OBIF: an omics-based interaction framework to reveal molecular drivers of synergy

Jezreel Pantaleón García [ORCID][1], **Vikram V. Kulkarni**[1,2], **Tanner C. Reese**[1,3], **Shradha Wali**[1,2], **Saima J. Wase**[1], **Jiexin Zhang**[4], **Ratnakar Singh**[5,6], **Mauricio S. Caetano**[1], **Humam Kadara**[7], **Seyed Javad Moghaddam** [ORCID][1,2], **Faye M. Johnson**[5], **Jing Wang**[4], **Yongxing Wang**[1] and **Scott E. Evans**[1,2,*]

[1]Department of Pulmonary Medicine, University of Texas MD Anderson Cancer Center, HoustonTX 77030, USA, [2]MD Anderson Cancer Center UT Health Graduate School of Biomedical Sciences, Houston, TX 77030, USA, [3]Rice University, Houston, TX 77005, USA, [4]Department of Bioinformatics and Computational Biology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA, [5]Department of Thoracic, Head and Neck Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA, [6]Department of Comparative Biosciences, University of Illinois at Urbana-Champaign, Urbana, IL 61802, USA and [7]Department of Translational Molecular Pathology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

ABSTRACT

Bioactive molecule library screening may empirically identify effective combination therapies, but molecular mechanisms underlying favorable drug–drug interactions often remain unclear, precluding further rational design. In the absence of an accepted systems theory to interrogate synergistic responses, we introduce Omics-Based Interaction Framework (OBIF) to reveal molecular drivers of synergy through integration of statistical and biological interactions in synergistic biological responses. OBIF performs full factorial analysis of feature expression data from single versus dual exposures to identify molecular clusters that reveal synergy-mediating pathways, functions and regulators. As a practical demonstration, OBIF analyzed transcriptomic and proteomic data of a dyad of immunostimulatory molecules that induces synergistic protection against influenza A and revealed unanticipated NF-κB/AP-1 cooperation that is required for antiviral protection. To demonstrate generalizability, OBIF analyzed data from a diverse array of Omics platforms and experimental conditions, successfully identifying the molecular clusters driving their synergistic responses. Hence, unlike existing synergy quantification and prediction methods, OBIF is a phenotype-driven systems model that supports multiplatform interrogation of synergy mechanisms.

INTRODUCTION

Superior treatment outcomes are achieved for many disease states when more than one therapeutic agent is administered (1–4). Indeed, there are many well documented instances when the therapeutic benefit of two agents administered together substantially exceeds the benefit that would be predicted by the additive effects of the agents administered individually. These clinical observations parallel the many instances of synergy observed in nature (1–5). Widespread availability of high throughput technologies has allowed multi-level study of complex biological responses from genome to phenome (6–8). Yet, there remains lack of consensus regarding the appropriate analysis of statistical and biological interactions found in antagonistic or synergistic responses (5,9,10). Moreover, previously proposed strategies to analyze these interactions frequently lack sufficient generalizability to study these processes outside of their home Omics platforms (1,10–14). Thus, while synergistic therapeutic combinations may be empirically derived from fortuitous clinical experiences or through screening of bioactive small molecule libraries, the absence of established means to investigate these favorable drug–drug interactions ultimately precludes understanding of their underlying mechanisms. Similarly, quantification of biologic synergy or antagonistic interactions can be accomplished with existing tools, but the mechanisms of their interactions will remain elusive. Consequently, development of a methodology to integrate the statistical and biological components of synergistic interactions in diverse Omics settings can advance the rational design of combination therapies

*To whom correspondence should be addressed. Tel: +1 713 563 7433; Email: seevans@mdanderson.org

while affording deeper understanding of molecular interactions in a broad range of biological contexts.

Pneumonia is a major worldwide cause of death (15–18). Further, the ongoing COVID-19 pandemic highlights the need to develop novel anti-pneumonia interventions with activity against both established and emerging pathogens (19–21). We have previously reported that a therapeutic dyad of immunostimulatory small molecules induces robust protection against a broad range of pneumonia-causing pathogens (22–25), including coronaviruses (26), while the individual components confer only modest protection. This combination (hereafter, 'Pam2-ODN') Comprises a Toll-like receptor (TLR) 2/6 agonist, Pam2CSK4 ('Pam2'), and a TLR 9 agonist, ODN M362 ('ODN'), which stimulate protective responses from lung epithelial cells (22). This biological response, termed inducible epithelial resistance, promotes survival benefits and microbicidal effects that significantly exceed the additive effects of the individual ligands (23,27). Thus, understanding the molecular mechanisms underlying this unanticipated synergy may allow optimized manipulation of epithelial antimicrobial responses and support new generations of host-based therapeutics against infections.

In the absence of a systems theory to interrogate synergistic mechanisms (5), we introduce Omics-Based Interaction Framework (OBIF) to identify molecular drivers through integration of statistical and biological interactions in synergistic biological responses. Unlike existing statistical models that are principally designed to either predict synergistic drug combinations (11–14,28–33) or to quantify synergy between drugs (1,34–36), OBIF is a phenotype-driven (6) synergy interrogation model that is applicable to any combination of exposure types (e.g. drugs, genotypes, environmental conditions, etc.). To explain the molecular mechanisms underlying empirically observed instances of synergy, OBIF performs full factorial analysis (37–39) of feature expression data from single vs. dual factor exposures to identify molecular clusters that reveal synergy-mediating pathways, functions and regulators. To demonstrate the utility of OBIF, we applied this strategy to multi-Omics experimental data from epithelial cells exposed to Pam2-ODN to identify biologically relevant, unanticipated cooperative signaling events that we subsequently confirmed to be required for the synergistic pneumonia protection. Then, to demonstrate generalizability, OBIF was applied to datasets from diverse types of Omics platforms and experimental models, successfully identifying molecular clusters driving their combined responses.

MATERIALS AND METHODS

Experimental models

All mouse experiments were performed with 6- to 10-week-old C57BL/6J mice of a single sex. Immortalized human lung epithelial (HBEC3-KT) cells were cultured in keratinocyte serum-free media (KSFM) supplemented with human epidermal growth factor and bovine pituitary extract. Mouse lung epithelial (MLE-15) cells were cultured in RPMI supplemented with 2% fetal bovine serum. Cultures were maintained in the presence of penicillin and streptomycin. The cell lines used were authenticated by the MD

Anderson Characterized Cell Line Core Facility. All mouse experiments were performed in accordance with the Institutional Animal Care and Use of Committee of The University of Texas MD Anderson Cancer Center, protocol 00000907-RN01.

Exposure to TLR ligands

S-[2,3-bis(palmitoyloxy)-propyl]-(R)-cysteinyl-(lysyl) 3-lysine (Pam2 CSK4) and ODN M362 were reconstituted in endotoxin-free water, then diluted to the desired concentration in sterile PBS. For *in vivo* experiments, as previously described (24,25), the indicated ligands were placed in an Aerotech II nebulizer driven by 10 L/min air supplemented with 5% CO_2 for 30 min. The nebulizer was connected by polyethylene tubing to a polyethylene exposure chamber. Twenty-four hours prior to infections, 10 ml of PBS, Pam2 (4 μM) alone, ODN alone (1 μM) or both (Pam2 4 μM and ODN 1 μM) was delivered via nebulization to unrestrained mice for 30 min, and then mice were returned to normal housing. For *in vitro* experiments, PBS, Pam2 (2.2 μM) alone, ODN alone (0.6 μM) or both (Pam2 2.2 μM and ODN 0.6 μM) was added into 2 ml of the culture media 4 h prior to inoculation with virus (24,25).

Reverse-phase protein array

To simultaneously evaluate the expression of 161 regulatory proteins and phospho-proteins in HBEC3-KT cells after exposure to either PBS, Pam2, ODN or Pam2-ODN, a targeted high-throughput screening proteomic assay was performed by the Reverse Phase Protein Array Core Facility at The University of Texas MD Anderson Cancer Center (40,41). The RPPA included 4 biological replicates per treatment condition, and data is available at GitHub (www.github.com/evanslaboratory/Datasource).

Infection models

For *in vivo* infections, frozen stock (2.8 × 10^7 50% tissue culture infective doses [TCID50] ml − 1) of influenza A H3N2 (Wide *et al*, 1977), virus was diluted 1:250 in 0.05% gelatin in Eagle's minimal essential medium and delivered by aerosolization for 30 min to achieve a 90% lethal dose (LD90) to LD100 (∼100 TCID50 per mouse). Mouse health was followed for 21 d post infection, $n = 15$ mice per condition. Animals were weighed daily and sacrificed if they met euthanasia criteria, including signs of distress or loss of 20% pre-infection body weight. For *in vitro* infections, IAV (multiplicity of infection [MOI] of 1.0) with or without NF-κB inhibitor IMD-0354 at 25 ng/μL was added to cells in submerged monolayer and viral burden was assessed 24 h post infection.

Pathogen burden quantification

To measure transcript levels of IAV nucleoprotein (NP) gene, samples were harvested in RNAlater and RNA was extracted using the RNeasy extraction kit. About 500 ng total RNA was reverse transcribed to cDNA by using an iScript cDNA synthesis kit and submitted to quantitative

reverse transcription-PCR (RT-PCR) analysis with SYBR green PCR master mix on an Bio-Rad CFX Connect Real-Time PCR Detection System. Host 18S rRNA was similarly probed to determine relative expression of viral transcripts (24,42).

Omics dataset formatting

To perform analysis of interaction effects across both the Omics and feature level, OBIF performs a sequence of transformations to the unscaled original data matrix m that are tailored to meet the statistical assumptions needed for two-way ANOVA analysis of interaction terms across (43,44) different Omics platforms: (i) background correction is applied if it has not been corrected for abundant negative or low signal intensities (45–47); (ii) a $\log_2$-transformation of count- (i.e. raw counts from sequencing-based platforms) or continuous-based (normalized counts, signal intensity and peak area from sequencing-, array- and spectrometry-based platforms respectively) data are applied to provide a Gaussian-like data distribution, if the expression values are not normally distributed (48,49); and (iii) quantile normalization is applied to all datasets to equalize inter-sample variances (50,51). These transformations allow OBIF to have an analysis-ready (non-negative, non-zero values) (52,53) data matrix m (Figure 1D) with expression values and of dimensions f x n, where f is the number of features as rows and n is the number of samples S as columns. The appropriate sample order in dimensions n of m is:

$$n = S_{(0,0)}{}^1 + \ldots + S_{(0,0)}{}^i + S_{(1,0)}{}^1 + \ldots + S_{(1,0)}{}^i + S_{(0,1)}{}^1$$
$$+ \ldots + S_{(0,1)}{}^i + S_{(1,1)}{}^1 + \ldots + S_{(1,1)}{}^i$$

The subscripts denote the condition of the samples: exposed to neither factor (0,0), exposed to factor A alone (1,0), exposed to factor B alone (0,1) or exposed to both factors A and B (1,1). The superscripts represent the sample replicates from 1 to i within each of the four conditions.

Synergy definition

To evaluate non-additive phenotypes of combined exposures as synergistic, we used an effect-based definition of synergy using the response additivity metric (43,44,54). First, the observed effects of single and dual factor exposures are calculated from their main differences relative to the unexposed (sham) group:

$E_A = \Delta (S_{(1,0)} - S_{(0,0)})$ for individual exposure to Factor A,
$E_B = \Delta (S_{(0,1)} - S_{(0,0)})$ for individual exposure to Factor B,
$E_{AB} = \Delta (S_{(1,1)} - S_{(0,0)})$ for combined exposure to Factor A and Factor B.

Hence, the observed combination effect (E_{AB}) is compared to the expected non-interacting additive effect of individual exposures ($E_{ADD} = E_A + E_B$) where $E_{AB} > E_{ADD}$ indicates synergism, while $E_{AB} = E_{ADD}$ indicates additivity, and $E_{AB} < E_{ADD}$ indicates antagonism. Then, a significant interaction effect (P-value < 0.05) is systematically evaluated (1,5) to discriminate true synergistic or antagonistic combination effects from a non-interacting cooperation between factors when $E_{AB} \neq E_{ADD}$ (43,44,54). Significant interaction effects were calculated from a multivariate Cox model with an interaction term for survival data, a two-way ANOVA for viral NP expression data and using full factorial analysis for feature expression data.

Omics-based interaction analysis

To evaluate significant interaction terms between factors at the Omics level within a dataset, OBIF performs a multiple linear regression across its feature expression values:

$$I_O \; 0 + F_A + F_B + F_A \cdot F_B$$

where the Omics-based interaction (I_O) regression formula is equivalent to a two-way ANOVA analysis where the intercept is referenced to the control samples (*0*) and returns a statistical summary of terms for the individual factor A (F_A), factor B (F_B) and their interaction ($F_A \cdot F_B$). Goodness of fit is calculated from the adjusted R^2 values, and overall significance is determined by the P-values of F-statistics that result from the I_O regression.

Full factorial analysis of Omics data

OBIF performs full factorial analysis across the expression values of m to calculate the factorial effects of both expression and contrast analysis. The fixed-effect model fitted to the expression levels of features is:

$$L_f = \beta_0 + \beta_1 \cdot F_A + \beta_2 \cdot F_B + \beta_3 \cdot F_{AB}$$

where the expression levels of features (L_f) requires an intercept parameter to reference the control samples (β_0) and estimates β coefficients for the main effects of individual factor A (F_A) and factor B (F_B) and their combination (F_{AB}). Subsequent model fitting for each treatment condition (Factor A, B and AB) involves ß$_0$ and their respective beta coefficient (ß$_{1-3}$). Therefore, calculation of main effects for each condition requires subtraction of ß$_0$ from their respective fitted model (ß$_0$ + ß$_1$ for Factor A, ß$_0$ + ß$_2$ for Factor B, or ß$_0$ + ß$_3$ for Factor AB) to accurately represent the group comparisons used for differential expression for selection of differentially expressed molecules (DEMs):

$$ME \; of \; F_A = (\beta_0 + \beta_1) - -\beta_0 = \beta_1$$
$$ME \; of \; F_B = (\beta_0 + \beta_2) - \beta_0 = \beta_2$$
$$ME \; of \; F_{AB} = (\beta_0 + \beta_3) - \beta_0 = \beta_3$$

To adjust for multiple testing, the resulting P-values are controlled for false discovery rate (FDR) using the Benjamini–Hochberg procedure to select F_A DEMs, F_B DEMs and F_{AB} DEMs, respectively. The Bonferroni procedure and Tukey's range test can be applied to datasets that require control for family-wise error rate. Similarly for detection of interactive DEMs (iDEMs), the remaining multi-factor effects are estimated using contrast analysis for the simple main effects (SME) and their interaction effect:

$$SME \; of \; F_A = (\beta_0 + \beta_3) - (\beta_0 + \beta_2) = \beta_3 - \beta_2$$
$$SME \; of \; F_B = (\beta_0 + \beta_3) - (\beta_0 + \beta_1) = \beta_3 - \beta_1$$

$$Interaction \; effect \; F_A \times F_B (\beta_4) = [(\beta_0 + \beta_3) - (\beta_0 + \beta_2)]$$
$$- [(\beta_0 + \beta_1) - (\beta_0)] = \beta_3 - \beta_2 - \beta_1$$

A significance threshold of *P*-value < 0.05 is used to determine if F_{AB} DEMs are susceptible to any SME. iDEMs are defined as F_{AB} DEMs with a statistically significant interaction term (*P*-value is < 0.05) during full factorial analysis that results in their synergistic or antagonist expression after combinatory dual exposures. Hence, iDEMs are more representative of synergistic or antagonistic statistical interactions than of molecular interactions between DEMs.

Synergy quantification

To evaluate non-additive expression levels of individual features, OBIF uses an Interaction Score (IS) where the absolute ratio of the Log_2 fold change with combined factors (FC_{AB}) over the linear sum of the Log_2 fold change of individual factors ($FC_A + FC_B$) can identify antagonistic, synergistic or additive features using fold change principles where additivity is represented when $FC_{AB} = FC_A + FC_B$. A subsequent log_2-transformation can then quantify the size of the interaction effect by re-scaling the additivity threshold around 0 for additive features, with IS > 0 for synergistic and IS < 0 for antagonistic features.

$$IS = Log_2\left(\left| \left(Log_2 FC_{AB} \right) / \left(Log_2 FC_A + Log_2 FC_B \right) \right|\right)$$

Antagonistic and synergistic features (iDEMs) are differentiated from additive features with a significant interaction effect (*P*-value < 0.05) (44,55) calculated during full factorial analysis.

Feature selection with OBIF

All differentially expressed molecules (DEMs) were represented in heatmaps after hierarchically clustering using Ward's minimum variance method with Euclidean distances of $Log_2 FC$ values to compute dissimilarity by rows (features) and by columns (samples). Column dendrograms were plotted to represent the distance between samples, vertical side bar colors summarize DEMs according to their and horizontal side bars colors represent sample types by factors. Color scale keys indicate the levels of feature expression with upregulation in red and downregulation in green. DEMs with F_{AB} (Pam2-ODN) were clustered by principal component analysis based on the mean linear fold change difference to reveal the expression patterns biologically present across all factors F_A, F_B and F_{AB}. Principal components 1 and 2 were used for plotting DEMs with F_{AB} and the variability between features is marked in each axis. Expression profiles (EPs) were identified in the clusters for each individual feature according to Figure 3 C.

Mixed-effects model

We compared OBIF's capacity to detect interaction effects at the level of individual features against a mixed-effect model (56):

$$L_{f-Mix} = \beta_0 + \beta_1 \cdot F_A + \beta_2 \cdot F_B + \beta_4 \cdot F_A \cdot F_B + (1|S)$$

where the expression level of features in a mixed-effect model ($L_{f\text{-Mix}}$) is a function of the estimated β coefficients for the fixed effects of individual factor A (F_A) and factor B (F_B) and their interaction ($F_A \cdot F_B$) with a random effect (1|S) for all sample conditions ($S_{(0,0)}$, $S_{(1,0)}$, $S_{(0,1)}$, $S_{(1,1)}$). After fitting the mixed-effect model, empirical Bayesian shrinkage is used to estimate the model fitness from the t- and F-statistics of the $L_{f\text{-Mix}}$ regression formula. The significance threshold of the interaction term uses a *P*-value < 0.05 to determine interaction between F_A and F_B.

Beta-uniform mixture model

To compare the performance of OBIF against other models and bioinformatic tools, we applied a beta-uniform mixture (BUM) model using the ClassComparison R package to distributions of *P*-values in order to compute the 'ground truth'. The BUM model estimates a ground truth reference for a large set of *P*-values by calculating the occurrence of true positive (TP), false positive (FP), true negative (TN) and false negative (FN) detections under the fundamental properties that such *P*-values are distributed uniformly and can be expressed as mixture of false null hypothesis and true null hypothesis (57,58). This approach has been validated to address the multiple-testing errors often found in large sets of *P*-values calculated from two-way ANOVA models (as OBIF), correlation analyses, Kruskal–Wallis tests and others (57,59). Using these estimates, we systematically compared their discrimination ability by calculating the following parameters:

$$Precision = TP / (TP + FP)$$
$$Recall\,(Sensitivity) = TP / (TP + FN)$$
$$False\,Discovery\,Rate = FP / (TP + FP)$$
$$Diagnostic\,Odds\,Ratio = (TP/FN) / (FP + TN)$$

Receiver operating characteristic curve

Applying BUM estimates, we compared the model fitness of OBIF against other models or bioinformatic tools by calculating the following parameters:

$$Sensitivity\,(True\,Positive\,Rate) = TP / (TP + FN)$$
$$Specificity\,(True\,Negative\,Rate) = TN / (FP + TN)$$

Then, a receiver operating characteristic area under the curve (ROC AUC) is calculated for each model or tool by plotting their sensitivity against their 1- specificity.

Enrichment analysis

To provide biological interpretation of the full factorial analysis and classification of features, enrichment analysis was integrated in the pipeline (60) to determine candidate effectors and regulators, biological pathways and functional processes. OBIF allows export of feature lists and expression values of DEMs, EPs and iDEMs. These are formatted to be used with common bioinformatic enrichment tools, including GSEA (61,62), EnrichR (63,64) and Ingenuity Pathway Analysis (IPA) (65). Data presented here were initially analyzed with IPA (QIAGEN Inc. https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis) using core analysis with the expression values of DEMs, EPs, and iDEMs. Both gene and chemical Ingenuity Knowledge Base modules are used

as enrichment reference, considering only experimentally observed confidence levels for identification of direct and indirect relationships. The thresholds of significance for canonical pathways, upstream analysis, diseases and functions, regulator effects and network analysis were selected from those with a *z*-score $\geq$ 2 for activation or $\leq$ -2 for inhibition, and < 5% false discovery rates for all predictions (65). Validation of enrichment results of iDEMs from IPA was accomplished with EnrichR using the TRRUST Transcription Factor 2019 reference database and the top enriched terms were selected and clustered using their combined score (*P*-value + z-score) (63,64).

DNA-binding ELISA

HBEC3-KT were grown to 80–100% confluence in 24-well plates and treated with PBS, Pam2, ODN or Pam2-ODN for the indicated durations. Measurements of DNA-binding activity of NF-κB and AP-1 transcription factor subunits were performed from whole cell lysates were made using their respective TransAM Kit according to product directions. For signal detection, samples were read immediately for absorbance at 450 nm with reference wavelength at 655 nm on a microplate reader. Experiments were repeated in triplicate and statistical analysis was performed with unpaired Student's *t* test using GraphPad Prism 8.0 with a significance threshold of *P*-value < 0.05.

Detection of Nuclear Translocation

HBEC3-KT were grown to 80–100% confluence in 100 mm dishes and treated for 30 min with PBS, Pam2, ODN or Pam2-ODN with or without pretreatment with NF-κB inhibitor IMD-0354 at 25 ng/μl for 16 h. Cells were detached from the plate with a 5 min incubation at 37°C degrees with 3 ml of Accutase to prevent additional activation of transcriptional activity. Cells were pelleted in individual 15 ml tubes at 500 g for 5 min and suspended in 500 μl of eBioscience FOXP3 fixation/permeabilization buffer for 15 min at room temperature. Cells were stained with a LIVE/DEAD Fixable Near IR Dead Cell Dye and with a 1:1000 dilution of NF-κB/p50 (E-10) Alexa Fluor 647, NF-κB/RelA (F-6) Alexa Fluor 488, AP-1/cJun (G-4) Alexa Fluor 594 and AP-1/cFos (D-1) Alexa Fluor 546 conjugated antibodies for 1 h on ice and protected from light. After incubation, cells were pelleted and washed with 200 μl of sterile PBS 4 times, then resuspended in 100 μl sterile PBS. After the last wash, cells were pelleted and resuspended in 50 μl of sterile PBS and nuclear DAPI staining at 0.5 μg/ml was performed just prior to data acquisition on ImageStreamX MII. HBEC3-KT images were acquired using INSPIRE software on the ImagestreamX Mark II imaging flow cytometer (Amnis Corporation) at 60 × magnification, with lasers 405 nm (50.00 mW), 488 nm (200.00 mW), 561 nm (200.00 mW), 642 nm (200.00 mW) and side scatter (782 nm) (2.25 mW). About 10 000–25 000 images per sample acquired include a brightfield image (Channel 1 and 9), p65 Alexa Fluor 488 (Channel 2), c-Fos Alexa Fluor 546 (Channel 3), c-Fos Alexa Fluor 594 (Channel 4), side scatter (Channel 6), DAPI (Channel 7) and p50 Alexa Fluor 647 (Channel 11). The laser outputs prevented saturation of pixels in the relevant detection channels as monitored by the corresponding Raw Max Pixel features during acquisition. For image compensation, single color controls were stained with all fluorochromes and 500 events were recorded with each laser for individual controls. Fluorescent images were taken in all channels with brightfield LEDs and scatter lasers turned off to accurately capture fluorescence. Individual single-color control file was then merged to generate a compensation matrix and all sample files were processed with this matrix applied. After compensation for spectral overlap based on single color controls, analysis was performed, and individual cell images were created using IDEAS® software version 6.1. Cell populations were hierarchically gated first by single cells, then cells in focus, then negative selected for live cells, and finally as double positive for both DAPI and the transcription factor subunit of interest (Supplementary File S3D). The spatial relationship between the transcription factors and nuclear images was measured using the 'Similarity' feature in the IDEAS software to quantitate the mean similarity score in the cell populations per sample. A similarity score > 1 represents nuclear translocation, and the shift in distribution of nuclear translocation between two samples was calculated using the Fisher's Discriminant ratio (Rd value) (66).

Statistical analysis and quantification

Statistical analyses were performed using Prism 8 (Graph-Pad, San Diego, CA, USA) and R. Analysis of survival was performed by calculation of Kaplan–Meier curves and multivariate Cox model was used for group comparisons relative to PBS. Analysis of viral NP expression was performed using a two-way ANOVA with post hoc Tukey analysis for paired comparisons that was adjusted for multiple testing. Analysis of DNA-binding activity *in vitro* was performed using a Student's *t* test for comparisons between two groups or using one-way ANOVA for comparison between multiple groups. Grouped data are shown as means $\pm$ standard error of the mean. To verify the statistical assumptions for each test, Gaussian distribution was evaluated with Saphiro–Wilk test, and equal variance between two samples was evaluated with F-tests, or for more than two samples with Barlett's or Levene's test. Simultaneous multiple outlier detection was performed using the robust regression and outlier removal (ROUT) method with a *q* value of 5% (maximum FDR). Treatment allocation of animals was randomized in the experiments, though assessment could not be blinded. A pre-specified minimum requirement of three biological replicates for *in vitro* studies and 10 for *in vivo* studies.

Computational requirements

Data preprocessing and analysis with OBIF was performed in RStudio on both Mac- and Windows-based computers. For Mac-based analysis, R version 3.6.3 (2020–02-29) with x86_64-apple-darwin15.6.0 platform was used on a personal computer equipped with 2.7 GHz Quad-Core Intel Core i7 with 16GB of RAM under macOS 11.6 operating system. For Windows-based analysis, R version 4.1.0 (2021–05-18) with x86_64-w64-mingw32 platform was used on a desktop

computer equipped with 3.4 GHz Intel Core i-6700 CPU with 16GB RAM under Windows 10 Enterprise operating system. To provide a processing power reference for OBIF users in other computers, we use the benchmarkme R package (https://github.com/csgillespie/benchmarkme) to assess the calculation speed of matrix benchmark functions. We used the bm_matrix_cal_lm() function to model a linear regression over a 5000×5000 matrix and obtained an elapsed time of 0.857 and 0.870 s in a Mac- and Windows-based computer, respectively. We used the bm_matrix_cal_manip() function to create, transpose and deform a 5000×5000 matrix and obtained an elapsed time of 0.458 and 0.630 s in a Mac- and Windows-based computer, respectively.

RESULTS

Pam2-ODN synergistically induced epithelial resistance against pneumonia

Our laboratory's interest in synergistic interactions arises from our experience investigating single versus dual immunostimulatory treatments to prevent pneumonia (9,22–25,27,67). As a demonstrative example, data are presented here from influenza A virus (IAV) challenges of different models following pretreatment with Pam2 alone, ODN alone or the Pam2-ODN combination (i.e. both treatments delivered together). When mice are challenged with IAV 24 h after the indicated inhaled treatments, we observed little increase in survival after the individual treatments compared to sham-treated control mice, whereas mice treated with the Pam2-ODN combination demonstrated profound antiviral protection (Figure 1A). Similarly, when isolated mouse lung epithelial (MLE-15) cells were challenged with IAV 4 h after pretreatment with the individual ligands, we observed no significant reductions in the viral burden relative to PBS-treated cells. However, cells pretreated with Pam2-ODN showed a significant reduction in viral nucleoprotein (NP) gene expression as assessed by qPCR relative to host 18s gene (Figure 1B). Comparing the effect of combination ligand treatment (E_{AB}) to the expected response additivity of the individual ligand treatments ($E_{ADD} = E_A + E_B$) (44) reveals synergistic effects on both *in vivo* survival benefits and *in vitro* viral clearance (Figure 1C). To better understand the molecular mechanisms driving such unanticipated synergy, we developed OBIF as a phenotype-driven model (6) to understand the mechanisms underlying observed favorable combination exposure interactions that mediate synergistic responses and outcomes.

Development of a systems synergy model from experimental Omics data

To formally test whether the effect of combined factors (F_{AB}: Pam2-ODN) is greater than the expected linear sum of its individual factors (F_A: Pam2; F_B: ODN), an initial 2-level 2-factor (22) factorial design is required (43,44). Our strategy adapts the traditional analysis of variance (ANOVA) approach into a model that links the empirical analysis of synergy (43,44) with the high-throughput capacity and high-dimensionality of Omics datasets (68,69) (Fig-

ure 1D). As summarized in Figure 1E, OBIF integrates statistical and biological interactions in Omics data matrices from single versus dual factor exposures, leveraging Omics screening to promote discovery of the molecular drivers of synergy, and facilitating the biological validation of such regulators. The analytical pipeline is freely available as an R package at GitHub (www.github.com/evanslaboratory/OBIF) (Table 1). Naturally, the experimental approach to validating computational findings must be tailored to the individual tools and characteristics of the biological responses being studied.

OBIF allows sequential transformation of an unscaled original data matrix with background correction, Log_2-transformation, and quantile normalization (Figure 2A) in order to improve performance of an Omics-based interaction analysis between the studied factors (Figure 2B). This allows better adjusted R^2 and F-statistic *P*-values of the two-way ANOVA with improved detection of interaction terms (Supplementary File S1A) across multiple platform types. We further evaluated OBIF's overall model fitting using multiple standard test performance metrics calculated from 40 publicly available (4,27,55,56,70–80) Omics datasets (Supplementary File S5) that contain a broad range of samples ($n = 4$–30) and data sizes (161 to 217 507 features). Here, we found that OBIF's overall performance is best when the total sample size is $n \geq 12$ (i.e. ≥ 3 biological replicates/condition) because that sample size adjusted R^2 values ≥ 0.6 (Supplementary File S1B), area under the receiver-operating characteristic (ROC) curve (AUC) analysis ≥ 0.8 (Supplementary File S1D) and diagnostic odds ratios (DOR) ≥ 10 (Supplementary File S1F).

Full factorial analysis dissects hidden effects of combined factor exposures

Analysis of factorial effects in a data matrix from single versus combined factor exposures can statistically differentiate whether stochastic feature expression in a combination is correlated with the effect of an individual factor (simple main effect, SME) or their influence on each other (interaction effect) (81–83). Based on this principle, OBIF performs full factorial analysis through paired comparisons of calculated β coefficients in each condition (84) to discover main effects during expression analysis and multi-factor effects (SMEs and interaction effect) from contrast analysis (Figure 2C) (85). These statistical effects can be readily visualized (Figure 2D) to aid investigators in evaluating the adequacy of model fit when using OBIF for their unique datasets. Main effect *P*-values and expression changes are represented in volcano plots to facilitate determination of cut offs for individual analyses during feature discovery. Simple main effect calculations are represented in quantile-quantile (Q-Q) plots to detect normal and/or skewed distributions of individual features.

After confirming adequate model fitness (i.e. adjusted $R^2 > 0.5$, *F*-test < 0.05), OBIF's full factorial analysis on scaled data identifies DEMs from a dataset with better discriminatory power for interaction effects than a mixed-effects model (Supplementary File S1C), particularly with datasets of at least $n \geq 12$ samples (Supplementary File
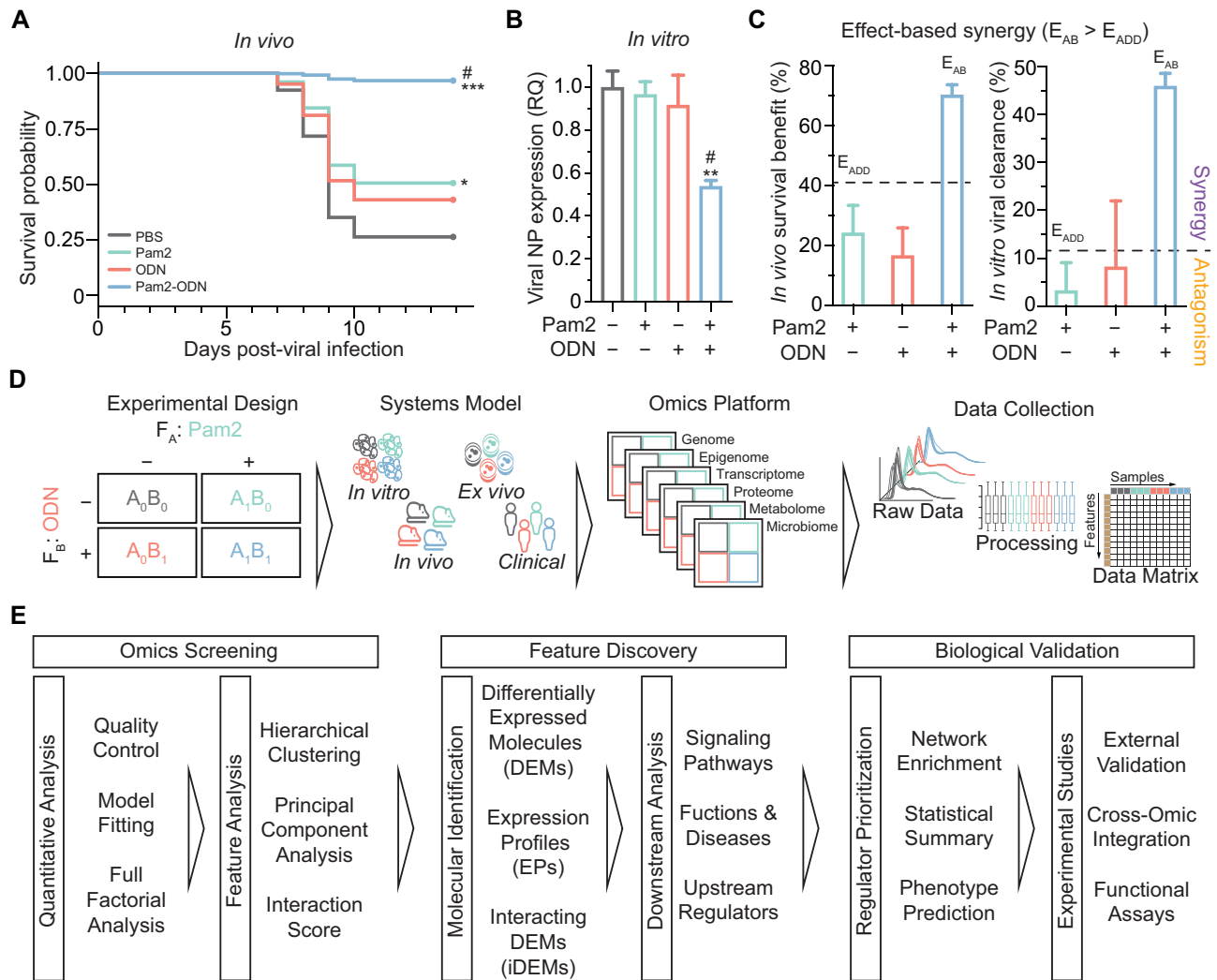
Figure 1. Omics-Based Interaction Framework: phenotype-driven synergy modeling and framework overview. (**A**) Mouse survival of influenza A challenge following the indicated pretreatments. $n = 30$ mice/group; aggregate of two independent experiments. (**B**) Virus burden of isolated mouse lung epithelial cells (MLE-15) after influenza A challenge following the indicated pretreatments. RQ, relative quantification of viral nucleoprotein (NP) expression to host 18s. $n = 6$ samples/condition. (**C**) Response additivity plots showing synergistic effects from antiviral responses in A (left) and B (right) relative to PBS. Dashed line indicates expected additive effect (E_{ADD}) of single ligand treatments over PBS. (**D**) Generic Omics workflow for phenotype-driven synergy modeling using a 2^2 experimental design. (**E**) Overview of OBIF, including (i) Omics screening of features in a data matrix, (ii) discovery of feature clusters of molecular drivers, and (iii) experimental validation of biologically relevant synergy regulators. Survival analysis was performed by Kaplan–Meier curves calculations. Group comparisons to PBS used a multivariate Cox model for survival data, and a two-way ANOVA with post hoc Tukey analysis for viral NP expression data: * $P < 0.05$, ** < 0.01, *** $P < 0.001$, # $P < 0.05$ of interaction effect. Grouped data are shown as means $\pm$ 95% confidence interval for survival data and standard error of the mean for viral NP expression data.

S1D). Similarly, OBIF outperforms other profiling tools commonly used to analyze multiple Omics data types, as demonstrated by superior values for area under the receiver operating characteristic (ROC) curve (AUC) and better false discovery rates, precision and recall (Supplementary File S1E) to select differentially expressed features from combination treatments with an improving overall performance in its diagnostic odds ratio as the sample size increases (Supplementary File S1F). Together, incorporation of full factorial analysis to dissect both main effects and hidden multi-factor effects allow enhanced analytical capacities over other available Omics profiling and processing tools (Table 2).

Differentially expressed molecules reveal synergy-specific pathways

To investigate the mechanisms underlying Pam2-ODN synergy, we used OBIF to re-analyze previously published (27) lung homogenate transcriptomic data from mice inhalationally treated with single versus combination ligands (GSE28994). Despite unsurprising overlap of detected differentially expressed features, OBIF outperformed commonly used methods of differential expression analysis with increased R2 values and more significant *P* values in all treatment conditions (Supplementary File S2A). Once feature expression is fitted (Figure 2D), this analysis identifies 3456 features as differentially expressed molecules (DEMs)

Table 1. OBIF package functions

Platform-dependent steps	Available functions
Background noise removal	Feature values import; RPKM/FPKM/TPM tresholding; Signal/Intensity/Area Normalization
Dataset transformation	Background correction; Linear/Log2 transformation; Quantile normalization
Quality control	Pre-/Post-scaling distribution plots; Clustering analysis; Outlier detection
Dataset management	Standardized analysis-ready output; Metadata support; Flexible data/figure export
Platform-independent steps	Available functions
Omics-level interaction analysis	Model fitness statistics; Two-way ANOVA test; Interaction plots
Feature-level interaction analysis	Full factorial analysis; Interaction scoring per feature; Feature discovery of DEMs, EPs and iDEMs
Quality assessment	Model diagnostic plots; Regression model comparisson; Interaction analysis performance evaluation
Data visualization	Hierarchical clustering; Heatmap; Venn/Euler diagram; Upset plot; PCA; Interaction score plot; OBIF Circos plot
Data analysis management	Standardized OBIF output; Data visualization tools; Flexible data/figure export

RPKM, Reads Per Kilobase Million; FPKM, Fragments Per Kilobase Million; TPM, Transcripts Per Kilobase Million; ANOVA, Analyisis of variance; DEMs, Differentially Expressed Molecules; EPs, Expression Profiles; iDEMs, interacting DEMs; PCA, Principal Component Analysis; OBIF, Omics-Based Interaction Framework.

2 h after treatment with Pam2, 2941 DEMs after ODN treatment, and 3138 DEMs after treatment with Pam2-ODN (Figure 3A). Despite the fact that 52% (1617/3138) of DEMs were shared by Pam2-ODN and the individual ligands, enrichment analysis using IPA software revealed an over-representation of 11 canonical cellular immune response pathways that were activated by Pam2-ODN (Figure 3B). Of these 11 immune and host defense pathways, most were only enriched by Pam2-ODN treatment, with a few enriched to a lesser degree by Pam2, and none by ODN.

Expression profiles summarize biological interactions and disentangle effectors of synergistic functions

Rather than relying on potentially redundant DEM clusters, OBIF classifies F_{AB} DEMs into eight expression profiles (EPs) that characterize cooperative and competitive biological interactions of individual factors (Figure 3C). EPs are defined by expression directionality (up- or down-regulation) of individual features and are not biased by the expression analysis of DEMs. Cooperative EPs have accordant expression directionality induced by F_A and F_B (i.e. both individual exposures induce increased feature expression or both induce decreased expression), while competi-

tive EPs have opposite directionalities induced by F_A and F_B. Among the cooperative EPs, concordant profiles result when F_{AB} directionality corresponds with the single factor effects (EPs I and II), and discordant profiles occur when F_{AB} directionality opposes the single factors (III and IV). Alternatively, among the competitive EPs, factor-dominant profiles are defined by F_{AB} directionality correspondence with one factor (F_A-dominant EPs V and VI; F_B-dominant EPs VII and VIII). Principal component analysis (Figure 3D) demonstrates that concordant EPs I and II were the most abundant in our dataset, followed by Pam2-dominant EPs V and VI. This abundance of EPs I and II better emphasizes the cooperative effects of both factors than does conventional DEM clustering alone (Supplementary File S2B).

Notably, enrichment analysis reveals that molecular effectors clustered by EPs correspond with Pam2-ODN-induced functions (Figure 3E), suggesting a biological basis for both the synergy and feature selection. Specifically, we found that features in EPs I and V contributed to host survival functions. Considered from an organizational perspective, induction of resistance to infection at the organismal level correlated with features in concordant EP I, at the cellular level with concordant EP II, and by leukocytes with Pam2-dominant EP V.

Multi-factor effects integrate biological and statistical interactions in EPs

Main effects determined significant DEMs per condition, while multi-factor effects explained whether Pam2-ODN DEMs and EPs resulted from SMEs and/or an interaction of individual ligands (Figure 3F). This analysis showed that most features in concordant EPs I and II are influenced by at least one multi-factor effect, while all features in discordant EPs III and IV are influenced by all multi-factor effects simultaneously. Not surprisingly, Pam2-dominant EPs V and VI and ODN-dominant EPs VII and VIII expression mainly results from their respective SMEs. This analysis also revealed that 67% (2116/3138) of Pam2-ODN DEMs are driven by the interaction effect of Pam2 and ODN as interactive DEMs (iDEMs). Thus, OBIF reconciled the biological interactions from EPs with the statistical interactions from multi-factor effects of Pam2-ODN.

iDEMs represent synergy drivers

Synergistic or antagonistic responses result from strong interaction effects between two factors in a combination (43,54). Hence, iDEMs integrate this principle during feature selection based on significant interaction effects between factors, allowing quantification of feature expression in a narrower set of DEMs. Using an interactions score (IS) that quantifies the effect size of the interaction effect relative to the additivity threshold, we can separate non-interactive DEMs (non-iDEMs) from true antagonistic ($IS < 0$) and synergistic ($IS > 0$) iDEMs (Figure 3G). This allows more focused enrichment analysis (Supplementary File S3A), in this case supporting NF-κB/RelA and AP-1/cJun as key transcriptional upstream regulators of Pam2-ODN's interaction effect and synergistic expression (Figure 3H), regardless of the enrichment method (Supplementary File
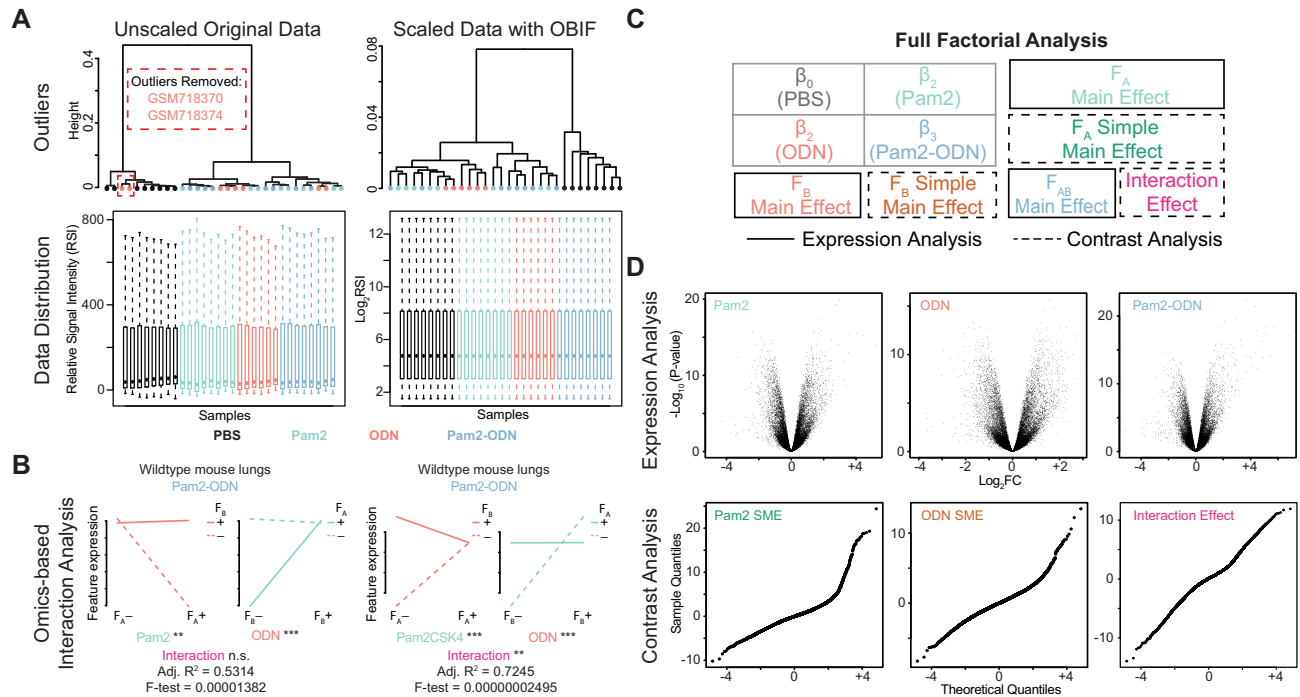
Figure 2. Overview of model fitting of Omics datasets during analysis with OBIF. (**A**) Quality control plots assess data distribution of GSE28994 with box plots (top) and detection of potential outliers by hierarchical clustering (bottom) in both the unscaled original data (left) and the scaled data with OBIF (right). (**B**) Omics-based interaction analysis of the data in (A) comparing interaction plots, coefficient significance and goodness of fit of unscaled and scaled data. (**C**) Scheme of full factorial analysis performed by OBIF from β coefficients. (**D**) Representative volcano plots of full factorial analysis from analysis ready data for each condition after expression analysis (top) and Q-Q plots of moderated *t*-statistics for each multi-factor effect after contrast analysis (bottom); FC, fold change; DEM, differentially expressed molecule; EP, expression profile; iDEM, interactive DEM; RSI, relative signal intensity; SME, simple main effect.

S3A,B). To visually represent the DEMs, EPs and iDEMs with the integrative analysis of biological and statistical interactions, OBIF generates a custom Circos plot that summarizes the main characteristics of these synergy-mediating feature clusters (Figure 3I).

Discovery of synergy-mediating feature clusters allows cross-Omics validation of molecular drivers

Downstream analyses of DEMs, EPs and iDEMs have the capacity to discern the contributing roles of individual factors to a combination treatment regardless of the Omics platform being studied. To demonstrate the cross-Omics function of OBIF, we analyzed our reverse-phase protein array (RPPA) data from single- or dual-treated human lung epithelial cells (Supplementary File S2C) and were able to visually compare the prevalence of DEMs, EPs and iDEMs in transcriptomics and proteomics datasets (Supplementary File S2D). Interestingly, the use of EPs improved the analysis of hierarchical clustering given that the latter approach alone did not reveal distinctive gene clusters to explain the synergistic response while the abundance of concordant EPs I and II highlighted the contribution of ODN to the protective combination that might otherwise be overlooked, as enrichment analysis showed far fewer signaling pathways (Figure 3B) and had a greater clustering distance from Pam2-ODN samples (Supplementary File S2B) from the microarray. Similar to our previous findings with iDEMs (Figure 3H), transcription factors from many pathways were involved, and NF-κB and AP-1 transcription factors subunits remained central elements of the regulatory network analysis from our microarray results while our RPPA data identified the top phospho-signaling DEMs and cross-validated STAT3, NF-κB/RelA and AP-1/cJun as transcriptional regulators involved in the Pam2-ODN signaling network (Supplementary File S2E).

Experimental validation of molecular regulators of Pam2-ODN synergy

Prompted by the foregoing results, we tested whether NF-κB and AP-1 subunits were biologically relevant synergy regulators of Pam2-ODN-induced epithelial resistance. DNA-binding activity of a panel of NF-κB and AP-1 family subunits in human lung epithelial cells (HBEC3-KT) after stimulation with Pam2-ODN confirmed that NF-κB/RelA and AP-1/cJun subunit activation was strongly increased after 15 min of treatment without significant contribution of other subunits (Figure 4A and Supplementary File S3C). Indeed, NF-κB/RelA and AP-1/cJun subunits exhibited surprisingly similar activation kinetics after Pam2-ODN treatment, further supporting cooperation or coordination (Figure 4B). Investigating this co-activation of non-redundant transcriptional families, single-cell nuclear translocation of canonical NF-κB (p50/RelA) and AP-1 (cFos/cJun) complexes was assessed in HBEC3-KT cells by imaging flow cytometry. We found that all transcriptional subunits exhibited an increased nuclear translocation (sim-

Table 2. Function comparison of Omics profiling tools

|		Omics profiling and processing tools									
		OF	LA	D2	ER	LE	DS	PI	MP	MA	MG
Omics input type	Genome	•	•	•	•	•	•	•			
	Epigenome	•	•	•	•	•	•	•			
	Transcriptome	•	•	•	•	•	•			•	
	Proteome	•							•		
	Metabolome	•								•	
	Microbiome	•									•
Preprocessing	Single-omics data import	•	•	•		•	•	•	•	•	•
	Multi-omics data import	•				•				•	
	Dataset aggregation	•	•	•	•	•				•	•
	Unique feature ID	•				•			•	•	•
	Signal normalization	•	•	•			•	•	•	•	•
	Background correction	•							•	•	•
	Log2 transformation	•			•				•		•
	Variance stabilization	•		•	•					•	•
	Outlier detection	•									
	Metadata support	•					•			•	
Analysis	Tests for interaction effects	•	•	•	•	•					
	Pairwise differential expression	•	•	•	•	•			•	•	•
	Multi-group differential expression	•	•	•	•	•			•	•	•
	Omics-based interaction analysis	•									
	Feature-level full factorial analysis	•									
	Fold change value calculations	•	•	•	•				•		•
	Combination index calculations	•									
	Interaction score calculations	•									
	P-value calculations/corrections	•	•	•	•	•			•	•	•
	Sample/feature clustering	•	•	•					•	•	•
Visualization	Data distribution plots	•					•	•	•	•	•
	Pre-/Post-scaling distribution plots	•					•		•		
	Dimensionality Reduction plots	•	•	•			•	•	•		
	Sample/Feature clustering plots	•	•	•			•	•	•		•
	Expression/Significance plots	•	•	•		•	•	•	•		•
	Heat maps	•	•	•			•	•	•	•	
	Feature intersection/overlap plots	•	•	•						•	
	Interaction score plots	•									
	Summarizing circos plot	•									
	Model fit diagnostic plots	•	•	•		•			•	•	•
	Regression model comparison plots	•				•					
	Interaction analysis performance plots	•									
Outputs	Standardized analysis-ready dataset	•	•	•					•	•	•
	Standardized analysis output data	•	•	•		•			•	•	•
	Standardized feature selection clusters	•							•	•	
	Source code file export	•	•	•		•				•	•
	Data analysis file export	•	•	•		•				•	•
	Flexible data/figure export	•	•	•		•	•	•	•	•	•

OF, Omics-based interaction framework (OBIF); LA, Limma/Glimma; D2, DESeq/DESeq2; ER, EdgeR;
LE, Lme4; DS, Devis; PI, PIVOT; MP, MaxQuant/Perseus; MA, Metabo Analyst; MG, MetagenomeSeq.

ilarity score > 2) after 15 min of Pam2-ODN treatment relative to the PBS-treated cells (Figure 4C). However, neither Pam2 nor ODN alone induced the same magnitude of nuclear translocation, whether assessed by change in similarity scores (Rd value) or by the percentage of translocated cells (Figure 4D) relative to PBS-treated cells.

Discovery of novel NF-κB and AP-1 cooperation required for antiviral protection

To differentiate transcriptional cooperation from coincidental transcriptional activation after Pam2-ODN treatment, we assessed the Pam2-ODN-induced nuclear co-translocation of NF-κB (p50/RelA) and AP-1 (cFos/cJun) complexes in the presence or absence of NF-κB inhibitor IMD-0354 (IMD). As expected, pre-treatment with IMD alone reduced the Rd Value and percentage of translocated cells for NF-κB/RelA and NF-κB/p50 subunits without significantly modifying the percentage of translocation for AP-1/cJun and AP-1/cFos subunits. However, NF-κB inhibition with IMD also unexpectedly reduced the Pam2-ODN-induced similarity score shifts and nuclear translocation of AP-1 subunits, particularly of AP-1/cFos (Figure 4E). This indicates that NF-κB inhibition impaired Pam2-ODN-induced AP-1 nuclear translocation, supporting the cooperative regulation of these two non-overlapping signaling pathways. Representative images shown in Figure 4F demonstrate that inhibition with IMD reduced Pam2-ODN-induced heterodimerization and nuclear translocation of NF-κB (p50/RelA) and AP-1 (cFos/cJun) complexes. Further, we confirmed that disruption of this transcriptional cooperation was sufficient to impair the in-
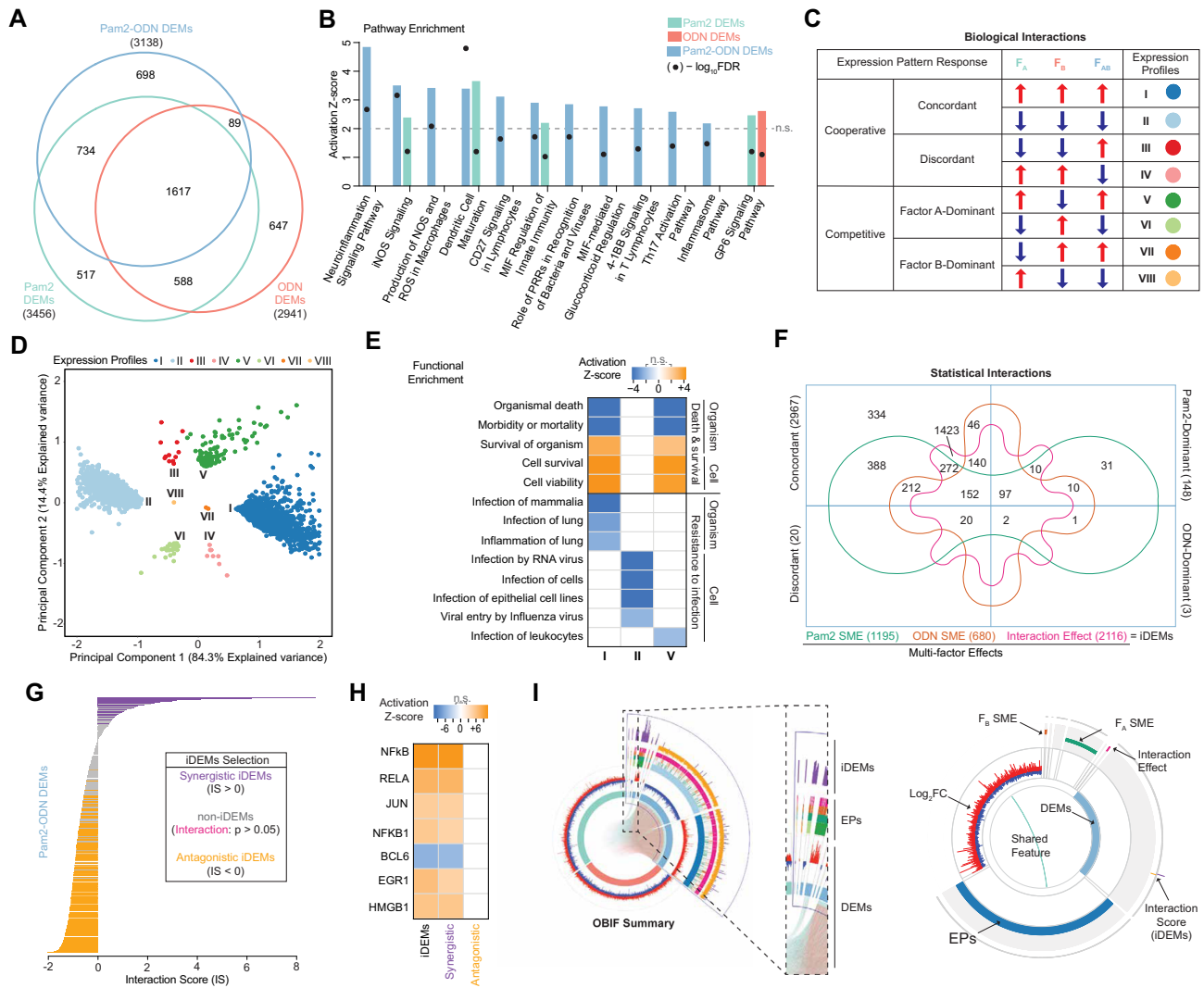
Figure 3. Full factorial analysis allows discovery of synergy-mediating feature clusters. (**A**) Euler diagram of DEMs following single or dual treatment in mouse lung homogenates. (**B**) Most overrepresented activated canonical pathways after IPA enrichment of DEMs. (**C**) Expression profiles depict biological interactions during dual factor exposure. (**D**) Principal component analysis of Pam2-ODN DEMs identified by expression profiles. (**E**) Top activated (orange) and inhibited (blue) diseases and functions after IPA enrichment of expression profiles. (**F**) Venn diagram of Pam2-ODN DEMs correlated by expression profiles classes and multi-factor effects. (**G**) Interaction score (IS) of Pam2-ODN DEMs, including antagonistic iDEMs (IS < 0, orange) and synergistic iDEMs (IS > 0, purple). (**H**) Top activated (orange) or inhibited (blue) transcriptional regulators after IPA enrichment of iDEMs. FC, fold change. DEMs, differentially expressed molecules. iDEMs, interactive DEMs. SME, simple-main effects. (**I**) Visual summary of OBIF's outputs including DEMs, EPs and iDEMs (left) plotted into 3 rings (right, from the inside out): (i) DEMs, inner links connect shared features between single (light green or orange) and dual (light blue) exposures followed by their Log₂FC values; (ii) EPs, where inner sectors represent individual profiles (I to VIII) followed by their statistically significant F_A-SME (green), F_B-SME (orange) or $F_A \cdot F_B$ interaction effect (pink); and (iii) iDEMs, represented by their synergistic (purple) or antagonistic (yellow) interaction scores. FC, fold change; DEMs, differentially expressed molecules; EPs, expression profiles; iDEMs, interactive DEMs; SME, simple main effect. n.s., non-significant threshold of Z-score (dashed line).

ducible viral burden reduction seen with Pam2-ODN (Figure 4G).

Applications of OBIF across multiple platforms and conditions

To demonstrate its generalizability, we used OBIF to interrogate the synergistic mechanisms of different datasets using a single- (Figure 5A) or a multi-Omics approach (Figure 5B). In our first case study, OBIF analyzed transcriptome data from untargeted total RNA-seq of an *in vivo* mouse model of K-ras mutant lung adenocarcinoma (CC-

LR) (56). In this model, the exposure factors are sex (F_A: Male) and genetic alterations (F_B: conditional Stat3 knockout lung epithelial cells), where only male mice without STAT3 gene expression (F_{AB}: Male + Stat3 K.O.) have a significant increase of tumor burden (Figure 5A). To investigate this currently unexplained interaction, we performed functional enrichment analysis of iDEMs, and disentangled a tumorigenic cluster (antagonistic iDEMs) from an immunosuppressive cluster (synergistic iDEMs). The investigators' previously suspected that both tumorigenic changes and failures of immune surveillance in the tumor microenvironment might result in the observed synergistic phenotype
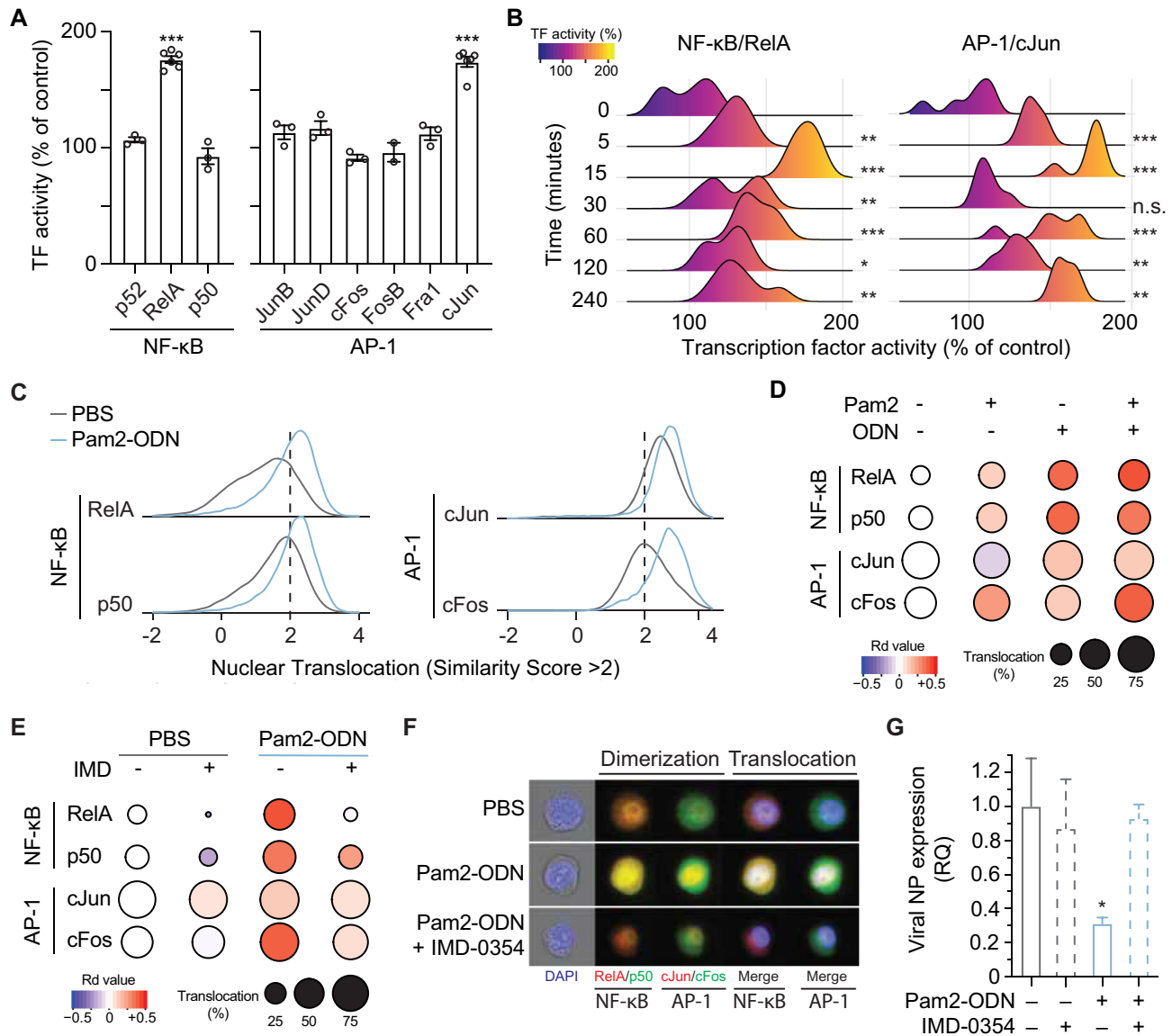
Figure 4. Regulators identified with OBIF uncovered cooperation between RelA and cJun that is required for synergistic antiviral protection. (**A**) Transcription factor (TF) activity of NF-κB and AP-1 subunits 15 min after treatment of human lung epithelial cells with Pam2-ODN. $n = 3$–6 samples/condition. (**B**) RelA and cJun activity at indicated times after Pam2-ODN treatment. $n = 6$ samples/conditsssion. (**C**) Nuclear translocation scores of NF-κB and AP-1 heterodimers after Pam2-ODN treatment. Shown as normalized frequency of similarity score per condition. (**D**) NF-κB and AP-1 subunit nuclear translocation in C (increased, red; decreased, blue) per condition. (**E**) NF-κB and AP-1 subunit nuclear translocation with or without NF-κB inhibition by IMD-0345. (**F**) Representative imaging flow cytometry images of hetero-dimerization and nuclear translocation of NF-κB and AP-1 in (D) and (E). (**G**) Virus burden of mouse lung epithelial cells challenged with influenza A with or without NF-κB inhibition. $n = 4$ samples/condition. Comparison to baseline activity used Student's t test for individual TF subunits, and one-way ANOVA for multiple TF timepoints: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; n.s., non-significant. Grouped data are shown as means ± standard error of the mean.

(56), but OBIF provides the first mechanistic support for that hypothesis. Additionally, network analysis of iDEMs revealed four previously unrecognized regulators (CCL28, CAGA, CAGB and CSF3R) that help explain the immunosuppressive molecular mechanisms that are permissive for the increased tumor burden in Male + Stat3 K.O. CC-LR mice.

In our second case study, OBIF analyzed both transcriptome data from an untargeted microarray and proteomic data from a targeted RPPA of an *in vitro* cell culture model of a chemotherapy-resistant adenosquamous lung carci-

noma. In this model, cancer cells are exposed to a drug (F_A: Volasertib, a polo-like kinase 1 inhibitor) and/or a cytokine (F_B: TGF-ß). It has been shown that tumor growth is inhibited only by combined exposure (F_{AB}: Volasertib + TGF-ß) (72). Here, iDEMs upstream regulator analysis and hierarchical clustering from both platforms co-identified transcription factor cJun from antagonistic iDEMs as a central regulator of the observed exposure synergy. Further, building on the investigators' report that the cMet/FAK/SRC axis regulates tumor growth inhibition in this model (72), network analysis of iDEMs yields downstream mechanistic
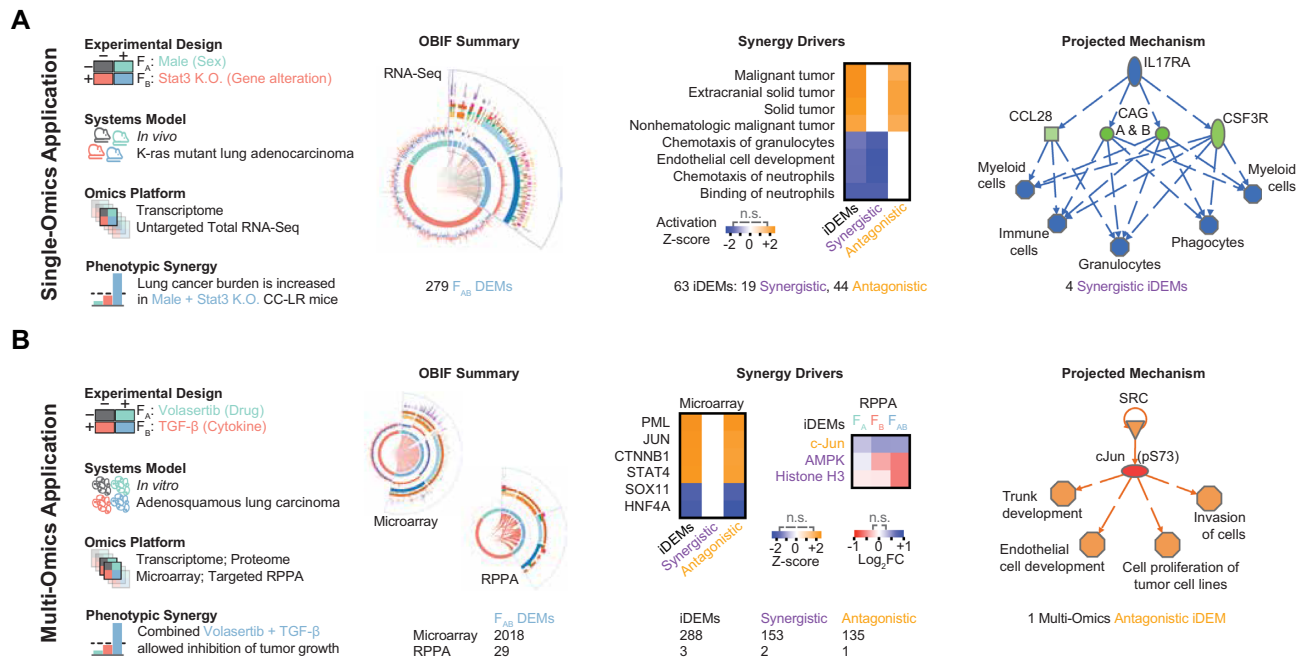
Figure 5. Single- and multi-Omics applications of OBIF in alternate systems. (**A**) Single-Omics application of OBIF in a K-ras mutant lung adenocarcinoma using RNA-seq data. The exposures investigated are presence of male sex (F_A) and presence of a Stat3 gene alteration (F_B). The displayed outputs are (i) OBIF summary analysis; (ii) functional enrichment and (iii) network analysis of iDEMs. (**B**) Multi-Omics application of OBIF in an adenosquamous lung carcinoma using Microarray and RPPA data. The exposures investigated are treatment with volasertib (F_A) and treatment with TGFβ (F_B). The displayed outputs are (i) OBIF summary analysis, (ii) Upstream regulator enrichment and heatmap clustering and (iii) network analysis of iDEMs.

insights into tumor growth regulation and epithelial to mesenchymal transformation through cJun that were not previously hypothesized.

OBIF was similarly applied to additional datasets derived from Omics platforms including microarray (27), RNA-seq (56), RPPA (72) and mass spectrometry-based metabolomics (4) investigations of diverse factor classes and biological systems (Supplementary File S4) to demonstrate that the synergy-mediating clusters can be universally identified. The statistical summary of the Omics-based interaction analysis demonstrates the effects of individual factors and interactions in other Omics platforms through interaction plots and statistical significance calculations (Supplementary File S4A). OBIF's full factorial analysis can be broadly applied to identify: (i) DEMs (Supplementary File S4B) from both individual and combined factors; (ii) EPs that represent the biological interactions in an unassisted fashion from F_{AB} DEMs (Supplementary File S4C); and iDEMs that represent the statistical interactions in feature expression through IS plots (Supplementary File S4D). These are displayed in a visual summary using a Circos plot containing co-expressed features, DEMs, Log$_2$FC, EPs, multi-factor effects and iDEMs with their interaction score (Supplementary File S4E).

DISCUSSION

Synergistic and antagonistic interactions are common in nature and frequently promote efficacy of therapeutic interventions (1–5). While synergy quantification from dose–response data, combinatorial screening of molecule libraries, and other synergy prediction may suggest poten-

tially beneficial conditions or treatments, they do not provide substantive insights into the underlying molecular mechanisms (1). Thus, synergy-mediating pathways cannot be strategically targeted in rational drug development only with synergy quantification or prediction models.

Our interest in synergy arose from our observations of the strikingly favorable interactions of one such empirically derived combination, Pam2-ODN. While we could easily quantify the superiority of protection conferred by the dual treatment, in the absence of a systems theory to interrogate synergistic mechanisms (1,5), we were limited in our capacity to use available Omics datasets to deduce the mechanisms mediating the synergy. This is important because, although this lack of mechanistic understanding does not limit the utility of the current combination, it precludes development of next generation interventions that more precisely (perhaps, more efficaciously) target the synergy-driving pathways with fewer off-target (potentially toxic) effects. In contrast to prediction and quantification models, OBIF was developed as a synergy interrogation model with the explicit intent to investigate observed synergistic events instead of combined exposures with additive-only effects or phenotypes (a potential misuse) for which generalized additive models are already available (86,87). As such, it is inherently a phenotype-driven model that performs full factorial analysis on feature expression data from single vs. dual factor exposures to identify molecular clusters that reveal synergy-mediating pathways, functions and regulators.

Current synergy quantification methods and tools (34–36) use accepted statistical approaches to detect drug–drug interactions (34) (e.g. Loewe additivity, Bliss additivity,

Chou-Talalay Method, response surface methods, isobologram analysis, etc.). These approaches use dose-effect metrics to identify unestablished interactions (screening) that cannot be easily applied to Omics-based datasets because of the large number of observations and resources needed, while OBIF was designed to investigate mechanisms of observed instances of synergy (phenotype-driven). Although OBIF is not designed to predict toxicities, as do some of these synergy quantification tools, enrichment analysis of cytotoxic pathways or functions may also provide insights into the safety and tolerability of certain combined exposures. Further, these approaches require a series of observations derived from dose–response data (continuous variables) to conform to their required principles for synergy detection, face validity and formal validity (34,43,44,54,88–90), whereas OBIF can be applied to beyond continuous variables (i.e. dichotomous or categorical factors). This principle expands the potential applications of OBIF to study multiple factor classes (i.e. cytokines, cell lines, drugs, environmental factors, genetic alterations, gender differences, time), system levels (i.e. *in vitro*, *in vivo*, *ex vivo*, *clinical*) and both targeted and untargeted platform formats (i.e. array-based, sequencing-based, mass spectrometry-based) (Supplementary File S5).

Other Omics-based models rely on molecular interactions (30), pathway- (28,32) or network-based analysis (12,28,31), and drug response similarity scores (29,33) to predict synergistic drug combinations. While these predictive strategies have facilitated the discovery of some novel drug combinations (1,7,10), they largely depend on modeling of drug targets, drug actions, and organism- or disease-specific data inputs that limit their capacity to explore potential mechanisms and hinder their transferability to other Omics platforms or research fields. In contrast, OBIF is capable of revealing mechanistically relevant molecular clusters (DEMs, EPs and iDEMs) directly from single- or multi-Omics expression data following combined exposures and can be easily transferred for downstream analysis with a wider span of bioinformatic tools. Thus, while established synergy prediction and quantification tools can offer important contributions to the discovery of drug–drug interactions, the objectives of a synergy interrogation tool are entirely different. In contrast to these types of tools, OBIF is applied only after synergy has been observed in any of a wide range of biological exposures and contexts, and focuses exclusively on the mechanistic interrogation of the drivers of synergistic interactions rather than on detecting undiscovered/unanticipated synergistic combinations *de novo*, and is capable of achieving this via independent single Omics dataset analysis with high correspondence when OBIF is serially applied to both cross-platform or multi-Omics analysis.

Using Pam2-ODN datasets as demonstrative examples, OBIF identified unanticipated transcriptional cooperation between non-redundant transcription factors, NF-κB (p50/RelA) and AP-1 (cFos/cJun) complexes, as a molecular mechanism of inducible synergistic protection against IAV. Thus, by facilitating understanding of combined factor exposures in terms of the individual components, a computational discovery allowed focused experimental validation of a discrete, novel mediator of a synergistic biological response.

Perhaps as importantly, the computational analyses were accomplished by integration of data from different Omics platforms, different specimen types, and even different host species (Figure 5 and Supplementary File S4). In all three detailed examples, OBIF revealed molecular mechanisms that were congruent with previously published findings and suggested novel mechanistic hypothesis for experimental validation. These applications already include single (Figure 3A–H) and multi-Omics (Supplementary File S2B–E) tested scenarios for Pam2-ODN's specific RelA and cJUN interaction (Figure 4A–F) and broad antiviral response (Figure 4G). These applications also include single (Figure 5A) and multi-Omics (Figure 5B) untested scenarios where specific (Figure 5B) and broad (Figure 5A) molecular drivers and mechanisms have been discovered using OBIF. Thus, OBIF can be used to rationally focus investigation into phenotypically observed synergy in a broad array of contexts, including currently unexplored synergistic gene mutation or copy number alterations from mapped reads in DNA copy number variation sequencing (CNV-seq) or from intensity signals obtained from array-based comparative genomic hybridization (aCGH) (91).

Unlike most 2^2 designs, OBIF dissects factorial effects of combined factor exposures through full factorial analysis of feature expression data in a single processing step better than conventional differential expression pipelines (i.e. LIMMA, DESeq2, MetaboAnalyst) that are limited to a few platform-specific Omics dataset. This allows simultaneous identification of DEMs directly from main effects of single or combined factors, overcoming pairwise comparisons to control and repetitive analysis of each condition. While this simultaneous identification of DEMs could alternately be performed with a mixed-effect model, we showed how the mixed-effect model is suboptimal to detect interaction effects at the level of individual features and to facilitate iDEM selection when compared with full factorial analysis. Additionally, clustering by DEMs, EPs and iDEMs improves the specificity of enrichment analysis to disentangle the signaling pathways, functions and regulators of this synergistic combination and to capture their specific driving features. Further, quantification of multi-factor effects (SMEs and interaction effects) reveals whether particular features, molecular clusters or functions enriched by combined exposures are the result of individual factors or their crosstalk.

These statistical relationships have biological analogues that are integrated during feature discovery of DEMs, EP and iDEMs. Because each study of combined exposures using Omics datasets is unique, fixed cut-offs are not appropriate for every application. These statistical effects are visualized in volcano and Q-Q plots during model fitting to tailor each application. In fact, concordant EPs I and II rescued the underrepresentation of ODN observed in distance-based clustering and enrichment analysis. Further, iDEMs derived from features with significant interaction effects allow focusing discovery on molecular regulators and the calculation of interaction scores allows quantification of their synergistic or antagonistic expression. Thus, unlike most systems models of synergy, OBIF facilitates integrative analyses of biological and statistical interactions that are easily discoverable and interpretable through molecular

clusters representing the complex dynamics of dual exposures.

OBIF is available as an open-source R package with a semi-automated pipeline to facilitate its broad application to unscaled original data from various targeted and non-targeted Omics platforms, factor classes and biological systems. Compared to other commonly used analysis tools and modeling methods, OBIF demonstrates superior differential expression analysis for combined factor exposures and detection of interaction effects at the whole Omics and feature level. We have shown that OBIF can be fitted to perform full factorial analysis and that it adequately identifies DEMs, EPs, iDEMs and their attendant values and scores to promote discovery of molecular drivers of synergy in multiple, diverse datasets.

In summary, OBIF provides a phenotype-driven systems biology model that allows multiplatform dissection of molecular drivers of synergy. And, we encourage the application of OBIF to provide holistic understanding in research fields where greater-than-additive beneficial combinations remain understudied.

DATA AVAILABILITY

The codes generated during the current study are available in the EvansLaboratory GitHub repository for the OBIF R Package (www.github.com/evanslaboratory/OBIF) and OBIF R Code Extensions (www.github.com/evanslaboratory/Extensions). The datasets generated during this study are available in the EvansLaboratory GitHub repository for the Reverse-phase protein array data (www.github.com/evanslaboratory/Datasource). Further information and reasonable requests for resources, data and reagents used in this work should be directed to and will be fulfilled by the corresponding author, Scott E. Evans (seevans@mdanderson.org).

The published datasets used and analyzed during this study are available at Gene Expression Omnibus (GEO) repository or with original publication for Microarray GSE28994 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc = GSE28994), RNA-seq GSE109000 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc = GSE109000), Reverse-phase protein array at EMBO Molecular Medicine (https://www.embopress.org/doi/abs/10.15252/emmm.201809960) and Metabolomics data at Frontiers in Pharmacology (https://www.frontiersin.org/articles/10.3389/fphar.2019.00754/full). A list of publicly available datasets tested to be compatible with OBIF is included in Supplementary File S5 for future research applications.

SUPPLEMENTARY DATA

Supplementary Data are available at NARGAB Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Chen,D., Liu,X., Yang,Y., Yang,H. and Lu,P. (2015) Systematic synergy modeling: understanding drug synergy from a systems biology perspective. *BMC Syst. Biol.*, **9**, 56.
2. Zappasodi,R., Merghoub,T. and Wolchok,J.D. (2018) Emerging concepts for immune checkpoint blockade-based combination therapies. *Cancer Cell*, **34**, 690.
3. Ronzitti,G. and Mingozzi,F. (2018) Combination therapy is the new gene therapy?*Mol. Ther.*, **26**, 12–14.
4. Han,M.-L., Liu,X., Velkov,T., Lin,Y.-W., Zhu,Y., Creek,D.J., Barlow,C.K., Yu,H.H., Zhou,Z., Zhang,J. *et al.* (2019) Comparative metabolomics reveals key pathways associated with the synergistic killing of colistin and sulbactam combination against multidrug-resistant acinetobacter baumannii. *Front. Pharmacol.*, **10**, 754.
5. Wei,P.L., Gu,H., Liu,J. and Wang,Z. (2018) Development of fangjiomics for systems elucidation of synergistic mechanism underlying combination therapy. *Comput. Struct. Biotechnol. J.*, **16**, 565–572.
6. Hasin,Y., Seldin,M. and Lusis,A. (2017) Multi-omics approaches to disease. *Genome Biol.*, **18**, 83.
7. Zhu,J., Shi,Z., Wang,J. and Zhang,B. (2015) Empowering biologists with multi-omics data: colorectal cancer as a paradigm. *Bioinformatics*, **31**, 1436–1443.
8. Bouhaddou,M., Barrette,A.M., Stern,A.D., Koch,R.J., DiStefano,M.S., Riesel,E.A., Santos,L.C., Tan,A.L., Mertz,A.E. and Birtwistle,M.R. (2018) A mechanistic pan-cancer pathway model informed by multi-omics data interprets stochastic cell fate responses to drugs and mitogens. *PLoS Comput. Biol.*, **14**, e1005985.
9. Alfaro,V.Y., Goldblatt,D.L., Valverde,G.R., Munsell,M.F., Quinton,L.J., Walker,A.K., Dantzer,R., Varadhachary,A., Scott,B.L., Evans,S.E. *et al.* (2014) Safety, tolerability, and biomarkers of the treatment of mice with aerosolized Toll-like receptor ligands. *Front. Pharmacol.*, **5**, 8.

10. Leung,E.L., Cao,Z.W., Jiang,Z.H., Zhou,H. and Liu,L. (2013) Network-based drug discovery by integrating systems biology and computational technologies. *Brief. Bioinform.*, **14**, 491–505.
11. Preuer,K., Lewis,R.P.I., Hochreiter,S., Bender,A., Bulusu,K.C. and Klambauer,G. (2018) DeepSynergy: predicting anti-cancer drug synergy with deep learning. *Bioinformatics*, **34**, 1538–1546.
12. Huang,L., Brunell,D., Stephan,C., Mancuso,J., Yu,X., He,B., Thompson,T.C., Zinner,R., Kim,J., Davies,P. *et al.* (2019) Driver network as a biomarker: systematic integration and network modeling of multi-omics data to derive driver signaling pathways for drug combination prediction. *Bioinformatics*, **35**, 3709–3717.
13. Lee,J.H., Kim,D.G., Bae,T.J., Rho,K., Kim,J.T., Lee,J.J., Jang,Y., Kim,B.C., Park,K.M. and Kim,S. (2012) CDA: combinatorial drug discovery using transcriptional response modules. *PLoS One*, **7**, e42573.
14. Tang,J., Karhinen,L., Xu,T., Szwajda,A., Yadav,B., Wennerberg,K. and Aittokallio,T. (2013) Target inhibition networks: predicting selective combinations of druggable targets to block cancer survival pathways. *PLoS Comput. Biol.*, **9**, e1003226.
15. Collaborators,G.L. (2017) Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the global burden of disease study 2015. *Lancet Infect. Dis.*, **17**, 1133–1161.
16. McAllister,D.A., Liu,L., Shi,T., Chu,Y., Reed,C., Burrows,J., Adeloye,D., Rudan,I., Black,R.E., Campbell,H. *et al.* (2019) Global, regional, and national estimates of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015: a systematic analysis. *Lancet Glob Health*, **7**, e47–e57.
17. Metlay,J.P., Waterer,G.W., Long,A.C., Anzueto,A., Brozek,J., Crothers,K., Cooley,L.A., Dean,N.C., Fine,M.J., Flanders,S.A. *et al.* (2019) Diagnosis and treatment of adults with Community-acquired pneumonia. An official clinical practice guideline of the american thoracic society and infectious diseases society of america. *Am. J. Respir. Crit. Care Med.*, **200**, e45–e67.
18. Shi,T., Denouel,A., Tietjen,A.K., Lee,J.W., Falsey,A.R., Demont,C., Nyawanda,B.O., Cai,B., Fuentes,R., Stoszek,S.K. *et al.* (2019) Global and regional burden of hospital admissions for pneumonia in older adults: a systematic review and meta-analysis. *J. Infect. Dis.*, **222**, S570–S576.
19. Fauci,A.S., Lane,H.C. and Redfield,R.R. (2020) Covid-19 - Navigating the uncharted. *N. Engl. J. Med.*, **382**, 1268–1269.
20. Lane,H.C., Marston,H.D. and Fauci,A.S. (2016) Conducting clinical trials in outbreak settings: points to consider. *Clin. Trials*, **13**, 92–95.
21. Guan,W.J., Ni,Z.Y., Hu,Y., Liang,W.H., Ou,C.Q., He,J.X., Liu,L., Shan,H., Lei,C.L., Hui,D.S.C. *et al.* (2020) Clinical characteristics of coronavirus disease 2019 in china. *N. Engl. J. Med.*, **382**, 1708–1720.
22. Cleaver,J.O., You,D., Michaud,D.R., Pruneda,F.A., Juarez,M.M., Zhang,J., Weill,P.M., Adachi,R., Gong,L., Moghaddam,S.J. *et al.* (2014) Lung epithelial cells are essential effectors of inducible resistance to pneumonia. *Mucosal Immunol.*, **7**, 78–88.
23. Duggan,J.M., You,D., Cleaver,J.O., Larson,D.T., Garza,R.J., Guzman Pruneda,F.A., Tuvim,M.J., Zhang,J., Dickey,B.F. and Evans,S.E. (2011) Synergistic interactions of TLR2/6 and TLR9 induce a high level of resistance to lung infection in mice. *J. Immunol.*, **186**, 5916–5926.
24. Kirkpatrick,C.T., Wang,Y., Leiva Juarez,M.M., Shivshankar,P., Pantaleon Garcia,J., Plumer,A.K., Kulkarni,V.V., Ware,H.H., Gulraiz,F., Chavez Cavasos,M.A. *et al.* (2018) Inducible lung epithelial resistance requires multisource reactive oxygen species generation to protect against viral infections. *MBio*, **9**, e00696-18.
25. Ware,H.H., Kulkarni,V.V., Wang,Y., Pantaleón García,J., Leiva Juarez,M., Kirkpatrick,C.T., Wali,S., Syed,S., Kontoyiannis,A.D., Sikkema,W.K.A. *et al.* (2019) Inducible lung epithelial resistance requires multisource reactive oxygen species generation to protect against bacterial infections. *PLoS One*, **14**, e0208216.
26. Evans,S.E., Tseng,C.K., Scott,B.L., Hook,A.M. and Dickey,B.F. (2020) Inducible epithelial resistance against coronavirus pneumonia in mice. *Am. J. Respir. Cell Mol. Biol.*, **63**, 540–541.
27. Tuvim,M.J., Gilbert,B.E., Dickey,B.F. and Evans,S.E. (2012) Synergistic TLR2/6 and TLR9 activation protects mice against lethal influenza pneumonia. *PLoS One*, **7**, e30596.
28. Chen,G., Tsoi,A., Xu,H. and Zheng,W.J. (2018) Predict effective drug combination by deep belief network and ontology fingerprints. *J. Biomed. Inform.*, **85**, 149–154.
29. Hsu,Y.C., Chiu,Y.C., Chen,Y., Hsiao,T.H. and Chuang,E.Y. (2016) A simple gene set-based method accurately predicts the synergy of drug pairs. *BMC Syst. Biol.*, **10**, 66.
30. Celebi,R., Bear Don't Walk,O.T., Movva,R., Alpsoy,S. and Dumontier,M. (2019) In-silico prediction of synergistic anti-cancer drug combinations using Multi-omics data. *Sci. Rep.*, **9**, 8949.
31. Cuvitoglu,A., Zhou,J.X., Huang,S. and Isik,Z. (2019) Predicting drug synergy for precision medicine using network biology and machine learning. *J. Bioinform. Comput. Biol.*, **17**, 1950012.
32. Regan-Fendt,K.E., Xu,J., DiVincenzo,M., Duggan,M.C., Shakya,R., Na,R., Carson,W.E. III, Payne,P.R.O. and Li,F. (2019) Synergy from gene expression and network mining (SynGeNet) method predicts synergistic drug combinations for diverse melanoma genomic subtypes. *NPJ. Syst. Biol. Appl.*, **5**, 6.
33. Yang,M., Jaaks,P., Dry,J., Garnett,M., Menden,M.P. and Saez-Rodriguez,J. (2020) Stratification and prediction of drug synergy based on target functional similarity. *NPJ Syst. Biol. Appl.*, **6**, 16.
34. Ianevski,A., Giri,A.K. and Aittokallio,T. (2020) SynergyFinder 2.0: visual analytics of multi-drug combination synergies. *Nucleic Acids Res.*, **48**, W488–W493.
35. Ianevski,A., He,L., Aittokallio,T. and Tang,J. (2017) SynergyFinder: a web application for analyzing drug combination dose-response matrix data. *Bioinformatics*, **33**, 2413–2415.
36. Di Veroli,G.Y., Fornari,C., Wang,D., Mollard,S., Bramhall,J.L., Richards,F.M. and Jodrell,D.I. (2016) Combenefit: an interactive platform for the analysis and visualization of drug combinations. *Bioinformatics*, **32**, 2866–2868.
37. Li,O.T.W., Chan,M.C.W., Leung,C.S.W., Chan,R.W.Y., Guan,Y., Nicholls,J.M. and Poon,L.L.M. (2009) Full factorial analysis of mammalian and avian influenza polymerase subunits suggests a role of an efficient polymerase for virus adaptation. *PLoS One*, **4**, e5658.
38. Antony,J. (2014) In: Antony,J. (ed) *Design of Experiments for Engineers and Scientists (Second Edition)*. Elsevier, Oxford, pp. 63–85.
39. Das,A.K. and Dewanjee,S. (2018) In: Sarker,S.D. and Nahar,L. (eds). *Computational Phytochemistry*. Elsevier, pp. 75–106.
40. Tibes,R., Qiu,Y., Lu,Y., Hennessy,B., Andreeff,M., Mills,G.B. and Kornblau,S.M. (2006) Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells. *Mol. Cancer Ther.*, **5**, 2512–2521.
41. Hennessy,B.T., Lu,Y., Gonzalez-Angulo,A.M., Carey,M.S., Myhre,S., Ju,Z., Davies,M.A., Liu,W., Coombes,K., Meric-Bernstam,F. *et al.* (2010) A technical assessment of the utility of reverse phase protein arrays for the study of the functional proteome in Non-microdissected human breast cancers. *Clin. Proteomics*, **6**, 129–151.
42. Taylor,S.C., Nadeau,K., Abbasi,M., Lachance,C., Nguyen,M. and Fenrich,J. (2019) The ultimate qPCR experiment: producing publication quality, reproducible data the first time. *Trends Biotechnol.*, **37**, 761–774.
43. Slinker,B.K. (1998) The statistics of synergism. *J. Mol. Cell Cardiol.*, **30**, 723–731.
44. Foucquier,J. and Guedj,M. (2015) Analysis of drug combinations: current methodological landscape. *Pharmacol. Res. Perspect.*, **3**, e00149.
45. Fernandes,A.D., Reid,J.N., Macklaim,J.M., McMurrough,T.A., Edgell,D.R. and Gloor,G.B. (2014) Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. *Microbiome*, **2**, 15.
46. Gloor,G.B., Macklaim,J.M., Vu,M. and Fernandes,A.D. (2016) Compositional uncertainty should not be ignored in high-throughput sequencing data analysis. *Austrian J. Stat.*, **45**, 73–87.
47. Pawlowsky-Glahn,V., Buccianti,A. and Pawlowsky-Glahn,V. (2011) In: *Compositional Data Analysis: Theory and Applications*. John Wiley & Sons, Incorporated, NY, UNITED KINGDOM.
48. Lo,S. and Andrews,S. (2015) To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Front Psychol.*, **6**, 1171.
49. Silverman,J.D., Washburne,A.D., Mukherjee,S. and David,L.A. (2017) A phylogenetic transform enhances analysis of compositional microbiota data. *Elife*, **6**, e21887.

50. Abrams,Z.B., Johnson,T.S., Huang,K., Payne,P.R.O. and Coombes,K. (2019) A protocol to evaluate RNA sequencing normalization methods. *BMC Bioinform.*, **20**, 679.

51. Lin,Y., Golovnina,K., Chen,Z.X., Lee,H.N., Negron,Y.L., Sultana,H., Oliver,B. and Harbison,S.T. (2016) Comparison of normalization and differential expression analyses using RNA-Seq data from 726 individual drosophila melanogaster. *BMC Genomics*, **17**, 28.

52. Tarazona,S., Balzano-Nogueira,L., Gomez-Cabrero,D., Schmidt,A., Imhof,A., Hankemeier,T., Tegner,J., Westerhuis,J.A. and Conesa,A. (2020) Harmonization of quality metrics and power calculation in multi-omic studies. *Nat. Commun.*, **11**, 3092.

53. Quinn,T.P., Erb,I., Gloor,G., Notredame,C., Richardson,M.F. and Crowley,T.M. (2019) A field guide for the compositional analysis of any-omics data. *Gigascience*, **8**, giz107.

54. Geary,N. (2012) Understanding synergy. *Am. J. Physiol. Endocrinol. Metab.*, **304**, E237–E253.

55. Goldstein,I., Paakinaho,V., Baek,S., Sung,M.-H. and Hager,G.L. (2017) Synergistic gene expression during the acute phase response is characterized by transcription factor assisted loading. *Nat. Commun.*, **8**, 1849.

56. Caetano,M.S., Hassane,M., Van,H.T., Bugarin,E., Cumpian,A.M., McDowell,C.L., Cavazos,C.G., Zhang,H., Deng,S., Diao,L. *et al.* (2018) Sex specific function of epithelial STAT3 signaling in pathogenesis of K-ras mutant lung cancer. *Nat. Commun.*, **9**, 4589–4589.

57. Pounds,S. and Morris,S.W. (2003) Estimating the occurrence of false positives and false negatives in microarray studies by approximating and partitioning the empirical distribution of *P*-values. *Bioinformatics*, **19**, 1236–1242.

58. Ji,Y., Wu,C., Liu,P., Wang,J. and Coombes,K.R. (2005) Applications of beta-mixture models in bioinformatics. *Bioinformatics*, **21**, 2118–2122.

59. Zhang,J. and Coombes,K.R. (2012) Sources of variation in false discovery rate estimation include sample size, correlation, and inherent differences between groups. *BMC Bioinform.*, **13** (Suppl.13), S1.

60. Reimand,J., Isserlin,R., Voisin,V., Kucera,M., Tannus-Lopes,C., Rostamianfar,A., Wadi,L., Meyer,M., Wong,J., Xu,C. *et al.* (2019) Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, cytoscape and enrichmentmap. *Nat. Protoc.*, **14**, 482–517.

61. Subramanian,A., Tamayo,P., Mootha,V.K., Mukherjee,S., Ebert,B.L., Gillette,M.A., Paulovich,A., Pomeroy,S.L., Golub,T.R., Lander,E.S. *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 15545–15550.

62. Mootha,V.K., Lindgren,C.M., Eriksson,K.F., Subramanian,A., Sihag,S., Lehar,J., Puigserver,P., Carlsson,E., Ridderstrale,M., Laurila,E. *et al.* (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.*, **34**, 267–273.

63. Chen,E.Y., Tan,C.M., Kou,Y., Duan,Q., Wang,Z., Meirelles,G.V., Clark,N.R. and Ma'ayan,A. (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.*, **14**, 128.

64. Kuleshov,M.V., Jones,M.R., Rouillard,A.D., Fernandez,N.F., Duan,Q., Wang,Z., Koplev,S., Jenkins,S.L., Jagodnik,K.M., Lachmann,A. *et al.* (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.*, **44**, W90–W97.

65. Kramer,A., Green,J., Pollard,J. Jr and Tugendreich,S. (2014) Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics*, **30**, 523–530.

66. Maguire,O., O'Loughlin,K. and Minderman,H. (2015) Simultaneous assessment of NF-κB/p65 phosphorylation and nuclear localization using imaging flow cytometry. *J. Immunol. Methods*, **423**, 3–11.

67. Evans,S.E., Tuvim,M.J., Fox,C.J., Sachdev,V., Gibiansky,L. and Dickey,B.F. (2011) Inhaled innate immune ligands to prevent pneumonia. *Br. J. Pharmacol.*, **163**, 195–206.

68. Coral,C. and Bokelmann,W. (2017) The role of analytical frameworks for systemic research design, explained in the analysis of drivers and dynamics of historic land-use changes. *Systems*, **5**, 20.

69. Bardini,R., Politano,G., Benso,A. and Di Carlo,S. (2017) Multi-level and hybrid modelling approaches for systems biology. *Comput. Struct. Biotechnol. J.*, **15**, 396–402.

70. Huang,K., Liu,X., Li,Y., Wang,Q., Zhou,J., Wang,Y., Dong,F., Yang,C., Sun,Z., Fang,C. *et al.* (2019) Genome-Wide CRISPR-Cas9 screening identifies NF-kappaB/E2F6 responsible for EGFRvIII-Associated temozolomide resistance in glioblastoma. *Adv Sci (Weinh)*, **6**, 1900782.

71. Kagohara,L.T., Zamuner,F., Davis-Marcisak,E.F., Sharma,G., Considine,M., Allen,J., Yegnasubramanian,S., Gaykalova,D.A. and Fertig,E.J. (2020) Integrated single-cell and bulk gene expression and ATAC-seq reveals heterogeneity and early changes in pathways associated with resistance to cetuximab in HNSCC-sensitive cell lines. *Br. J. Cancer*, **123**, 101–113.

72. Singh,R., Peng,S., Viswanath,P., Sambandam,V., Shen,L., Rao,X., Fang,B., Wang,J. and Johnson,F.M. (2019) Non-canonical cMet regulation by vimentin mediates plk1 inhibitor–induced apoptosis. *EMBO Mol. Med.*, **11**, e9960.

73. Guillon,A., Arafa,E.I., Barker,K.A., Belkina,A.C., Martin,I., Shenoy,A.T., Wooten,A.K., Lyon De Ana,C., Dai,A., Labadorf,A. *et al.* (2020) Pneumonia recovery reprograms the alveolar macrophage pool. *JCI Insight*, **5**, e133042.

74. Van Goor,A., Ashwell,C.M., Persia,M.E., Rothschild,M.F., Schmidt,C.J. and Lamont,S.J. (2017) Unique genetic responses revealed in RNA-seq of the spleen of chickens stimulated with lipopolysaccharide and short-term heat. *PLoS One*, **12**, e0171414.

75. Blanco-Melo,D., Nilsson-Payant,B.E., Liu,W.C., Uhl,S., Hoagland,D., Moller,R., Jordan,T.X., Oishi,K., Panis,M., Sachs,D. *et al.* (2020) Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell*, **181**, 1036–1045.

76. Glover,K.P., Chen,Z., Markell,L.K. and Han,X. (2015) Synergistic gene expression signature observed in TK6 cells upon co-exposure to UVC-Irradiation and protein kinase C-Activating tumor promoters. *PLoS One*, **10**, e0139850.

77. Maifiah,M.H., Creek,D.J., Nation,R.L., Forrest,A., Tsuji,B.T., Velkov,T. and Li,J. (2017) Untargeted metabolomics analysis reveals key pathways responsible for the synergistic killing of colistin and doripenem combination against acinetobacter baumannii. *Sci. Rep.*, **7**, 45527.

78. Han,M.L., Liu,X., Velkov,T., Lin,Y.W., Zhu,Y., Li,M., Yu,H.H., Zhou,Z., Creek,D.J., Zhang,J. *et al.* (2018) Metabolic analyses revealed time-dependent synergistic killing by colistin and aztreonam combination against multidrug-resistant acinetobacter baumannii. *Front Microbiol*, **9**, 2776.

79. Malik,A.A., Swenson,T., Weihe,C., Morrison,E.W., Martiny,J.B.H., Brodie,E.L., Northen,T.R. and Allison,S.D. (2020) Drought and plant litter chemistry alter microbial gene expression and metabolite production. *ISME J.*, **14**, 2236–2247.

80. Burberry,A., Wells,M.F., Limone,F., Couto,A., Smith,K.S., Keaney,J., Gillet,G., van Gastel,N., Wang,J.Y., Pietilainen,O. *et al.* (2020) C9orf72 suppresses systemic and neural inflammation induced by gut bacteria. *Nature*, **582**, 89–94.

81. Li,C., Shen,W., Shen,S. and Ai,Z. (2013) Gene expression patterns combined with bioinformatics analysis identify genes associated with cholangiocarcinoma. *Comput. Biol. Chem.*, **47**, 192–197.

82. Mihret,A., Loxton,A.G., Bekele,Y., Kaufmann,S.H.E., Kidd,M., Haks,M.C., Ottenhoff,T.H.M., Aseffa,A., Howe,R. and Walzl,G. (2014) Combination of gene expression patterns in whole blood discriminate between tuberculosis infection states. *BMC Infect. Dis.*, **14**, 257.

83. Zhang,X., Cha,I.-H. and Kim,K.-Y. (2017) Use of a combined gene expression profile in implementing a drug sensitivity predictive model for breast cancer. *Cancer Res. Treat.*, **49**, 116–128.

84. Hassall,K.L. and Mead,A. (2018) Beyond the one-way ANOVA for 'omics data. *BMC Bioinform.*, **19**, 199.

85. Mee,R.W. (2009) In: Mee,R. (ed). *A Comprehensive Guide to Factorial Two-Level Experimentation*. SpringerNY, pp. 27–74.

86. Van den Berge,K., Roux de Bezieux,H., Street,K., Saelens,W., Cannoodt,R., Saeys,Y., Dudoit,S. and Clement,L. (2020) Trajectory-based differential expression analysis for single-cell sequencing data. *Nat. Commun.*, **11**, 1201.

87. Ren,X. and Kuan,P.F. (2020) Negative binomial additive model for RNA-Seq data analysis. *BMC Bioinform.*, **21**, 171.

88. Chou,T.C. and Talalay,P. (1984) Quantitative analysis of dose-effect relationships: the combined effects of multiple drugs or enzyme inhibitors. *Adv. Enzyme Regul.*, **22**, 27–55.

89. Yadav,B., Wennerberg,K., Aittokallio,T. and Tang,J. (2015) Searching for drug synergy in complex dose-response landscapes using an interaction potency model. *Comput. Struct. Biotechnol. J.*, **13**, 504–513.

90. Chou,T.C. (2010) Drug combination studies and their synergy quantification using the chou-talalay method. *Cancer Res.*, **70**, 440–446.

91. Xie,C. and Tammi,M.T. (2009) CNV-seq, a new method to detect copy number variation using high-throughput sequencing. *BMC Bioinform.*, **10**, 80.