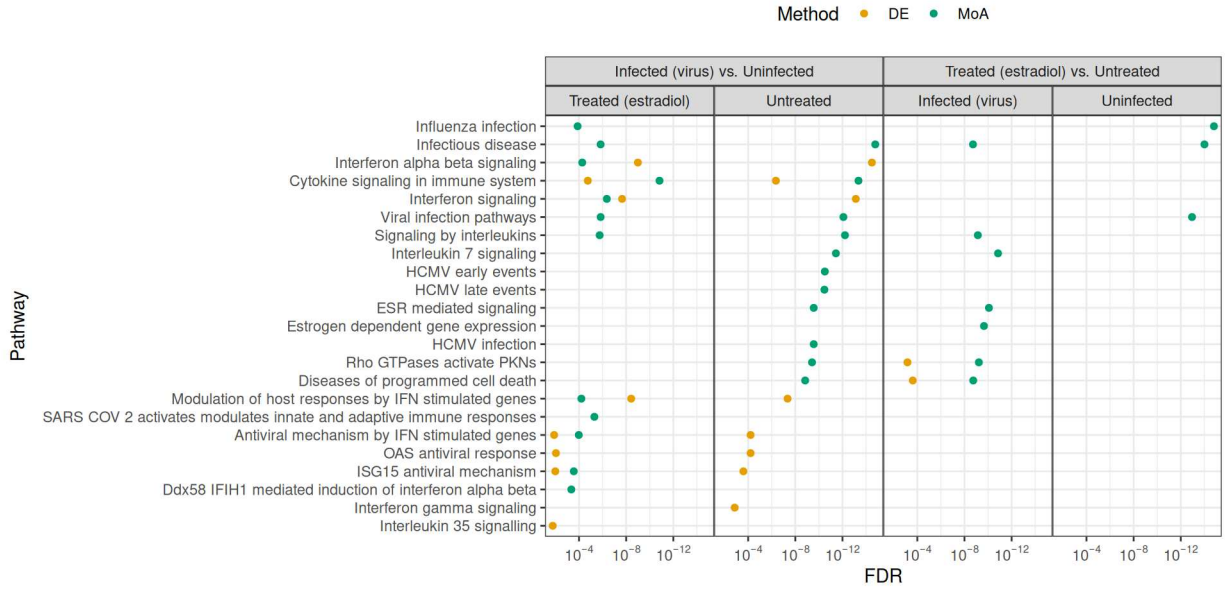


Figure 5: Among the top 20 hits at 48 hours, DeMAND finds relevant estrogen-related pathways, but both methods detect type I interferon and cytokine signaling. The Y-axis pathways are ranked by significance (FDR < 0.001) and curated from the top 20 Reactome pathways identified by each of the methods at the 48-hour timepoint. We subset the top 20 pathway hits to curated selections of particular interest to the experimental system: these include host immune pathways, cytoskeletal modulation by IAV infection, and estrogen-related pathways. Unlike the gene ontology analysis in Figure S5, where there was little agreement among the top hits, in this pathway analysis there is generally more agreement between the two methods. Equivalent Figure S4 does not apply any subsetting.

We inspect estrogen-dependent gene expression pathway in Figure 6 to better understand the contribution of the enriched genes across methods. Differential expression identifies histone H2B and H3 proteins in the estrogen-dependent gene expression pathway. Network analysis uncovers heat shock protein 90 (HSP90AB1) that is essential for ER-α dimerization[56] and is upregulated in airway epithelial cells during asthma[57]. The network methods (KLD, MoA) additionally identify several interesting host factors: Argonaute-1 (AGO1) is known to be an estrogen-dependent transcription coactivator[58], and structural maintenance of chromosomes 1A (SMC1A) interacts with BRCA1[59]. The mechanism-of-action (MoA) enrichment step additionally implicates estrogen receptor 1 (ESR1). Thus, while both methods find histone proteins, network analysis better captures the estrogen-dependent gene expression pathway by detecting relevant transcription factors and transcription co-activators.

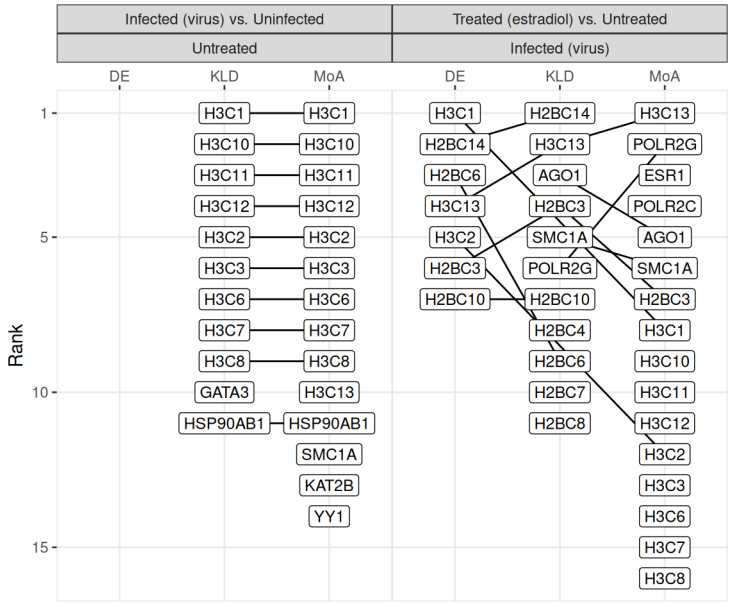


Figure 6: DeMAND's mechanism of action (MoA) metric best detects the pathway of estrogen-dependent gene expression at 48 hours. Plot columns show ranks of the significant genes (FDR < 0.05) of the estrogen-dependent gene expression pathway (Reactome R-HSA-9018519) identified by each of the methods at the 48-hour timepoint. Lines connect common genes to illustrate any changes in rank among the methods. Note that although differential expression includes significant genes (FDR < 0.05), these genes were not sufficient to enrich the estrogen-related gene expression pathway in Figure 5. DE = differential expression; KLD = Kullback–Leibler divergence (see Methods); MoA = mechanism of action.

3.4 Lower-significance antiviral responses with estradiol

Gene ontology of biological processes in Figure S5 shows agreement between differential expression and network analysis of a less significant antiviral response with estrogen treatment of IAV infected cells, including type I interferon signaling, as seen in Figure 5 (interferon alpha beta signaling). DeMAND neither quantifies up- or down-regulation nor magnitude of response, limiting the conclusions possible here with network analysis. Unlike the pathway comparison, there are no common hits when reversing the comparison by contrasting the effect of estradiol instead of contrasting the effect of infection; contrasting the effect of estradiol primarily highlights housekeeping biological processes. As with the pathway analysis, here we also consider only the 48-hour timepoint because differential expression found no significant hits at the 24-hour timepoint using FDR < 0.001. Therefore, we find at 48 hours post-infection that estradiol treatment results in a lower significance antiviral response and a lower significance type I interferon response.

3.5 Candidate transcription factors specific to epithelial cell types with few prior viral infection associations

We found eight transcription factors enriched in the mechanism of action gene set (FDR < 0.001). Their details, including their contributing genes, are summarized in Table 2. In bronchial epithelial cells, RAG1 acts in concert with X-chromosome miRNA-3202 to promote cell death by generating reactive oxygen species (ROS)[60]. TET1 is downregulated by IAV infection in A549 human lung epithelial cells[61]. DLX4 knock-down suppressed proliferation of A549 cells[62] and blocked TGF-β-based proliferation in Mv1Lu mink lung epithelial cells[63]. Little is known about the alkylation DNA repair gene ALKBH3 in epithelial cells or the lung, but it has been implicated in prostate and breast cancer[64,65]. Similarly, NRF1 (or NFE2L1) has not been investigated in the contexts of influenza infection or bronchial epithelium, but its deletion has recently been found to exacerbate oxidative stress[66,67]. Little is known about ZNF597 in humans except that overexpression has been implicated in clinical phenotypes of congenital malformations[68] and lower expression with higher risk of osteosarcoma[69]. The LIM/homeobox protein 3 (LHX3) transcription factor of the surface ectoderm, including nasal epithelial cells, has primarily been studied in the pituitary gland in humans[70], and has been found significant in animal models of avian influenza although the significance directionality changes with the animal model[71]. Not much about PSMB5 is known in this context because it was previously only known to be T-cell-specific. Collectively, most of these transcription factors uncovered by network analysis had no prior link to influenza or the bronchial epithelium, suggesting gene regulation programs missed by differential expression.

The differential expression found three transcription factors, DYRK1A, TBX3, and BANP (FDR < 0.001), that do not overlap with any of those found by MoA. DYRK1A is a multifunctional protein kinase implicated in many

disorders and, although no association has been reported between DYRK1A and estrogen, a CRISPR screen recently found that DYRK1A is required for coronavirus replication in humans[72,73]. T-box transcription factor 3 (TBX3) is a transcriptional repressor that regulates development, including development of the mammary gland, and its overexpression or mutation in mammary epithelium has been associated with estrogen-receptor-positive breast tumors and advanced stage breast cancer[74–77]. Another cell-proliferation-related transcription factor, ZNF471, is suspected to inactivate the PI3K/AKT/mTOR signaling pathway[78]. The B-cell translocation gene 3 (BTG3)-associated nuclear protein (BANP) transcription factor controls essential metabolic genes[79]. Thus, differential expression enriches for transcription factors related to infection, proliferation, and metabolism.

Table 2: Novel candidate transcription factors uncovered, including those specific to lung epithelial cells and the host response to viral infection. The candidate transcription factors were identified from estradiol treated, IAV infected vs. uninfected cells (FDR < 0.001). Transcription factors are ordered by method and FDR. A brief interpretation of the hit is summarized in the description column along with the significantly enriched genes among those provided to ToppGene responsible for enriching the candidate transcription factor. Most of these candidate transcription factors were not previously linked to influenza and some were unstudied in lung epithelial cells.

Method	Time	Name	Description (Genes)	FDR	Comment
MoA	48	PSMB5	Previously only known to be T cell specific (NPM1, FLNA, RPL28, RPL32, ADM, H2AC13, H2AC16, H2AC17, H3C1, H3C12, H3C10, SERPINE1, CDC25B, HSP90AB1) MSigDB	0.0000001	unstudied
	48	ZNF597	Overexpression implicated in clinical phenotypes of congenital malformations and lower expression with higher risk of Osteosarcoma (CACNA1D, H3C7, RARA, NELFB, NDUFB7, STAT1, GSK3B, IFIH1, H2AC13, H2AC15, H2AC16, H2AC17, SIRT2, H3C1, H3C11, H3C12, PLCG2, FYN, NOTCH1, HSP90AB1) MSigDB	0.0000022	unstudied
	48	LHX3	Found in epithelial surface ectoderm that includes nasal epithelial cells and involved in avian influenza (H2AC15, H2AC17, H3C1, H3C6, H3C12) MSigDB	0.0001516	avian flu signif, nasal epithelium
	24	RAG1	Acts with X-chromosome miRNA-3202 to promote cell death by generation of reactive oxygen species (ROS) (KRAS, ATR, CWC27, CASP3, CBX3, BIRC2, ITGAV, CUL5, RBM25, DDX46, GCLM, PIK3CB, NFE2L2, ATG12, BRD4, GLS, ROBO1, SUZ12, CDC5L) MSigDB	0.0004352	bronchial epithelium
	24	TET1	Downregulated by IAV infection in A549 human lung epithelial cells (NELFE, CUL4B, SMARCA5, TANK, WDR36, ATR, BIRC2, UPF3A, COPS2, DHX15, RBM25, PRPF3, NSA2, NFE2L2, CTR9, CENPH, CRBN) MSigDB	0.0004352	IAV signif, lung epithelium
	24	ALKBH3	Alkylation DNA repair gene implicated in prostate and breast cancer (WDR36, SRSF5, CWC27, COPS2, CCNT2, RBM25, NSA2, CTR9, FUS, TACC3, SNRPB, SEC13) MSigDB	0.0005181	unstudied
	24	NRF1	Deletion has recently been found to exacerbate oxidative stress (SMARCA5, FMR1, RBBP8, FADD, U2AF2, SRPK1, JAK2, TIA1, WDHD1, PRPF3, EIF3A, FUS, CDC5L, FTSJ3) TFBS	0.0008344	unstudied
	24	DLX4	Knock-down suppressed proliferation of A549 cells; TGF-beta based proliferation in mink lung epithelial cells (CWF19L2, WDR36, LEO1, SRSF11, NUP42, SYF2, DNTTIP2, PPIG, GTF2H1, RBM25, PIK3C3, CCDC47, NSA2, PUM3, SNRPB, FTSJ3) MSigDB	0.0008344	lung epithelium
DE	24	DYRK1A	Multifunctional protein kinase required for coronavirus replication in humans (LYRM7, GPBP1, BCLAF1, TTC14, ZNF302, SMC5, DNTTIP2, NKTR, PIBF1, WASHC4, IFT74, AHSA2P, SMC4, TEX9, BROX, PAXBP1, TIA1, MAP4K5, MOB1B, SEC62, CDC5L, SEPTIN7, SEPTIN10, SENP7) MSigDB	0.0000241	coronavirus replication
	24	TBX3	Transcription repressor that regulates development including mammary gland (LYRM7, GPBP1, CLK4, CBX3, SRSF11, SMARCAD1, NKTR, TAF1D, RBL1, FAM227B, ESF1, PAXBP1, RNF6, CEBPZ, EMC2) MSigDB	0.0000979	mammary epithelium
	24	BANP	Controls metabolic genes; suspected to inactivate PI3K/AKT/mTOR pathway (LYRM7, GPBP1, BCLAF1, TTC14, SMARCAD1, NSRP1, DNTTIP2, NKTR, PIBF1, TAF1D, IFT74, U2SURP, SMC4, TEX9, PAXBP1, SEC62, UGGT2, CDC5L, SEPTIN7, SEPTIN10, SENP7) MSigDB	0.0003034	unstudied

3.6 Fulvestrant is the only estrogen-relevant drug found by network analysis

We found an estrogen-relevant drug, fulvestrant, from a large list of drug interactions using network discovery and additional criteria as follows. There are many drug interactions associated with E2 or viral-infection response genes summarized in Figure S6. We narrowed the large list of drug interactions, which included detrimental interactions, to medicinal drug candidates by searching with drug descriptions for "estra-" or "estro-" and cross-referenced those hits with PubChem and summarize the single result in Table S3. Fulvestrant is described by DrugBank as both down-regulating and degrading estrogen receptor. According to PubChem, it has been used in many clinical trials, but none mention "viral" or "infection"; four clinical trials mention "lung". Therefore, identifying the FDA-approved drug fulvestrant by DeMAND, but not differential expression, supports the relevance of the network enrichment approach.

4 Discussion

In this work, we performed a meta-analysis using a protein-protein network informed approach to identify novel candidate proteins, pathways, and processes in cells derived from female donors that may explain how estradiol (E2) alters influenza virus replication. The original work by Peretz et al.[12] showed that treating nasal epithelial cells from female donors with E2 suppressed virus replication. Their differential expression analysis found that the E2 treatment "had no effect on interferon or chemokine secretion but significantly downregulated cell metabolic processes, including genes that encode for zinc finger proteins"[12]. Given that E2 can modify interferon-stimulated genes (ISGs)[80] and other inflammatory pathways (described in the introduction), this work was predicated on the idea that a network-informed analytical approach may identify pathways not detected by more traditional differential expression analyses. Moreover, several influenza virus host replication screens have been published that identify human proteins that influenza virus either physically interacts with or that influenza is dependent on for replication[6–11]. We therefore reanalyzed the published transcriptomic data to determine which of these host factors may be regulated by E2 and may explain the suppressed virus replication in E2-treated cells.

In general, there was little overlap between the results of differential expression and network analysis (only 5%–11% overlap; Figure 2), indicating that these two approaches capture different aspects of the host response to virus infection and estradiol treatment. Additionally, the experimental data are strongly time-dependent, because each method identified distinct results at the 24- and 48-hour timepoints. The set of transcription factors we uncovered with network analysis had no overlap with those found using differential expression (Table 2). While the original analysis of this dataset highlighted 38 zinc finger estrogen response elements[12] that are generally known to induce or repress host transcription[81], our network analysis identified a single zinc finger transcription factor ZNF597 and our differential expression analysis implicated ZNF471 by BANP, however neither of these zinc finger proteins was enriched in the original analysis. Thus, our analyses point to mechanisms beyond the zinc finger response highlighted in the original study.

Network analysis recovered known influenza–host interaction proteins and replication factors at rates well above chance (Figure 3), suggesting many possible routes by which E2 may alter virus replication (Figure 4). Of the over 1,200 host replication factors[47], network analysis identifies up to 20% of candidates per comparison. These are proteins that have been experimentally determined to suppress virus replication when knocked down, typically by siRNA screens. Thus, our top 1-percentile of network analysis hits contain a significant E2-regulated subset of host factors that inform subsequent analyses.

A good indicator that network analyses are detecting E2-relevant biology is that network analysis identified estrogen signaling pathways and suggested the drug fulvestrant. We continued our analyses of the genes identified by network analysis and differential expression analysis by completing more standard enrichment analyses, such as pathway enrichment, transcription factor site enrichment, and drug interaction enrichment. The genes identified as significant by DeMAND were enriched for estrogen-related pathway proteins at the 48-hour timepoint (Figure 5, Figure 6). Moreover, drug target enrichment analysis of the genes identified by DeMAND found an enriched signature for fulvestrant, an FDA-approved drug classified as an estrogen receptor antagonist and selective estrogen receptor degrader. Fulvestrant itself lacks therapeutic properties suitable to control virus replication[82–86], but its enrichment is consistent with DeMAND capturing E2-associated gene networks.

The pathway results of the differential expression and network analysis methods both detected type I interferon and cytokine signaling, but they diverged in the inflammatory and estrogen-related pathways they enriched (Figure 5). Differential expression uniquely enriched for IL-35 anti-inflammatory signaling whereas network analysis uniquely enriched for influenza infection, IL-7 pro-inflammatory signaling, and estrogen signaling (Figure S4). The transcription factors we identified with network analysis were either specific to bronchial epithelial cells (LHX3, RAG1, TET1, DLX4) or previously not studied in the context of viral infection (PSMB5, ALKBH3, NRF1, ZNF597). Differential expression identified DYRK1A activity at 24 hours; DYRK1A is required for coronavirus replication in humans and may play a similarly important role for IAV. Also at 24 hours, we found RPS7 using network enrichment, which

was previously identified as a COVID-19 hub gene[54] (Figure 4). Together, alongside differential expression, network analysis uncovers a rich picture of E2-associated pathways and candidate transcription regulators with cell type and viral infection relevance.

Network analysis identified estrogen-involved host factors (AGO1, ESR1, HSP90AB1) (Figure 6), of which HSP90AB1 was a top enriched host replication factor when contrasting infected with uninfected cells (Figure 4). HSP90 is known to be important for expression of viral genes[52,53,87]; thus our analysis supports prior interest in inhibiting the HSP90 host protein for IAV infection control[88,89]. Both methods found enriched histone-related genes (Figure 6) in the estrogen-dependent gene expression pathway. Consistent with previous studies, we find Rho GTPase modulation of the cytoskeleton in the context of IAV infection[55] and our analysis uncovers an effect of E2 (Figure 5). Together, HSP90 and Rho GTPases are host proteins of interest for further study of E2 influence on IAV replication.

A likely reason network analysis recovered relevant host factors is that the DeMAND algorithm uses replicate-level variation directly to detect changes between pairs of genes. In the experimental system under study, this direct measure of dysregulation significantly identified relevant host factors that co-immunoprecipitated with viral proteins, including those involved in viral replication (Figure 3). In contrast, network or pathway analysis is conventionally run after differential gene expression analysis, as in edgeR KEGG pathway analysis[90]. This conventional workflow collapses the variability between replicates into per-gene summaries before the network step, losing information about how pairs of genes behave together. It would be of interest to test the performance of DeMAND alone against a yet-to-be-developed network analysis algorithm that can make use both, replicate-level variation like DeMAND as well as differential expression analysis, to better accommodate a conventional workflow.

In all, our study suggests several new mechanisms by which E2 may regulate influenza replication. Each protein and pathway identified here requires further study to confirm its roles and to assess whether it is suitable for therapeutic development.

There are notable limitations in this study that we aim to address in the future. The data are limited to female nasal epithelial cells. Future studies should consider whether the observed hormonal effects translate across the sexes. Currently, network analysis tools lack the flexibility to specify a design matrix describing experimental sample conditions and timepoints, as differential gene expression tools allow. Like the differential gene expression tools, network analysis calculations also use three or more replicates of normalized log-fold gene expression values. We suspect that this missing capability to specify a design matrix was responsible for our network analysis of the clinical hNEC cell cultures not having a clearer separation between viral and estrogen interactions (Figure 3). The missing capability may also explain why the antiviral responses captured by DeMAND's Kullback–Leibler divergence (KLD) statistic had low conversion to mechanism of action (MoA) (see Methods), which instead identified constitutively active and housekeeping biological processes (Figure S5). For some analyses such as gene ontology, there is value in looking at the intermediate KLD measurement of gene pair dysregulation (Figure S5). Because of the inherent limitations of undirected graph-based network enrichment and the lack of fold-change output, our results are best read as hypothesis-generating candidates rather than evidence of the direction or magnitude of E2 effects.

Future work should improve network analysis tools and test our predictions with a view toward drug discovery. An alternative to our manual PubChem search would be to use AI agents, with a pharmacologist's input, to screen a larger set of compounds while minimizing false positives of compounds with known adverse reactions. CRISPR or siRNA screens of our predicted pathways and transcription factors, beginning with HSP90AB1, would test how well these E2-regulated candidates contribute to influenza A replication.

5 Data availability

Raw and processed data files are available under the Gene Expression Omnibus (GEO) accession GSE75208[12]. Accompanying code is available to review and download at https://github.com/ImmuSystems-Lab/hnec.

6 Competing interests

There are no competing interests.

7 Author Contributions

P.N.: investigation, methodology, software, visualization, writing - original draft. L.M.N.: writing - review & editing. T.S.L.: writing - review & editing. J.F.A.: funding acquisition, writing - review & editing. R.G.: funding acquisition, methodology, writing - review & editing. J.E.S.: conceptualization, funding acquisition, methodology, project administration, resources, supervision, writing - review & editing.

8 Funding

This research was supported by NIH grant R21 AI174080.

9 References

1. Iuliano AD, Roguski KM, Chang HH, Muscatello DJ, Palekar R, Tempia S, Cohen C, Gran JM, Schanzer D, Cowling BJ, Wu P, Kyncl J, Ang LW, Park M, Redlberger-Fritz M, Yu H, Espenhain L, Krishnan A, Emukule G, Van Asten L, Pereira Da Silva S, Aungkulanon S, Buchholz U, Widdowson MA, Bresee JS, Azziz-Baumgartner E, Cheng PY, Dawood F, Foppa I, Olsen S, Haber M, Jeffers C, MacIntyre CR, Newall AT, Wood JG, Kundi M, Popow-Kraupp T, Ahmed M, Rahman M, Marinho F, Sotomayor Proschle CV, Vergara Mallegas N, Luzhao F, Sa L, Barbosa-Ramírez J, Sanchez DM, Gomez LA, Vargas XB, Acosta Herrera aBetsy, Llanés MJ, Fischer TK, Krause TG, Mølbak K, Nielsen J, Trebbien R, Bruno A, Ojeda J, Ramos H, An Der Heiden M, Del Carmen Castillo Signor L, Serrano CE, Bhardwaj R, Chadha M, Narayan V, Kosen S, Bromberg M, Glatman-Freedman A, Kaufman Z, Arima Y, Oishi K, Chaves S, Nyawanda B, Al-Jarallah RA, Kuri-Morales PA, Matus CR, Corona MEJ, Burmaa A, Darmaa O, Obtel M, Cherkaoui I, Van Den Wijngaard CC, Van Der Hoek W, Baker M, Bandaranayake D, Bissielo A, Huang S, Lopez L, Newbern C, Flem E, Grøneng GM, Hauge S, De Cosío FG, De Moltó Y, Castillo LM, Cabello MA, Von Horoch M, Medina Osis J, Machado A, Nunes B, Rodrigues AP, Rodrigues E, Calomfirescu C, Lupulescu E, Popescu R, Popovici O, Bogdanovic D, Kostic M, Lazarevic K, Milosevic Z, Tiodorovic B, Chen M, Cutter J, Lee V, Lin R, Ma S, Cohen AL, Treurnicht F, Kim WJ, Delgado-Sanz C, De Mateo Ontañón S, Larrauri A, León IL, Vallejo F, Born R, Junker C, Koch D, Chuang JH, Huang WT, Kuo HW, Tsai YC, Bundhamcharoen K, Chittaganpitch M, Green HK, Pebody R, Goñi N, Chiparelli H, Brammer L, Mustaquim D. Estimates of global seasonal influenza-associated respiratory mortality: A modelling study. The Lancet. 2018 Mar;391(10127):1285–1300. PMID: 29248255

2. Tenforde MW, Noah KP, O'Halloran AC, Kirley PD, Hoover C, Alden NB, Armistead I, Meek J, Yousey-Hindes K, Openo KP, Witt LS, Monroe ML, Ryan PA, Falkowski A, Reeg L, Lynfield R, McMahon M, Hancock EB, Hoffman MR, McGuire S, Spina NL, Felsen CB, Gaitan MA, Lung K, Shiltz E, Thomas A, Schaffner W, Talbot HK, Crossland MT, Price A, Masalovich S, Adams K, Holstein R, Sundaresan D, Uyeki TM, Reed C, Bozio CH, Garg S. Timing of influenza antiviral therapy and risk of death in adults hospitalized with influenza-associated pneumonia, influenza hospitalization surveillance network (FluSurv-NET), 2012–2019. Clinical Infectious Diseases. 2025 Feb 24;80(2):461–468. PMID: 39172994

3. Tokars JI, Olsen SJ, Reed C. Seasonal incidence of symptomatic influenza in the united states. Clinical Infectious Diseases. 2018 May 2;66(10):1511–1518. PMID: 29206909

4. Klein SL, Hodgson A, Robinson DP. Mechanisms of sex disparities in influenza pathogenesis. Journal of Leukocyte Biology. 2012 Jul 1;92(1):67–73. PMID: 22131346

5. Klein SL, Pekosz A, Passaretti C, Anker M, Olukoya P. Sex, gender and influenza. World Health Organization [Internet]. 2010;58 p. Available from: https://iris.who.int/handle/10665/44401

6. Sui B, Bamba D, Weng K, Ung H, Chang S, Van Dyke J, Goldblatt M, Duan R, Kinch MS, Li WB. The use of random homozygous gene perturbation to identify novel host-oriented targets for influenza. Virology. 2009 May 10;387(2):473–481. PMCID: PMC2674145

7. Brass AL, Huang IC, Benita Y, John SP, Krishnan MN, Feeley EM, Ryan BJ, Weyer JL, Weyden L van der, Fikrig E, Adams DJ, Xavier RJ, Farzan M, Elledge SJ. The IFITM proteins mediate cellular resistance to influenza a H1N1 virus, west nile virus, and dengue virus. Cell. 2009 Dec 24;139(7):1243–1254. PMCID: PMC2824905

8. Shapira SD, Gat-Viks I, Shum BOV, Dricot A, Grace MM de, Wu L, Gupta PB, Hao T, Silver SJ, Root DE, Hill DE, Regev A, Hacohen N. A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. Cell. 2009 Dec 24;139(7):1255–1267. PMCID: PMC2892837

9. Karlas A, Machuy N, Shin Y, Pleissner KP, Artarini A, Heuer D, Becker D, Khalil H, Ogilvie LA, Hess S, Mäurer AP, Müller E, Wolff T, Rudel T, Meyer TF. Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. Nature. 2010 Feb 11;463(7282):818–822. PMID: 20081832

10. König R, Stertz S, Zhou Y, Inoue A, Hoffmann HH, Bhattacharyya S, Alamares JG, Tscherne DM, Ortigoza MB, Liang Y, Gao Q, Andrews SE, Bandyopadhyay S, De Jesus P, Tu BP, Pache L, Shih C, Orth A, Bonamy G, Miraglia L, Ideker T, García-Sastre A, Young JAT, Palese P, Shaw ML, Chanda SK. Human host factors required for influenza virus replication. Nature. 2010 Feb 11;463(7282):813–817. PMCID: PMC2862546

11. Watanabe T, Kawakami E, Shoemaker JE, Lopes TJS, Matsuoka Y, Tomita Y, Kozuka-Hata H, Gorai T, Kuwahara T, Takeda E, Nagata A, Takano R, Kiso M, Yamashita M, Sakai-Tagawa Y, Katsura H, Nonaka N, Fujii H, Fujii K, Sugita Y, Noda T, Goto H, Fukuyama S, Watanabe S, Neumann G, Oyama M, Kitano H, Kawaoka Y. Influenza virus-host interactome screen as a platform for antiviral drug development. Cell Host & Microbe. 2014 Dec;16(6):795–805. PMID: 25464832

12. Peretz J, Pekosz A, Lane AP, Klein SL. Estrogenic compounds reduce influenza a virus replication in primary human nasal epithelial cells derived from female, but not male, donors. American Journal of Physiology-Lung Cellular and Molecular Physiology. 2016 Mar 1;310(5):L415–L425. PMID: 26684252

13. Bouman A, Heineman MJ, Faas MM. Sex hormones and the immune response in humans. Human Reproduction Update. 2005;11(4):411–423. PMID: 15817524

14. Straub RH. The complex role of estrogens in inflammation. Endocrine Reviews. 2007 Aug;28(5):521–574. PMID: 17640948

15. Kovats S. Estrogen receptors regulate innate immune cells and signaling pathways. Cellular Immunology. 2015 Apr;294(2):63–69. PMID: 25682174

16. Robinson DP, Lorenzo ME, Jian W, Klein SL. Elevated 17β-estradiol protects females from influenza a virus pathogenesis by suppressing inflammatory responses. PLoS Pathog. 2011 Jul;7(7):e1002149. PMCID: PMC3145801

17. Vermillion MS, Ursin RL, Attreed SE, Klein SL. Estriol reduces pulmonary immune cell recruitment and inflammation to protect female mice from severe influenza. Endocrinology. 2018 Sep 1;159(9):3306–3320. PMCID: PMC6109301

18. Harding AT, Heaton NS. The impact of estrogens and their receptors on immunity and inflammation during infection. Cancers (Basel). 2022 Feb 12;14(4):909. PMID: PMC8870346

19. Harrington WR, Kim SH, Funk CC, Madak-Erdogan Z, Schiff R, Katzenellenbogen JA, Katzenellenbogen BS. Estrogen dendrimer conjugates that preferentially activate extranuclear, nongenomic versus genomic pathways of estrogen action. Molecular Endocrinology. 2006 Mar 1;20(3):491–502. PMID: 16306086

20. Adlanmerini M, Solinhac R, Abot A, Fabre A, Raymond-Letron I, Guihot AL, Boudou F, Sautier L, Vessières E, Kim SH, Lière P, Fontaine C, Krust A, Chambon P, Katzenellenbogen JA, Gourdy P, Shaul PW, Henrion D, Arnal JF, Lenfant F. Mutation of the palmitoylation site of estrogen receptor α in vivo reveals tissue-specific roles for membrane versus nuclear actions. Proceedings of the National Academy of Sciences. 2014 Jan 14;111(2). PMID: 24371309

21. Gilliver SC. Sex steroids as inflammatory regulators. The Journal of Steroid Biochemistry and Molecular Biology. 2010 May;120(2):105–115. PMID: 20045727

22. Klein SL, Flanagan KL. Sex differences in immune responses. Nature Reviews Immunology. 2016 Oct;16(10):626–638. PMID: 27546235

23. Lengi AJ, Phillips RA, Karpuzoglu E, Ahmed SA. Estrogen selectively regulates chemokines in murine splenocytes. Journal of Leukocyte Biology. 2007 Apr 1;81(4):1065–1074. PMID: 17185357

24. Luo J, Liu D. Does GPER really function as a g protein-coupled estrogen receptor in vivo? Frontiers in Endocrinology. 2020 Mar 31;11:148. PMID: 32296387

25. Hoffmann JP, Liu JA, Seddu K, Klein SL. Sex hormone signaling and regulation of immune function. Immunity. 2023 Nov;56(11):2472–2491. PMID: 37967530

26. Berghöfer B, Frommer T, Haley G, Fink L, Bein G, Hackstein H. TLR7 ligands induce higher IFN-α production in females. Journal of Immunology. 2006;177(4):2088–2096. PMID: 16887967

27. Koyama S, Ishii KJ, Coban C, Akira S. Innate immune response to viral infection. Cytokine. 2008 Sep;43(3):336–341. PMID: 18694646

28. Pujantell M, Skenteris NT, Claussen JM, Grünhagel B, Thiele RJ, Altfeld M. Sex-dependent differences in type i IFN-induced natural killer cell activation. Frontiers in Immunology. 2023;14:1277967. PMCID: PMC10757368

29. Abdullah M, Chai PS, Chong MY, Tohit ERM, Ramasamy R, Pei CP, Vidyadaran S. Gender effect on in vitro lymphocyte subset levels of healthy individuals. Cellular Immunology. 2012;272(2):214–219. PMID: 22078320

30. Gupta S, Nakabo S, Blanco LP, O'Neil LJ, Wigerblad G, Goel RR, Mistry P, Jiang K, Carmona-Rivera C, Chan DW, Wang X, Pedersen HL, Gadkari M, Howe KN, Naz F, Dell'Orso S, Hasni SA, Dempsey C, Buscetta A, Frischmeyer-Guerrerio PA, Kruszka P, Muenke M, Franco LM, Sun HW, Kaplan MJ. Sex differences in neutrophil biology modulate response to type i interferons and immunometabolism. Proceedings of the National Academy of Sciences. 2020 Jul 14;117(28):16481–16491. PMCID: PMC7368314

31. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: A network-based approach to human disease. Nature Reviews Genetics. 2011 Jan;12(1):56–68. PMID: 21164525

32. Perrin-Cocon L, Diaz O, Jacquemin C, Barthel V, Ogire E, Ramière C, André P, Lotteau V, Vidalain PO. The current landscape of coronavirus-host protein–protein interactions. Journal of Translational Medicine. 2020 Dec;18(1):319. PMID: 32811513

33. Rask-Andersen M, Almén MS, Schiöth HB. Trends in the exploitation of novel drug targets. Nature Reviews Drug Discovery. 2011 Aug;10(8):579–590. PMID: 21804595

34. Kotlyar M, Fortney K, Jurisica I. Network-based characterization of drug-regulated genes, drug targets, and toxicity. Methods. 2012 Aug;57(4):499–507. PMID: 22749929

35. Shoemaker JE, Fukuyama S, Eisfeld AJ, Muramoto Y, Watanabe S, Watanabe T, Matsuoka Y, Kitano H, Kawaoka Y. Integrated network analysis reveals a novel role for the cell cycle in 2009 pandemic influenza virus-induced inflammation in macaque lungs. BMC Systems Biology. 2012 Dec;6(1):117. PMID: 22937776

36. Nguyen TM, Shafi A, Nguyen T, Draghici S. Identifying significantly impacted pathways: A comprehensive review and assessment. Genome Biology. 2019 Dec;20(1):203. PMID: 31597578

37. Badkas A, Pacheco MP, Sauter T. Network modeling approaches for metabolic diseases and diabetes. Current Opinion in Systems Biology. 2024 Dec;39:100530.

38. Woo JH, Shimoni Y, Yang WS, Subramaniam P, Iyer A, Nicoletti P, Rodríguez Martínez M, López G, Mattioli M, Realubit R, Karan C, Stockwell BR, Bansal M, Califano A. Elucidating compound mechanism of action by network perturbation analysis. Cell. 2015 Jul;162(2):441–451. PMID: 26186195

39. Klipp E, Liebermeister W. Mathematical modeling of intracellular signaling pathways. BMC Neuroscience. 2006 Oct;7:S10. PMID: 17118154

40. Aldridge BB, Burke JM, Lauffenburger DA, Sorger PK. Physicochemical modelling of cell signalling pathways. Nature Cell Biology. 2006 Nov;8(11):1195–1203. PMID: 17060902

41. Wang M, Luciani LL, Noh H, Mochan E, Shoemaker JE. TREAP: A new topological approach to drug target inference. Biophysical Journal. 2020 Dec;119(11):2290–2298. PMID: 33129831

399 42. Noh H, Shoemaker JE, Gunawan R. Network perturbation analysis of gene transcriptional profiles reveals protein targets and mechanism of action of drugs and influenza a viral infection. Nucleic Acids Research. 2018 Apr 6;46(6):e34–e34. PMID: 29325153

400 43. Ackerman EE, Shoemaker JE. Network controllability-based prioritization of candidates for SARS-CoV-2 drug repositioning. Viruses. 2020 Sep 26;12(10):1087. PMCID: PMC7650805

401 44. MacDonald JW. hugene20sttranscriptcluster.db [Internet]. Bioconductor; 2017. Available from: https://bioconductor.org/packages/hugene20sttranscriptcluster.db

402 45. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research. 2015;43(7):e47. PMID: 25605792

403 46. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, Bork P, Jensen LJ, Mering C von. The STRING database in 2023: Protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Research. 2023 Jan 6;51:D638–D646. PMID: 36370105

404 47. Ackerman EE, Kawakami E, Katoh M, Watanabe T, Watanabe S, Tomita Y, Lopes TJ, Matsuoka Y, Kitano H, Shoemaker JE, Kawaoka Y. Network-guided discovery of influenza virus replication host factors. Bloom JD, Harwood CS, editors. mBio. 2018 Dec 21;9(6):e02002–18. PMID: 30563907

405 48. Chen J, Xu H, Aronow BJ, Jegga AG. Improved human disease candidate gene prioritization using mouse phenotype. BMC Bioinformatics. 2007 Dec;8(1):392. PMID: 17939863

406 49. Nanda P, Shoemaker JE. Toppgene: Gene list enrichment analysis using the ToppGene suite [Internet]. Bioconductor 3.23; 2026. Available from: https://bioconductor.org/packages/toppgene/

407 50. Kaimal V, Bardes EE, Tabar SC, Jegga AG, Aronow BJ. ToppCluster: A multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems. Nucleic Acids Research. 2010 Jul 1;38:W96–W102. PMID: 20484371

408 51. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J, Bolton EE. PubChem 2025 update. Nucleic Acids Research. 2025 Jan 6;53:D1516–D1525. PMID: 39558165

409 52. Zhao Y, Yuan J, Xiao D, Zhang L, Li C, Hu J, Chen R, Song D, Wen Y, Wu R, Zhao Q, Du S, Yan Q, Han X, Wen X, Cao S, Huang X. HSP90AB1 is a host factor that promotes porcine deltacoronavirus replication. The Journal of Biological Chemistry. 2024 Jan;300(1):105536. PMCID: PMC10789647

410 53. Oh KK, Adnan M, Cho DH. Drug-repurposing against COVID-19 by targeting a key signaling pathway: An in silico study. Medical Hypotheses. 2021 Oct;155:110656. PMCID: PMC8349734

411 54. Qi P, Huang M, Li T. Screening the potential biomarkers of COVID-19-related thrombosis through bioinformatics analysis. Front Genet. 2022;13:889348. PMCID: PMC9174658

412 55. Bedi S, Ono A. Friend or foe: The role of the cytoskeleton in influenza a virus assembly. Viruses. 2019 Jan 10;11(1):46. PMCID: PMC6356976

413 56. Picard D. Chaperoning steroid hormone action. Trends in endocrinology and metabolism: TEM. 2006 Aug;17(6):229–235. PMID: 16806964

414 57. Huang H, Qiao Y, Chu L, Ye C, Lin L, Liao H, Meng X, Zou F, Zhao H, Zou M, Cai S, Dong H. Up-regulation of HSP90α in HDM-induced asthma causes pyroptosis of airway epithelial cells by activating the cGAS-STING-ER stress pathway. International Immunopharmacology. 2024 Apr 20;131:111917. PMID: 38527402

58. Gómez Acuña LI, Nazer E, Rodríguez-Seguí SA, Pozzi B, Buggiano V, Marasco LE, Agirre E, He C, Alló M, Kornblihtt AR. Nuclear role for human argonaute-1 as an estrogen-dependent transcription coactivator. The Journal of Cell Biology. 2020 Sep 7;219(9):e201908097. PMID: 32673398

59. Yazdi PT, Wang Y, Zhao S, Patel N, Lee EY-HP, Qin J. SMC1 is a downstream effector in the ATM/NBS1 branch of the human s-phase checkpoint. Genes & Development. 2002 Mar 1;16(5):571–582. PMID: 11877377

60. Zeng L, Zhang S, Fu N, Gao F, Ren N, Zheng W, Lin B, Chen H. miR-3202 inhibits bronchopulmonary dysplasia-mediated apoptosis and oxidative stress in bronchial epithelial cells via targeting RAG1. Pathology - Research and Practice. 2024 Sep;261:155482. PMID: 39067173

61. Bery DE, El-Masry SA, Guirgis AA, Zain AM, Khalil H. Epigallocatechin-3-gallate inhibits replication of influenza a virus via restoring the host methylated genes following infection. International Microbiology. 2025 Apr 15; PMID: 40232535

62. Wan Z, Wang J, Liu Q, Yang D, Li P, Wang L. Knockdown of DLX4 inhibits non-small cell lung cancer tumor growth by downregulating CKS2. Open Life Sciences. 2023 Sep 19;18(1):20220720. PMID: 37744456

63. Trinh BQ, Barengo N, Naora H. Homeodomain protein DLX4 counteracts key transcriptional control mechanisms of the TGF-β cytostatic program and blocks the antiproliferative effect of TGF-β. Oncogene. 2011 Jun 16;30(24):2718–2729. PMID: 21297662

64. Koike K, Ueda Y, Hase H, Kitae K, Fusamae Y, Masai S, Inagaki T, Saigo Y, Hirasawa S, Nakajima K, Ohshio I, Makino Y, Konishi N, Yamamoto H, Tsujikawa K. Anti-tumor effect of AlkB homolog 3 knockdown in hormone- independent prostate cancer cells. Current Cancer Drug Targets. 2012 Aug 1;12(7):847–856. PMID: 22515525

65. Stefansson OA, Hermanowicz S, Van Der Horst J, Hilmarsdottir H, Staszczak Z, Jonasson JG, Tryggvadottir L, Gudjonsson T, Sigurdsson S. CpG promoter methylation of the ALKBH3 alkylation repair gene in breast cancer. BMC Cancer. 2017 Dec;17(1):469. PMID: 28679371

66. Hirakawa T, Taniuchi M, Iguchi Y, Bogahawaththa S, Yoshitake K, Werellagama S, Uemura T, Tsujita T. NF-E2-related factor 1 suppresses the expression of a spermine oxidase and the production of highly reactive acrolein. Scientific Reports. 2025 Apr 21;15(1):12405. PMID: 40258928

67. Lyu D, Wang M, Qiu L, Deng R, Hu S, Zhang Y. Deletion of Nrf1α exacerbates oxidative stress-induced cellular senescence by disrupting cell homeostasis. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 2025 Apr;119970. PMID: 40280334

68. Schulze KV, Szafranski P, Lesmana H, Hopkin RJ, Hamvas A, Wambach JA, Shinawi M, Zapata G, Carvalho CMB, Liu Q, Karolak JA, Lupski JR, Hanchard NA, Stankiewicz P. Novel parent-of-origin-specific differentially methylated loci on chromosome 16. Clinical Epigenetics. 2019 Dec;11(1):60. PMID: 30961659

69. Sun X, Zheng D, Guo W. Comprehensive analysis of a zinc finger protein gene–based signature with regard to prognosis and tumor immune microenvironment in osteosarcoma. Frontiers in Genetics. 2022 Feb 25;13:835014. PMID: 35281811

70. Branch MC, Weber M, Li MY, Flora P, Ezhkova E. Overview of chromatin regulatory processes during surface ectodermal development and homeostasis. Developmental Biology. 2024 Nov;515:30–45. PMCID: PMC11317222

71. Klees S, Schlüter JS, Schellhorn J, Bertram H, Kurzweg AC, Ramzan F, Schmitt AO, Gültas M. Comparative investigation of gene regulatory processes underlying avian influenza viruses in chicken and duck. Biology. 2022 Jan 29;11(2):219. PMID: 35205087

72. Yang Y, Fan X, Liu Y, Ye D, Liu C, Yang H, Su Z, Zhang Y, Liu Y. Function and inhibition of DYRK1A: Emerging roles of treating multiple human diseases. Biochemical Pharmacology. 2023 Jun;212:115521. PMID: 36990324

73. Fu Z, Xiang Y, Fu Y, Su Z, Tan Y, Yang M, Yan Y, Baghaei Daemi H, Shi Y, Xie S, Sun L, Peng G. DYRK1A is a multifunctional host factor that regulates coronavirus replication in a kinase-independent manner. Journal of Virology. 2024 Jan 23;98(1):e0123923. PMID: 38099687

74. Douglas NC, Papaioannou VE. The t-box transcription factors TBX2 and TBX3 in mammary gland development and breast cancer. Journal of Mammary Gland Biology and Neoplasia. 2013 Jun;18(2):143–147. PMID: 23624936

75. Arendt LM, St Laurent J, Wronski A, Caballero S, Lyle SR, Naber SP, Kuperwasser C. Human breast progenitor cell numbers are regulated by WNT and TBX3. PloS One. 2014;9(10):e111442. PMID: 25350852

76. Kunasegaran K, Ho V, Chang TH-T, De Silva D, Bakker ML, Christoffels VM, Pietersen AM. Transcriptional repressor Tbx3 is required for the hormone-sensing cell lineage in mammary epithelium. PloS One. 2014;9(10):e110191. PMID: 25343378

77. Aliwaini S, Lubbad AM, Shourfa A, Hamada HAA, Ayesh B, Abu Tayem HEM, Abu Mustafa A, Abu Rouk F, Redwan MM, Al-Najjar M. Overexpression of TBX3 transcription factor as a potential diagnostic marker for breast cancer. Molecular and Clinical Oncology. 2019 Jan;10(1):105–112. PMID: 30655984

78. Wang X, Yao L, Li Z, Zhang J, Ruan M, Mulati Y, Gan Y, Zhang Q. ZNF471 interacts with BANP to reduce tumour malignancy by inactivating PI3K/AKT/mTOR signalling but is frequently silenced by aberrant promoter methylation in renal cell carcinoma. International Journal of Biological Sciences. 2024;20(2):643–663. PMID: 38169650

79. Grand RS, Burger L, Gräwe C, Michael AK, Isbel L, Hess D, Hoerner L, Iesmantavicius V, Durdu S, Pregnolato M, Krebs AR, Smallwood SA, Thomä N, Vermeulen M, Schübeler D. BANP opens chromatin and activates CpG-island-regulated genes. Nature. 2021 Aug;596(7870):133–137. PMID: 34234345

80. Kanda N, Watanabe S. 17beta-estradiol inhibits the production of interferon-induced protein of 10 kDa by human keratinocytes. J Invest Dermatol. 2003 Mar;120(3):411–419.

81. Laity JH, Lee BM, Wright PE. Zinc finger proteins: New insights into structural and functional diversity. Current opinion in structural biology. 2001 Feb;11(1):39–46. PMID: 11179890

82. Curran M, Wiseman L. Fulvestrant. Drugs. 2001;61(6):807–813; discussion 814. PMID: 11398912

83. Fulvestrant. Drugs and lactation database (LactMed). Bethesda (MD): National Institute of Child Health and Human Development; 2006. PMID: 40127363

84. Di Leo A, Jerusalem G, Petruzelka L, Torres R, Bondarenko IN, Khasanov R, Verhoeven D, Pedrini JL, Smirnova I, Lichinitser MR, Pendergrass K, Garnett S, Lindemann JPO, Sapunar F, Martin M. Results of the CONFIRM phase III trial comparing fulvestrant 250 mg with fulvestrant 500 mg in postmenopausal women with estrogen receptor-positive advanced breast cancer. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology. 2010 Oct 20;28(30):4594–4600. PMID: 20855825

85. Tancredi R, Furlanetto J, Loibl S. Endocrine therapy in premenopausal hormone receptor positive/human epidermal growth receptor 2 negative metastatic breast cancer: Between guidelines and literature. The Oncologist. 2018 Aug;23(8):974–981. PMID: 29934412

86. Sammons S, Kornblum NS, Blackwell KL. Fulvestrant-based combination therapy for second-line treatment of hormone receptor-positive advanced breast cancer. Targeted Oncology. 2019 Feb;14(1):1–12. PMID: 30136059

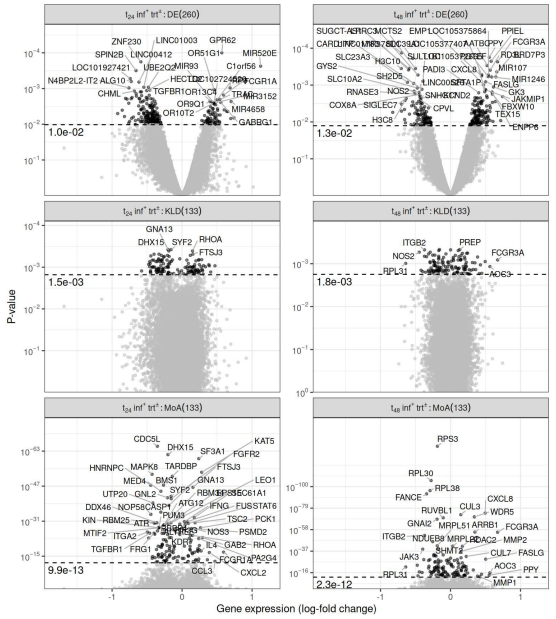
87. Lubkowska A, Pluta W, Strońska A, Lalko A. Role of heat shock proteins (HSP70 and HSP90) in viral infection. International Journal of Molecular Sciences. 2021 Aug 29;22(17):9366. PMID: 34502274

88. Kumar P, Gaur P, Kumari R, Lal SK. Influenza a virus neuraminidase protein interacts with Hsp90, to stabilize itself and enhance cell survival. Journal of Cellular Biochemistry. 2019 Apr;120(4):6449–6458. PMID: 30335904

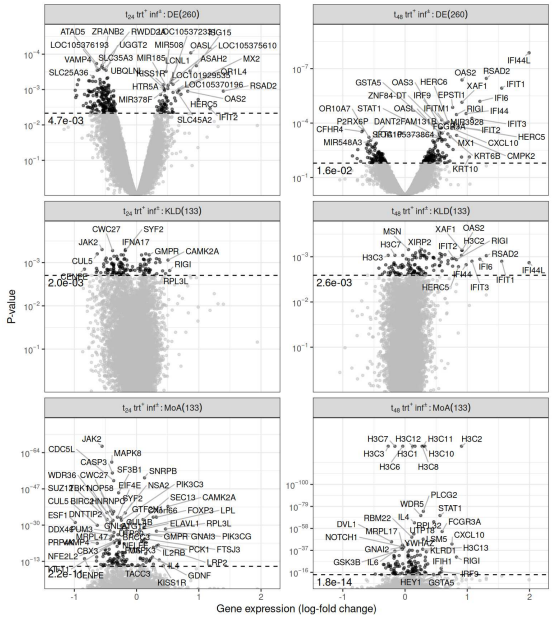
446 89. Wan Q, Song D, Li H, He M. Stress proteins: The biological functions in virus infection, present and challenges for target-based antiviral drug development. Signal Transduction and Targeted Therapy. 2020 Jul 13;5:125. PMID: 32661235

447 90. Chen Y, Chen L, Lun ATL, Baldoni PL, Smyth GK. edgeR v4: Powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. Nucleic Acids Research. 2025 Jan 11;53(2):gkaf018. PMID: 39844453

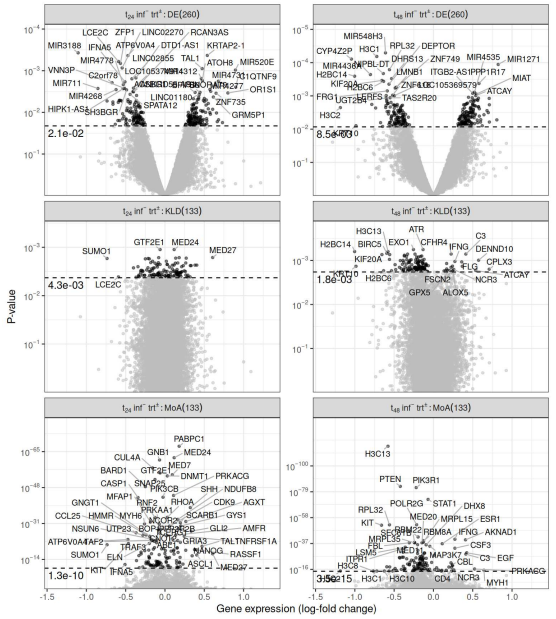
Supplementary Information



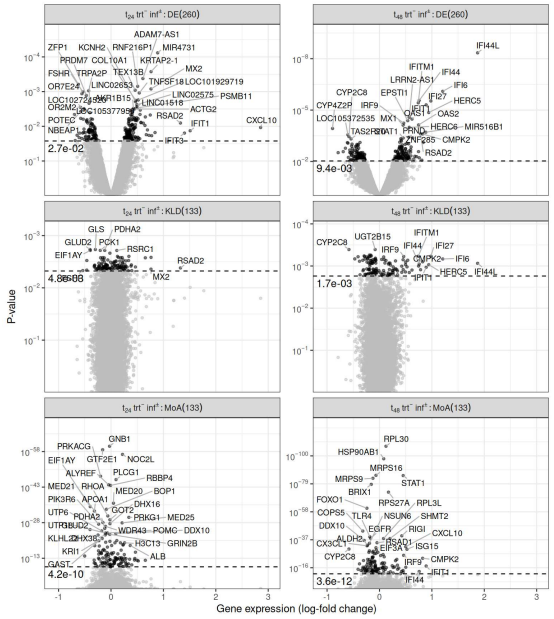
(a) trt^{-}inf$^{\pm}$: Untrt, Inf (virus) vs. Uninf



(b) trt^{+}inf$^{\pm}$: Trt, Inf (virus) vs. Uninf

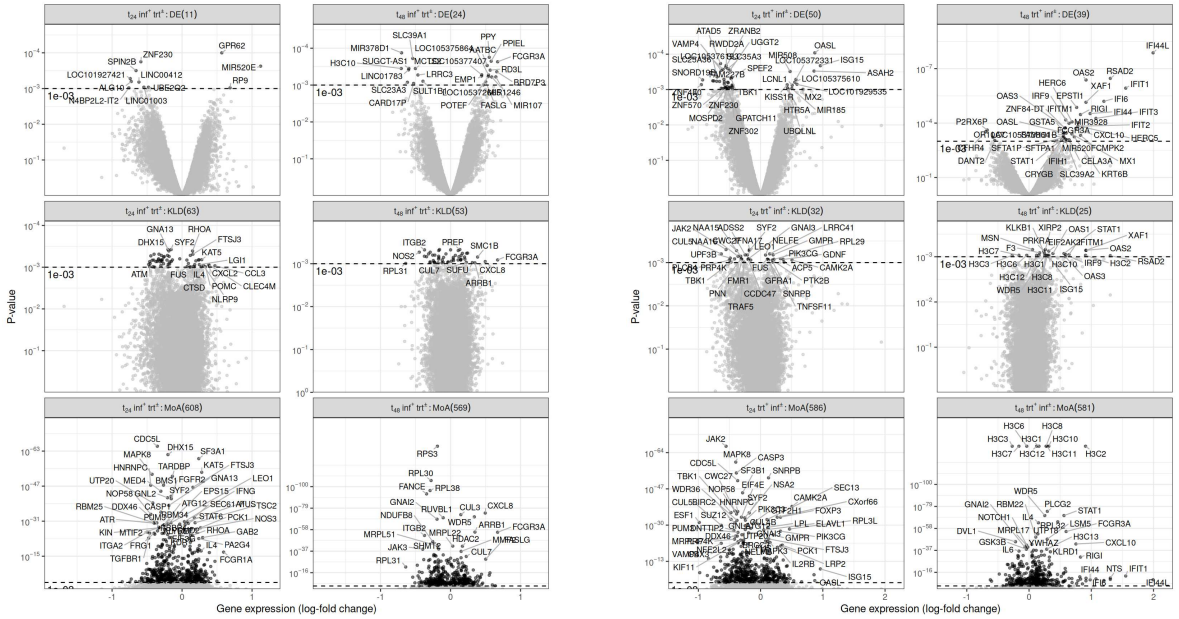


(c) inf^{-}trt$^{\pm}$: Uninf, Trt (estradiol) vs. Untrt



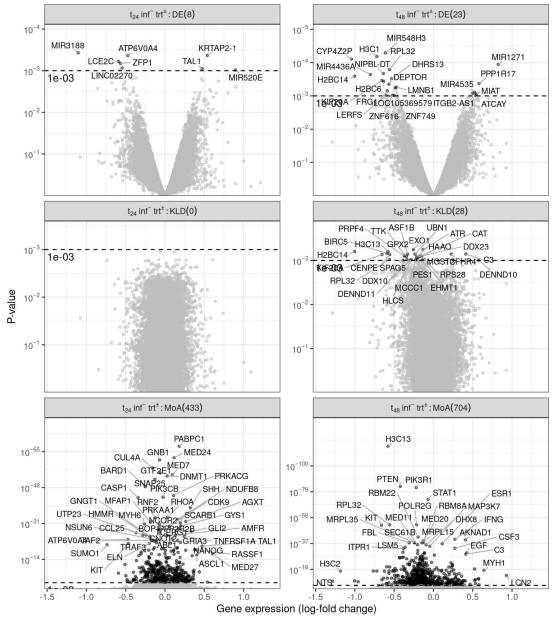
(d) inf^{+}trt$^{\pm}$: Inf, Trt (estradiol) vs. Untrt

Figure S1: Volcano plots comparing significance of expressed genes found by the three methods at the 24- and 48-hour timepoints for each condition. Each condition was subset and independently enriched—using the comparisons indicated in the sub-captions (a–d)—with differential expression, network-based raw Kullback–Leibler divergence (KLD) gene pairs, and mechanism of action (MoA) gene sets. Number of significant genes (n) shown in the sub-plot titles were chosen as the top 1 percentile, and the corresponding p-value is shown by the number below the dashed line. Significant genes meeting this criterion are colored as black dots and labeled where possible, whereas non-significant genes are colored as gray dots. Note that the DeMAND methods (KLD, MoA) only calculate significance (y-axis), therefore the plots use the same log-fold change calculated by DE (x-axis).

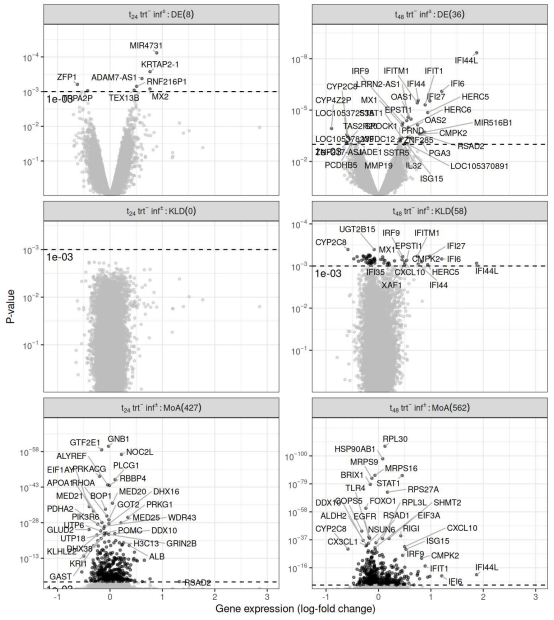


(a) $\text{trt}^{-}\text{inf}^{\pm}$: Untrt, Inf (virus) vs. Uninf

(b) $\text{trt}^{+}\text{inf}^{\pm}$: Trt, Inf (virus) vs. Uninf



(c) $\text{inf}^{-}\text{trt}^{\pm}$: Uninf, Trt (estradiol) vs. Untrt



(d) $\text{inf}^{+}\text{trt}^{\pm}$: Inf, Trt (estradiol) vs. Untrt

Figure S2: Recreation of Figure S1 using a fixed p-value < 0.001 threshold showing 0–704 significant genes. Due to significance values for the three methods (DE, KLD, MoA) being generated by different computations with their associated different distributions, we opted to use the ranking-based 1 percentile threshold of Figure S1 instead of a fixed threshold.

Table S1: Individual enriched genes across methods found among known host factors. Known host factors consist of pooled genes from Watanabe et al. 2014 and Ackerman et al. 2018. A "Count" of three indicates the known host factor is present in all three sets of known host factors (Watanabe interaction and replication, and Ackerman replication; see Figure 3).

`See separate file attachment "enriched-host-factors.csv"`

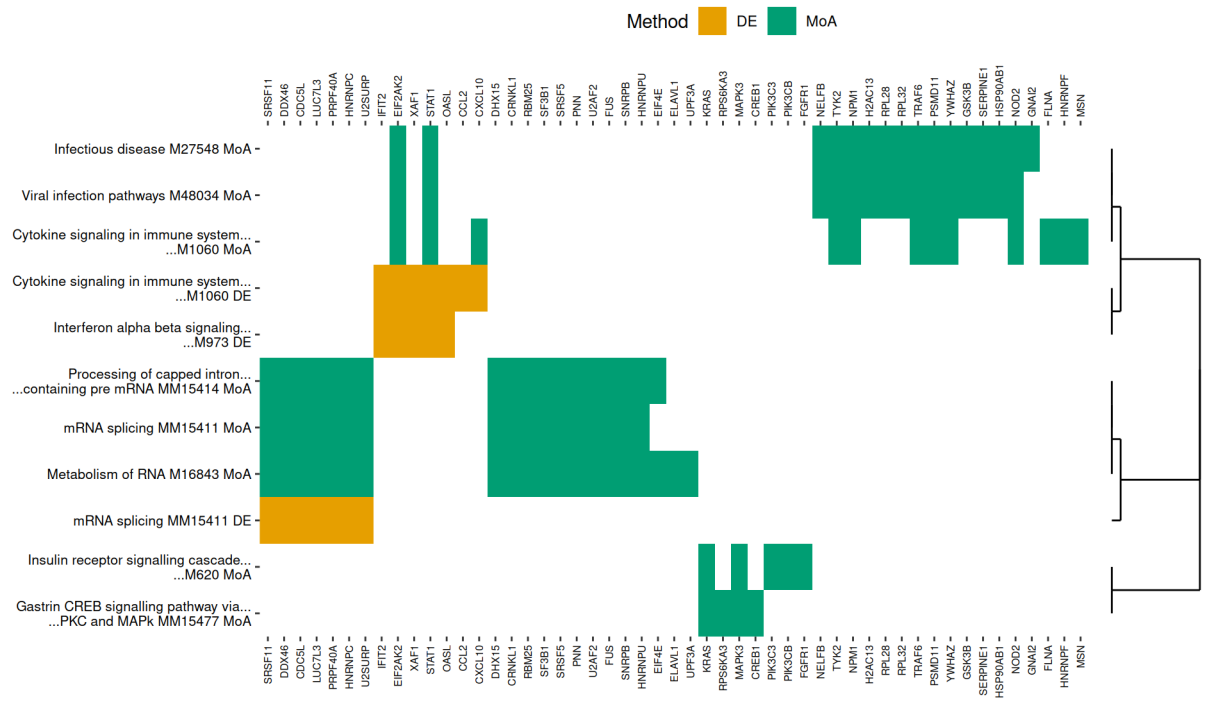


Figure S3: Individual known host factors in pathways after pooling timepoints for E2-treated, infected vs. uninfected condition. Rows are Reactome pathways and columns are method-identified genes present in the known host factor set. Filled boxes are genes enriched by the methods. There are very few enriched genes from the differential expression method present among the known host factors, therefore the enriched genes were pooled from all timepoints.



Figure S4: Among the top-20 hits at 48 hours, only DeMAND finds relevant estrogen-related pathways, but both methods detect type I interferon and cytokine signaling. The Y-axis order follows ranks of the significant genes (FDR < 0.001) of the top 20 Reactome pathways identified by each of the methods at the 48-hour timepoint. Pathways of particular interest to the experimental system are indicated with bold text: these include host immune pathways, cytoskeletal modulation by IAV infection, and estrogen-related pathways. Unlike the gene ontology analysis in Figure S5, where there was little agreement among the top hits, in this pathway analysis there is generally more agreement between the two methods.



Figure S5: Gene ontology analysis reveals lower significance antiviral responses with estrogen treatment for both DE and DeMAND. For gene ontology enrichment, the KLD measure of gene-gene pair network dysregulation shows higher sensitivity. The figure shows significant genes (FDR < 0.001) of the top 10 gene ontology biological processes identified by each of the methods at the 48-hour timepoint; DE had no significant genes at the 24-hour timepoint.

Table S2: Excerpt of ToppGene results table of gene set enrichment. The results show the single top result for each of the categories returned by the ToppGene web API. Results include the queried genes in the enrichment identifier as well as the provenance in the database. Columns not shown include the numbers of genes participating in set comparisons and alternative p/q-values other than the false discovery rate (FDR). In this work, we use ToppGene to query the top 100 results for each category for all experimental conditions (Table 1) and methods (DE, KLD, MoA).

Category	ID, Name	Source, URL	FDRBH	GenesSymbol
Coexpression	M11961 PUJANA_CHEK2_...	C2:CGP: Chemi... gsea-msigdb.o...	9.50e-15	MRE11, SFPQ, MTF2, RBBP4, CPSF6, TO...
CoexpressionA...	facebase_RNAs... facebase_RNAs...	Gudmap RNAseq gudmap.org	3.98e-06	KIF18A, CWF19L2, SFPQ, SFRP1, PTBP2...
Computational	GNF2_XRCC5 Neighborhood ...	C4:CGN: Cance... gsea-msigdb.o...	1.27e-05	SFPQ, CPSF6, PSMC2, HNRNPC, DHX15, ...
Cytoband	6p22.1 6p22.1		1.63e-06	H4C1, H4C4, H4C6, H4C3, H4C8, H4C2,...
Disease	DOID:10652 (i... Alzheimer's d...	AllianceGenome fms.allianceg...	3.88e-06	MRE11, BRIP1, CASP1, TSC2, TARDBP, ...
Domain	IPR000504 RRM_dom	InterPro ebi.ac.uk/int...	4.29e-07	HNRNPDL, SFPQ, PTBP2, CPSF6, SNRNP3...
Drug	ctd:D009538 Nicotine	CTD ctd.mdibl.org...	2.29e-06	SFPQ, MTF2, RBBP4, PTBP2, TOPBP1, C...
GeneFamily	725 RNA binding m...	genenames.org genenames.org...	4.14e-09	HNRNPDL, SFPQ, PTBP2, CPSF6, SNRNP3...
GOBiologicalP...	GO:0008380 RNA splicing		9.37e-14	CWF19L2, SFPQ, PTBP2, SF3A1, SNRNP3...
GOCellularCom...	GO:0140513 nuclear prote...		3.24e-17	CARM1, HNRNPDL, MRE11, CWF19L2, SFP...
GOMolecularFu...	GO:0001221 transcription...		3.49e-04	MTF2, CBX3, STAT6, PHF1, ATXN3L, UB...
HumanPheno	HP:0002366 Abnormal lowe...		1.27e-02	GRN, TARDBP, TBK1, ATXN3, FUS
Interaction	int:EP300 EP300 interac...		2.18e-15	CARM1, HNRNPDL, MRE11, RPL34, SFPQ,...
MicroRNA	mmu-miR-669b-3p	MirDB mirdb.org/cgi...	1.32e-10	HNRNPDL, STAM2, KIF23, MTF2, PTBP2,...
MousePheno	MP:0011975 neuronal cyto...		1.29e-03	GRN, ATXN3L, TARDBP, ATXN3
Pathway	M14033 REAC-TOME_MRNA...	Reactome Path... gsea-msigdb.o...	6.86e-10	CWF19L2, SF3A1, SNRNP35, SYF2, HNRN...
Pubmed	24711643 Identifying b...	Pubmed pubmed.org/24...	1.16e-24	HNRNPDL, MRE11, SFPQ, HYOU1, RBBP4,...
TFBS	M30034 KAT5_TARGET_G...	MSigDB gtrd.biouml.org/	1.28e-03	SNF8, RPL34, SFPQ, KIF23, MTF2, BMS...
ToppCell	f20b0d71f857a... Striatum-Macr...	Mouse Adult B... toppcell.cchm...	3.03e-06	SFPQ, KIF23, BRIP1, ECT2, HNRNPC, P...

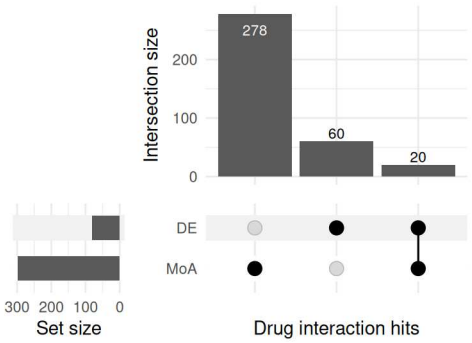


Figure S6: DeMAND finds many more candidate drugs. Drug candidates (FDR < 0.001) found by differential expression and network analysis enrichment methods that are associated with estradiol (E2) or virus conditions. Network enrichment (MoA) predicts many more drug candidates (n = 278) than differential expression (DE) alone (n = 60) in the ToppGene databases with 20 drug predictions common to both methods.

Table S3: Estradiol-related medicinal drug interactions at 48 hours with PubChem support. The single hit is from estradiol-treated, infected vs. uninfected samples enriched by the mechanism of action network analysis method. Drug names are listed with their PubChem compound identifiers (CIDs). The cut-off for significant drugs queried from ToppGene (FDR < 0.001) was applied to exclude biologically irrelevant hits of 17alpha-estradiol, diethylstilbestrol, ethylnylestradiol and Phytoestrogens.

Drug (PubChem)	FDR	Genes in query relevant to the ToppGene drug hit
Fulvestrant (CID 104741)	3.14e-07	NPM1, PCLAF, CXCL10, SRC, POMC, GRIN2B, RET, AR, STAT1, GSK3B, BMP2, EGF, CSF3, GNAQ, LTF, CDC25B, CXCL11, IL6, HSP90AB1

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- enrichedhostfactors.csv