SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES

» Cell biology

» Molecular evolution

» Research data

» Gene regulatory networks

The species translation challenge—A systems biology perspective on human and rat bronchial epithelial cells

Carine Poussin[1,5], Carole Mathis[1,5], Leonidas G. Alexopoulos[2,3,5], Dimitris E. Messinis[2,3], Rémi H.J. Dulize[1], Vincenzo Belcastro[1], Ioannis N. Melas[2,3], Theodore Sakellaropoulos[3], Kahn Rhrissorrakrai[4], Erhan Bilal[4], Pablo Meyer[4], Marja Talikka[1], Stéphanie Boué[1], Raquel Norel[4], John J. Rice[4], Gustavo Stolovitzky[4], Nikolai V. Ivanov[1], Manuel C. Peitsch[1] and Julia Hoeng[1]

Received: 20 December 2013
Accepted: 25 April 2014
Published: 10 June 2014

The biological responses to external cues such as drugs, chemicals, viruses and hormones, is an essential question in biomedicine and in the field of toxicology, and cannot be easily studied in humans. Thus, biomedical research has continuously relied on animal models for studying the impact of these compounds and attempted to 'translate' the results to humans. In this context, the SBV IMPROVER (Systems Biology Verification for Industrial Methodology for PROcess VErification in Research) collaborative initiative, which uses crowd-sourcing techniques to address fundamental questions in systems biology, invited scientists to deploy their own computational methodologies to make predictions on species translatability. A multi-layer systems biology dataset was generated that was comprised of phosphoproteomics, transcriptomics and cytokine data derived from normal human (NHBE) and rat (NRBE) bronchial epithelial cells exposed in parallel to more than 50 different stimuli under identical conditions. The present manuscript describes in detail the experimental settings, generation, processing and quality control analysis of the multi-layer omics dataset accessible in public repositories for further intra- and inter-species translation studies.

Design Type(s)	*in vitro* design • compound treatment design • cell type comparison design
Measurement Type(s)	transcription profiling assay • protein expression profiling
Technology Type(s)	DNA microarray • sandwich ELISA
Factor Type(s)	compound • dose • organism
Sample Characteristic(s)	Homo sapiens • Rattus norvegicus • bronchus • bronchial epithelial cell line

[1]Philip Morris International R&D, Philip Morris Products S. A., Quai Jeanrenaud 5, 2000 Neuchâtel, Switzerland. [2]ProtATonce Ltd, Scientific Park Lefkippos, Patriarchou Grigoriou & Neapoleos, 15343 Ag. Paraskevi, Attiki, Greece. [3]National Technical University of Athens, Heroon Polytechniou 9, Zografou 15780, Greece. [4]IBM Computational Biology Center, Yorktown Heights, NY 10598, USA. [5]These authors contributed equally to this work.. Correspondence and requests for materials should be addressed to C.P. (email: carine.poussin@pmi.com)

Background & Summary

Animal models have been used intensively to understand biological mechanisms associated with diseases and to unravel toxic effects of drugs or environmental agents. Biological processes in mice or rats have been generally assumed to reflect biological processes in humans under analogous conditions. A natural question in this context is the degree to which biological perturbations observed in rodents can be translated to humans. Such knowledge is important since it can reduce uncertainties in species extrapolations (Fig. 1).

The Systems Biology Verification for Industrial Methodology for Process Verification in Research (SBV IMPROVER) initiative[1,2] (https://www.sbvimprover.com/) opened a challenge called Species Translation Challenge (STC) to the scientific community to identify compound-specific biological mechanisms of actions (MoA) that are common to different species, in this case, humans and rats. The challenge consisted of four sub-challenges whereby the interspecies pathway perturbation prediction challenge sought to explore whether responsive gene sets and related processes in humans can be inferred based upon the corresponding data in rats.

To address the question of species translatability at different molecular layers of the biological system in the context of STC, an experiment was designed to generate human and rat multi-layer datasets consisting of phosphoproteomics, transcriptomics and cytokine level measurements. To ensure that the generated datasets were comparable and that the proof of concept predictions across species was valid, experiments with well-controlled conditions were designed and conducted using an *in vitro* system. This chemical testing strategy is aligned with the effort to 'Replace, Reduce, Refine' animal experiments (the '3R' approach) (http://ihcp.jrc.ec.europa.eu/our_activities/alt-animal-testing-safety-assessment-chemicals/alternative-testing-strategies-progress-report-2009.-replacing-reducing-and-refining-use-of-animals-in-research) and to use more appropriate cell-based assays that have the potential to provide more relevant data on the effects of short- and long-term exposure to toxicants. Primary normal human bronchial epithelial cells (NHBE) and primary normal rat bronchial epithelial cells (NRBE) were exposed in parallel to various types of stimuli, which were selected ensuring a broad perturbation spectrum of the cellular system, under identical experimental conditions (duration of exposure, concentration of stimuli and cell culture parameters).

The challenge aimed to investigate whether the phosphorylation signals could be inferred from gene expression data within species (reverse engineering) and the translatability of phosphorylation signals across species, and also to better understand the level at which translation across species is more robust (e.g., individual molecules, predefined gene sets representative of canonical pathways or higher-order processes). These questions have been articulated around four sub-challenges proposed to the scientific community (https://www.sbvimprover.com/challenge-2/challenge-2-challenge). The second SBV IMPROVER symposium was held in Greece at the end of October 2013 to announce the results of the Species Translation Challenge and to discuss the topic extensively with all participants (http://www.bio-itworld.com/2013/11/8/sometimes-you-can-trust-rat.html; http://www.genomeweb.

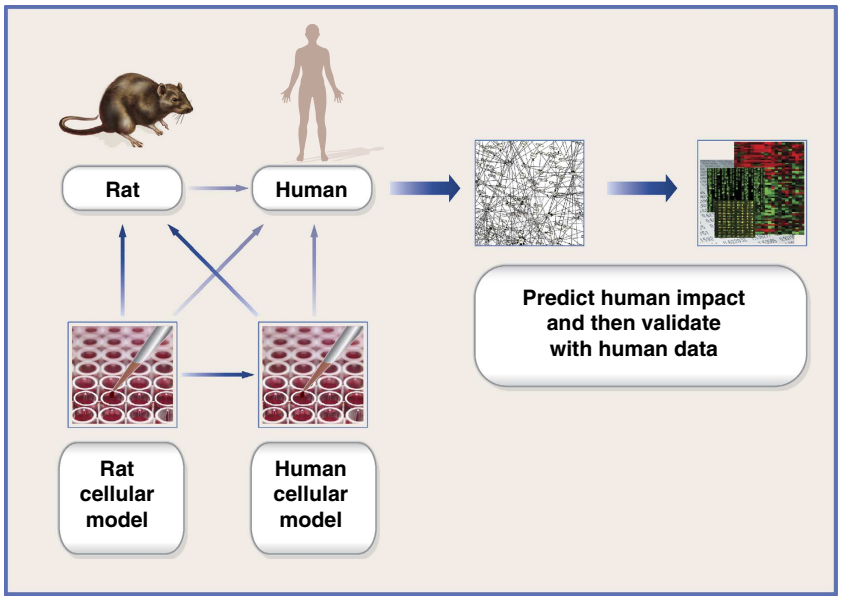

Figure 1. Concept of translatability. The arrows indicate potential routes of translation between *in vitro* and *in vivo* systems and/or across species.

com/informatics/improver-species-translation-challenge-results-released; http://www.americanlaboratory.com/913-Technical-Articles/149618-Results-are-in-for-the-Second-sbv-IMPROVER-Challenge-on-Species-Translation/).

The present manuscript describes the experimental design, optimization steps and data quality checks necessary to generate a multi-layer systems biology data compendium suitable for computational crowdsourcing challenges such as the Species Translation Challenge. The experimental settings and protocols as well as the generation, processing and quality control analysis of the raw data are detailed. The raw data (168 and 164 CEL files for human and rat respectively) and processed data (e.g., normalized gene expression data) are freely available in public repositories such as ArrayExpress for transcriptomics data (Data Citation 1). Human and rat proteomics data are deposited in the figshare public repository (Supplementary Table 1) (Data Citation 2).

The unique multi-omics dataset presented in this manuscript is of great value for the computational community to develop new modelling capabilities to address the important topic of species translatability at different molecular levels of the human and rat bronchial epithelial cellular system. A better understanding of the range of applicability of the translation concept will impact the predictability of signaling responses, mode of action and efficacy of drugs in the field of systems pharmacology as well as increase the confidence in the estimation of human risk from rodent data in the context of toxicological risk assessment. It provides a unique translational compendium with applicability in systems biology and toxicology, fully aligned with the Tox21 initiatives[3].

Methods

Cell culture

NHBE cells were purchased from Lonza (Catalog number CC-2540, Lonza Inc., Switzerland). These cells, obtained from different Caucasian, disease-free and non-smoker donors, were isolated from airway (tracheal/bronchial) epithelial tissue located above the bifurcation of the lungs. NRBE cells were purchased from CHI Scientific Inc. (Catalog number 4-61391, Maynard, Maryland, USA) and isolated from pooled tracheobronchial tissue of adult inbred AGA rats. The stocks of NHBE and NRBE cells were stored in liquid nitrogen with 10% (v/v) dimethylsulfoxyde (DMSO). Vials of stock cells were rapidly thawed and diluted in 20 ml of bronchial epithelial cell growth medium with supplements (Lonza, BulletKit CC-3170). Both cell types were seeded in flasks (T75) coated with rat tail collagen type I from BD (catalog number: 354236) and grown in the same growth medium with supplements at 37.0 ± 1 °C in a humidified incubator with $5.0 \pm 0.5\%$ CO_2 in air. After 24 h, the medium was changed and cells were regularly checked during proliferation using a microscope. Once reaching confluence, cells were split into subcultures. Briefly, cells were washed with HEPES Buffered Saline Solution, then trypsinized with Trypsin/EDTA that was neutralized using a Trypsin Neutralizing Solution (TNS) (The 3 solutions are included in Clonetics™ ReagentPack™ from Lonza; catalog number CC-5034). Cells were expanded for 10 days (including 1 split) to reach the final number needed for screening or main experiments. Cells were seeded into pre-coated rat tail collagen type I 96-well plates (BD BioCoat™, catalog number: 356649) testing different cell densities ranging from 2,500 to 50,000 cells per well (in 100 µl). The range of optimal seeding densities was determined by microscopic inspection to be 25,000–35,000 cells/well (80–90% confluence). However, optimal yield of RNA extraction used for transcriptomics analysis was obtained with 50,000 cells/well corresponding to 100% confluency. For the screening of the main experiment, cells were re-suspended in bronchial epithelial cell growth medium with supplements, and seeded at a cell density of 50,000 cells/well in pre-coated rat tail collagen type I 96-well plates (BD BioCoat™, catalog number: 356649). After 24 h, NHBE and NRBE were treated in parallel with selected stimuli or DME.

Systems biology data generation

Due to the high number of stimuli and experimental conditions described above, the main phase was conducted in two experiments (40 stimuli used for the experiment 1 and 12 stimuli for the experiment 2) to generate all samples required to produce the entire systems biology dataset.

Measurements of phosphoproteomics and cytokines using xMAP beads. For phosphoproteomics measurements, NRBE and NHBE cell cultures were removed from the incubator and placed on ice. The cells were washed with 100 µl of ice-cold phosphate-buffered saline (PBS) and cells were lysed using 60 µl of Tris-HCL supplemented with inhibitors of proteases and phosphatases in a 96-well plate on ice for 20 min. The plates were incubated overnight at −20 °C and then rapidly thawed in a 37 °C water bath for 2 min, followed by sonication. Cell debris was removed following a centrifugation at $2700 \times g$ for 20 min at 4 °C. For cytokines measurements, cell supernatants were collected 24 h post-treatment. For the bead-based enzyme-linked immunosorbent assay (ELISA) procedure, 50 µl of non-diluted cell lysates or supernatants were incubated with the xMAP beads (4000 beads/well for each protein) for 1.5 h to capture target proteins with specific antibodies coupled to the beads. The beads were washed twice with 100 µl of PBS. The beads were then incubated for 1.5 h with 20 µl of detection antibodies targeting different epitopes than the bead-coupled capture antibodies (average concentration, 1 µg/ml) followed by washing steps. Subsequently, 50 µl of streptavidin-phycoerythrin (PE) (at a final concentration of 5 µg/µl) were

added and the mixture was incubated for 20 min. The beads were then washed and re-suspended in 130 µl of PBS-bovine serum albumin (BSA) assay buffer.

Transcriptomics. Total RNA was isolated from NHBE and NRBE cells using the QIAGEN RNeasy 96 Kit (Catalog number 74181). For each sample, isolated RNA was quantified using the Nanodrop 1000 Spectrophotometer (Thermo Scientific) and quality checked using the Agilent 2100 Bioanalyzer. Twenty nanograms of total RNA were reverse-transcribed into cDNA and amplified using the NuGEN™ Ovation™ RNA Amplification System V2 (Catalog number 3100-A01). The cDNA was then purified using magnetic beads (Agencourt RNAClean XP, Catalog number A63987 from Beckman Coulter GmbH, Krefeld, Germany) to remove unincorporated nucleotide triphosphates, salts, enzymes and inorganic phosphates. Purified cDNA was quantified, quality checked and fragmented (at least 3.75 µg of cDNA is needed) with a combined chemical and enzymatic reaction, and finally labeled using enzymatic attachment of nucleotides coupled to biotin. Fifty microliters of fragmented and labeled cDNA were added to 170 µl of Master Mix Hybridization Cocktail Assembly (Affymetrix GeneChip® Hybridization, Wash, and Stain Kit; Catalog number 900720). After denaturation reaction (2 min at 99 °C and 5 min at 45 °C) followed by centrifugation at Vmax for 1 min, 200 µl of the cDNA cocktail were hybridized on Affymetrix® HG-U133 Plus2 or Rat 230 2.0 GeneChips. The arrays were incubated in the GeneChip® Hybridization Oven 645 (Catalog number 00-0331) for 18 ± 2 h at 45 °C with a rotation speed of 60 rpm. After the hybridization step, the arrays were washed and stained on a Fluidics Station FS450 (Catalog number 00-0335) using Affymetrix® GeneChip Command Console™ Software (AGCC software version 3.2) with protocol FS450_0004. Finally, the arrays were scanned using the GeneChip® Scanner 3000 7 G (Catalog number 00-0210).

Raw images from the scanner were saved as DAT files. Using the AGCC Viewer software application, each image was checked for artifacts, overall intensity distribution, checkerboards at the corners, a central cross to ensure adequate grid alignment and readability of the array name. The AGCC Viewer software automatically gridded the DAT file image and extracted probe cell intensities into a CEL file. The CEL files were further processed (MAS5.0) with Affymetrix® Expression Console™ software (version Build 1.3.1.187) for a first quality check of the data. Materials and reagent kits were purchased from Affymetrix, Inc. (Santa Clara, CA, USA), NuGen (San Carlos, CA, USA) and QIAGEN GmbH (Hilden, Germany).

Due to the high number of samples collected for experiment 1, it was not possible to process all samples at once. Therefore, mRNA samples were processed in three batches (samples were randomized within each batch). Each batch contained human and rat mRNAs (in triplicate) for a subset of randomly selected stimuli among those tested (Fig. 2). The same DME control mRNA samples (four replicates) were re-hybridized for each batch. For experiment 2, all mRNA samples were processed together at a single point in time. This included the DME control mRNA samples (four replicates) obtained for this second experiment (Fig. 2).

Data Records

Phosphoproteomics and cytokine data

Raw data processing, normalization and active signals analysis. Phosphoproteomics and cytokine release data were measured using xMAP technology on a Luminex FlexMAP3D® system and the software used was the Luminex xPONENT® for FLEXMAP3D®, Version 4.2. Custom software was developed to analyse the raw data following the standard LXB format data extraction (http://cran.r-project.org/web/packages/lxb/README.html). Following data acquisition, the raw measurements corresponding to the fluorescence intensity of each bead for each individual analyte (protein), were exported. At least 100 events (counts) were measured for each analyte. The median statistic (median fluorescence intensity, MFI) less sensitive to outliers was chosen to summarize data as a representative value of the protein measurements upon Luminex recommendations. To remove the effects of non-specific binding of proteins to beads in lysates, negative control 'naked' beads (BSA-coated beads devoid of antibody that corresponds to Control B) were prepared using standard coupling procedures. Phycoerythrin-coated beads were also prepared and used as positive control (Control A). Both positive and negative control beads were mixed with the other beads in the multiplex assay. The signal intensities of the negative control beads were found to positively correlate with the signal intensities of the phosphoproteins, which were corrected using a robust linear regression on all replicates[4]. The dependent variable was the signal intensity of a phosphoprotein across stimuli and DME controls (including replicates), and the independent variable was the signal intensity of the 'naked' bead (robust Tukey biweight regressions were calculated with data from experiments 1 and 2, independently). The final normalized signal intensity values for phosphoproteins were taken as the ratio between the residuals and the Root Mean Square Error (RMSE) that resulted from the regression fit. The cytokine data corresponded to the median of the distribution of bead signal intensities measured for each protein in all supernatant samples. In the context of the supernatant, the chance of non-specific binding was reduced when compared to the cell lysate context. Therefore, it was not necessary, to use 'naked' beads (control B) to control for this effect. The median signal intensity values were normalized by calculating z-scores for each cytokine across all stimuli including DME controls. This score was independently calculated for experiment 1 and 2 by taking the ratio of difference between the signal and the mean as well as and the standard deviation calculated for

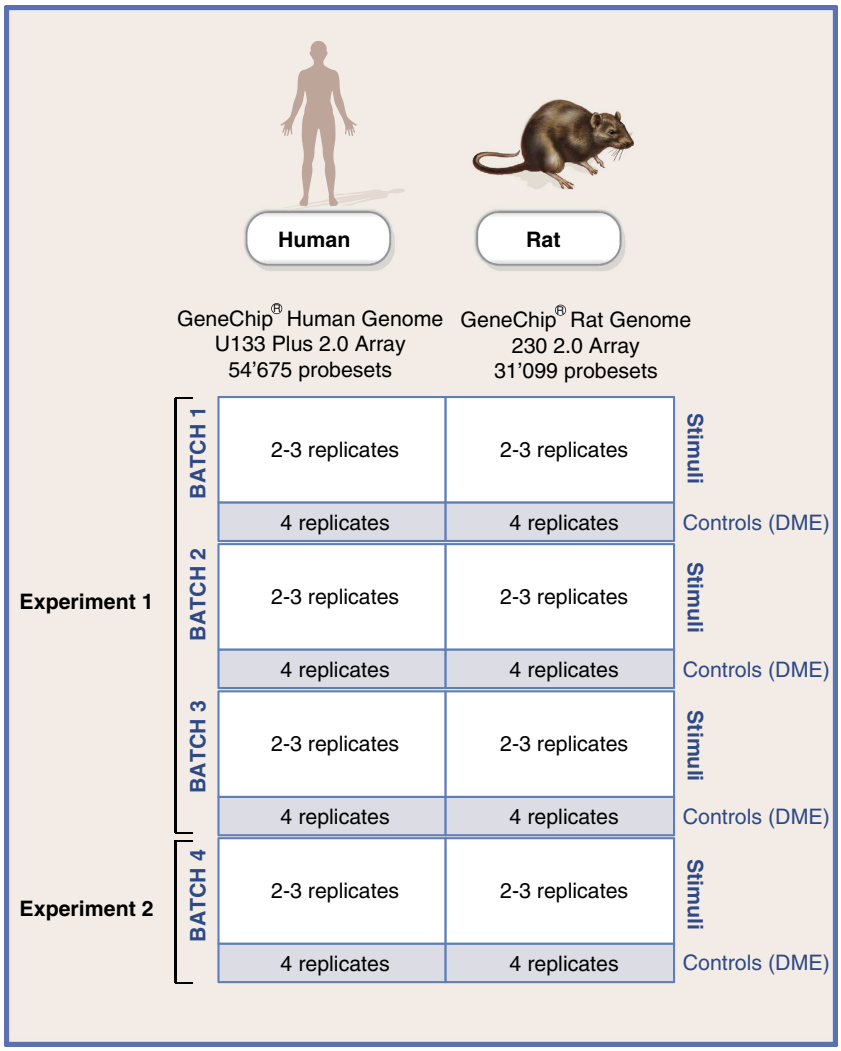

Figure 2. Schema of the mRNA processing to generate the gene expression dataset avoiding confounding effects between species and between DME controls and treatment conditions.

one cytokine across stimuli. For phosphoproteins and cytokines, normalized values beyond 3 standard deviations were considered as active signals.

Quality control analysis. Both individual signals and clusters of signals were examined to achieve the highest quality of the datasets. For each analyte, the final reported signal value was the median of the distribution of individual bead counts because it is less sensitive to outliers and distribution skewness. The minimum number of beads that should be counted for each analyte was also an important parameter to ensure robustness and reliability of the reported median. The effect of the minimum number of beads required to detect a robust signal was investigated by bootstrapping analysis. The analysis showed that the bead count could greatly affect the reported median, particularly for those with low protein concentrations. Therefore, to further increase robustness, the minimum bead count was increased from 25–50 beads to 100 beads. Furthermore, the distribution of raw (bead signal) measurements for each analyte was examined to evaluate skewness and bi-modality. If the distribution of a bead signal was significantly distorted, the analyte was excluded from the dataset. Finally, to evaluate the precision as well as the robustness of the dataset, each measurement was performed in triplicate while the measurement of the control state (basal level—no treatment) that is crucial in determining the fold increase of the signal from the basal level, was performed in six-plicate. The variability of the replicates of each signal (expressed as median coefficient of variation (CV) across all conditions) served as an estimate of the measurement precision (Supplementary Figure 1). RPS6 was excluded for further analysis due to high median CV (Supplementary Figure 1).

Data storage. The data are reported as the median of bead signal intensities for each phosphoprotein or cytokine of the panel that have been measured in each sample. For each stimulus, at least 3 sample

replicates have been measured for the main experimental phase (Supplementary Table 1). The DME control included 5 to 6 sample replicates depending on the experiment. For the screening phase, a single sample was measured for each phosphoprotein and stimulus (Supplementary Table 2). Supplementary Table 1 (Results of proteomics data including phosphoproteomics and cytokine level measurements for NHBE and NRBE cells exposed to 52 stimuli) and 2 (Results of phosphoprotein measurements in NHBE and NRBE using antibody-bead based assays for the experimental screening of 270 stimuli) have been deposited in the figshare public repository (Data Citation 2).

Gene expression data

Raw data processing and normalization. For each species, all CEL files were processed and normalized together using GC robust multiarray averaging (GCRMA)[5,6]. The data were processed using the GCRMA R package (v2.32) from Bioconductor.

Quality control analysis and differential gene expression analysis. The quality of the chip was assessed at the probe- and probeset-levels by generating different diagnostic plots (chip images, probe-signal intensity distribution, pseudo-images, NUSE (Normalized Unscaled Standard Error) and RLE (Relative Log Expression) plots, correlation matrix) (Supplementary Figure 2). Chips exceeding a NUSE median value of 1.05 were considered to be outliers and excluded. Remaining CEL files were re-normalized together per species using GCRMA. The Principal Component Analysis (PCA) of normalized expression data revealed batch effects in both the human and rat gene expression datasets, which were expected, because the samples were processed as distinct batches (Supplementary Figure 2i and j). No batch correction was done. Instead, batches were treated separately in all analyses. This was made possible by the presence of corresponding 'DME control' samples (at least four replicates) within each batch. Differentially expressed genes were identified by comparing normalized data from DME control with data from each stimulus using limma R-package from Bioconductor. Provided as Supplementary Figure 3 and 4, volcano plots indicate the magnitude and the confidence of gene expression regulation for each stimulus (relative to DME control) in human and rat cells, respectively.

Data storage. The raw (CEL. files) and processed (matrices of human and rat gene expression normalized separately; values correspond to $\log_2$ expression) gene expression data (Data Citation 1) have been submitted to ArrayExpress database (http://www.ebi.ac.uk/arrayexpress/), and are available with the accession number E-MTAB-2091. Metadata were stored in a MAGE-TAB file (SDRF and IDF tabs) supportive of MIAME format for microarray data[7].

Technical Validation

The large multi-omics dataset was generated from *in vitro* cultures to feasibly test a large number of stimuli that were needed to perturb various biological pathways under controlled conditions in both human and rat systems.

Immortalized cell lines have been used in the scientific community for decades in different cellular assays due to their commercial availability at very affordable prices and the ease to culture them. While immortalized cell lines often originate from primary cells/tissues, they have gone through significant mutations, leading to genotypic and phenotypic drifting and eventually to the loss of tissue specific function. In a systems biology perspective, genetic and phenotypic modifications of cell lines have an impact on genome-wide expression profiles and probably also on other large-scale omics approaches and thus could bias data aimed to understand how cellular responses may translate from one species to another[8–12]. For example, it has been shown that the expression profile of primary airway epithelial cells and immortalized cells were different since their expression profiles did not group together using an unsupervised hierarchical clustering approach[13]. Therefore, despite the wide use of immortalized cells, it was decided to work with primary cells that constitute more suitable *in vitro* models to mimic *in vivo* behaviour.

Bronchial epithelial cells were selected as the cell system used for our experiments. The choice for these primary cell types was driven by the fact that these primary cells are at the critical interface between the body and the external environment and were commercially available in both species.

The detailed experimental workflow is described in Fig. 3a starting from the optimization phase to the execution of the main experiment that generated the final datasets for the challenge. The experimental workflow involves various steps, including (i) optimization of the cell culture and experimental conditions; (ii) validation of protein assays (Fig. 3b); (iii) identification, screening and selection of stimuli; and (iv) generation, processing and quality control of omics data.

I-Optimization experiments

Optimization of the cell culture and experimental conditions. Adaptation and optimization of the cell culture conditions originally provided by the vendor for both NHBE and NRBE cells were conducted to avoid spurious differences not associated with the origin of the cells as described in detail in the 'Methods' section.

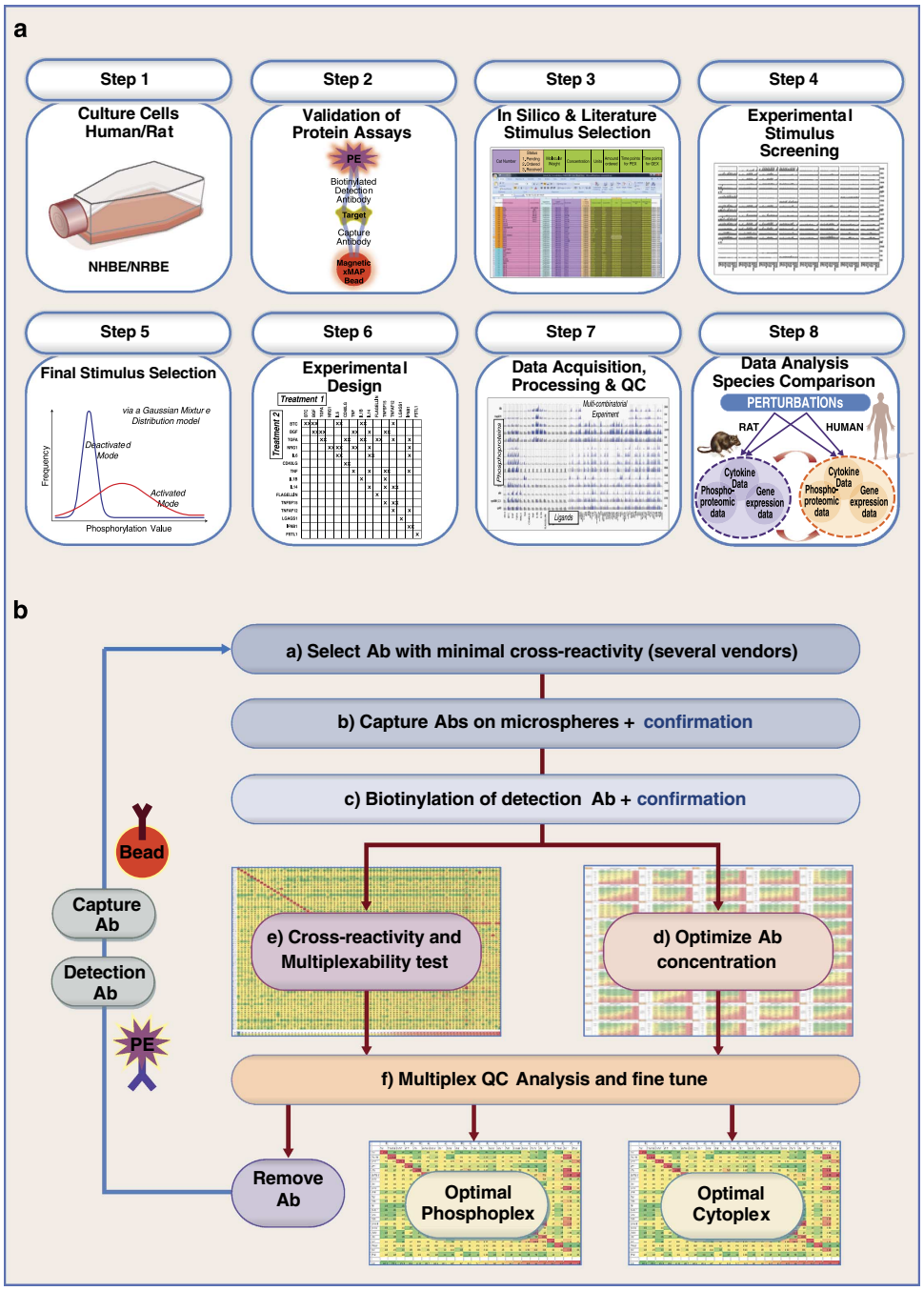

Figure 3. Overall experimental workflow. (**a**)- Experimental steps followed to generate the STC multi-layer omics dataset compendium for translational systems biology. (**b**)- Pipeline for the development and optimization of antibody-based multiplexed assays (detailed description of step 2 'Validation of protein assays').

Bead assay optimization for phosphoprotein and cytokine measurements. Using Luminex's xMAP technology (Luminex Corp, Austin, TX, USA) and ProtATonce multiplex assay optimization (ProtATonce, Athens, Greece), sandwich antibody multiplex assays were employed for the acquisition of both phosphoproteomics and cytokine data. Distinct sets of colour-coded beads with a unique colour-ID formed the solid support for antibody coupling to enable the binding of specific sample proteins on the beads (Supplementary Table 3). A biotinylated detection antibody and a streptavidin-reporter dye (phycoerythrin) completed the sandwich enzyme-linked immunosorbent assay (ELISA). The Luminex analyser works as a fluorescence-activated-cell-sorting instrument that simultaneously measures the intensity of the reporter dye and identifies the colour-ID of the bead. The xMAP technology enables the

measurement of up to 500 different analytes in a single sample but antibody cross-reactivity limits the simultaneous measurements of analytes to a few dozen. Because the quality of the data is dependent on the quality of antibodies with minimum cross-reactivity (high specificity and minimal background noise), a large number of antibodies from several vendors was purchased and first validated for the xMAP technology according to a six-step process: 1) antibodies were screened to identify optimal antibody pairs, 2) antibodies were captured on different colour-coded magnetic microspheres and the capturing efficiency was confirmed, 3) detection antibodies were biotinylated, and successful biotinylation was confirmed, 4) the concentrations of the capture and detection antibodies were optimized, 5) multiplexability issues caused by the cross-reactivity of antibodies were identified, 6) the assays were optimized to specific sample requirements (Fig. 3b).

The quality of the antibodies was validated by performing cross-reactivity (assess the specificity of an antibody) experiments in which single purified recombinant proteins were measured using the whole panel of beads (http://www.protatonce.com/#!assay-development/czoq). The antibody selection was based on an optimization algorithm that selected the maximum number of best-performing antibodies and retained the largest possible multiplexability without compromising the signal-to-noise ratio (SNR) (calculated as the ratio between the signal measured for a single purified recombinant protein and the average of signals measured in wells that do not contain the recombinant protein corresponding to background signal). Every antibody was tested against every possible antigen/antibody substrate to create a large matrix representing the specificity of each antibody to each substrate. An optimization problem was formulated to identify antibody pairs with the lowest possible off-target specificity. So if $x_i \in \{0, 1\}$ is the decision whether to include antibody i in the final assay and $C_{i,j}$ is the specificity of antibody i for substrate j, then the problem to solve is: $\min_x \sum_{i,j} x_i x_j C_{i,j}$.The problem was bound to yield a multiplex assay of size N ($\sum_i x_i = N$) and iteratively solve the problem for every N. Finally, the largest multiplex assay that yielded an acceptable background signaling level was selected. The antibodies selected were then tested for their sensitivity to their target protein and those that gave large signal to noise ratios were selected for the final experiments.

This procedure was only possible for phosphoproteomics experiments for which recombinant phosphorylated proteins were available. An alternative solution was to use cell lysates generated from cells exposed to prototypical stimuli known to modulate the phosphorylation of measured proteins. Signal-to-noise ratio (SNR) was used as an assay quality indicator. SNR values, which ranged from 5 to 1700 (no unit), strongly depend on the affinity and concentration of capture and detection antibodies. Multiplexed assays were optimized for the bronchial epithelial cell lysates and supernatants. When low signals were obtained, the concentration of detection antibody was increased to compensate for the low signals. In the absence of signal for all treatments and conditions, the assay was removed. In total, 41 different multiplexed assays for phosphoproteomics were evaluated and 19 met the criteria described above and were further used in the main experiment (Table 1). Eighty cytokine assays were evaluated out of which 22 cytokine assays were selected for the main experiment (Table 2).

Phosphoproteomics and cytokine assay variability assessment for NHBE cells

Potential sources of variability when measuring both phosphoproteins and cytokines were investigated including the inter-donor variability (for human derived primary cells) as well as different factors contributing to the technical variability[14].

Inter-donor variability. Inter-donor variability was investigated in NHBE cells from four different donors under the same conditions. These cells were stimulated with human TNF-alpha (100 ng/ml) for 20 min to measure phosphoprotein HSP27 levels or PolyI:C (10 μg/ml) and for 4 h to measure the secretion of CXCL10 protein ($n = 8$ wells of treated cells per donor and $n = 4$ wells of untreated cells per donor). The experiment demonstrated that the donor-to-donor variability was low. Coefficients of variation equal to 11 and 24% for HSP27 and CXCL10, respectively, were calculated together with the mean and standard deviation of the signal values measured in lysates of NHBE cells from the four donors. The main source of variability originated from one particular donor with systematic lower signals for HSP27 and CXCL10. NHBE cells from two donors, which gave similar results for HSP27 and CXCL10, were pooled 1+1 to obtain sufficient cells for the main phase.

Technical variability. As variability may arise from pipetting errors during cell plating, cell lysis or various ELISA technique steps, for each donor, samples from eight different wells were processed identically in parallel. The measurement of the phosphorylated HSP27 and CXCL10 level in these samples were used to quantify well-to-well variability from the same donor. The coefficients of variation for HSP27 measurement after TNF-alpha treatment ranged from 6 to 33%. For CXCL10 measurement, the coefficients of variation ranged from 7 to 14%.

To assess the variability of the ELISA technique, HSP27 was measured in 12 samples derived from a mixed pool of NHBE cell lysates that were prepared from cells treated with human TNF-alpha (100 ng/ml) for 20 min. A coefficient of variation equal to 7% was observed.

Reading the same well several times lowered the fluorescence intensity due to a photo-bleaching effect which prevented to determine the instrument variability.

Overall, the technical variability resulting from sample handling and the methods used to determine phosphoprotein and cytokine levels was lower than biological variability, ensuring robustness of data.

Common name (Target)	Residue	Uniprot ID Human	Uniprot ID Rat	Protein name
AKT1	S473	P31749	P47196	RAC-alpha serine/threonine-protein kinase
CREB1	S133	P16220	P15337	Cyclic AMP-responsive element-binding protein 1
EGFR	Y1068	P00533	Q9WTS1	Epidermal growth factor receptor
ERK1 (MAPK3)	T202/Y204	P27361	P21708	Mitogen-activated protein kinase 3
FAK1	Y397	Q05397	O35346 (FADK 1)	Focal adhesion kinase 1
GSK3B	S21/S9	P49841	P18266 (GSK-3 beta)	Glycogen synthase kinase-3 beta
HSP27 (HspB1)	S78	P04792	P42930 (HspB1)	Heat shock protein beta-1
IKBA	S32/S36	P25963	Q63746 (IkB-alpha)	NF-kappa-B inhibitor alpha
JNK2 (MAPK9)	T183/Y185	P45984	P49186 (MAPK 9)	Mitogen-activated protein kinase 9
MEK1 (MAPKK1)	S217/S221	Q02750	Q01986 (MAPKK 1)	Dual specificity mitogen-activated protein kinase kinase 1
MKK6 (MAPKK6)	S207/T211	P52564	Q925D6 (-)	Dual specificity mitogen-activated protein kinase kinase 6
NFKB	S536	Q04206	O88619 (-)	Transcription factor p65
p38MAPK	T180/Y182	Q16539 (MAPK 14)/ Q15759 (MAPK 11)	P70618 (MAPK 14)/ (MAPK11)	Mitogen-activated protein kinase 14/11
P53	S46	P04637	P10361	Cellular tumor antigen p53
RPS6KB1 (P70S6K, S6K1)	T421/S424	P23443	P67999	Ribosomal protein S6 kinase beta-1
RPS6	S235/S236	P62753	P62755	40S ribosomal protein S6
RPS6KA1 (RSK1)	S380	Q15418	Q63531	Ribosomal protein S6 kinase alpha-1
SHP2	Y542	Q06124	P41499	Tyrosine-protein phosphatase non-receptor type 11
WNK1	T60	Q9H4A3	Q9JIH7	Serine/threonine-protein kinase WNK1

Table 1. Final phosphoproteomics assay panel.

Optimization of the experimental design for the phosphoproteomics measurements

To capture the highest number of protein phosphorylation events linked to pathways perturbed upon stimulus exposure, it was essential to determine the optimal time points to harvest both cell types for the main experimental phase. To select these time points, NHBE and NRBE cells were cultured in parallel under the optimized conditions as previously determined followed by exposure to seven prototypical stimuli: TNF-alpha, TGF-alpha, insulin, IL-6, IL1-alpha, IL1-beta, IFN-gamma and a vehicle control (Dulbecco's Modified Eagle's Medium; DME). The concentration of each stimulus was chose based on literature review as described in Step 2 of the study workflow (Fig. 3a). Five different time points were selected (0, 5, 15, 20 and 25 min) to measure 41 different phosphoproteins that were plotted using a modified version of DataRail[15] (Fig. 4). For each of the time point, the fold increase of each signal was calculated as compared to the basal level. The two time points (5 and 25 min) with the maximum number of activated signals and the largest fold increase for both cell types were selected as the optimal time points for phosphoproteomics measurements in the main experiment. The selection of 5 and 25 min was also consistent with earlier studies done with hepatocytes[16].

II-Stimulus selection process

The goal of the Species Translation challenge was to understand and provide insight on how far the translation concept can be applied between rodents and humans at various layers of biological molecules

Target	Uniprot ID Human	Uniprot ID Rat	Name
CCL2 (MCP-1)	P13500	P14844	C-C motif chemokine 2
CCL20 (MIP3-alpha)	P78556	P97884	C-C motif chemokine 20
CCL3 (MIP1-alpha)	P10147	P50229	C-C motif chemokine 3
CCL5	P13501	P50231	C-C motif chemokine 5
CNTF	P26441	P20294	Ciliary neurotrophic factor
CRP	P02741	P48199	C-reactive protein
CXL10 (IP10)	P02778	P48973	C-X-C motif chemokine 10
EGF	P01133	P07522	Pro-epidermal growth factor
GROA (CXCL1)	P09341	P14095	Growth-regulated alpha protein
HAVR1	Q96D42	O54947	Hepatitis A virus cellular receptor 1 (Human) Hepatitis A virus cellular receptor 1 homolog (Rat)
ICAM1	P05362	Q00238	Intercellular adhesion molecule 1
IFNG	P01579	P01581	Interferon gamma
IL10	P22301	P29456	Interleukin-10
IL1A	P01583	P16598	Interleukin-1 alpha
IL1B	P01584	Q63264	Interleukin-1 beta
IL6	P05231	P20607	Interleukin-6
LYAM1	P14151	P30836	L-selectin
NGF	P01138	P25427	Beta-nerve growth facto
AGER (RAGE)	Q15109	Q63495	Advanced glycosylation end product-specific receptor
TNFA	P01375	P16599	Tumor necrosis factor
VEGFB	P49765	O35485	Vascular endothelial growth factor B
X3CL1	P78423	O55145	Fractalkine

Table 2. Final cytokine assay panel.

measured *in vitro*. This implied to activate or repress a large range of pathways/biological functions to ensure broad perturbation coverage of the biological system. Therefore, Fig. 4 illustrates our strategy in selecting various stimuli to perturb as many pathways as possible and how an initial experiment was performed to screen various phosphorylated proteins following stimuli exposure. This strategy provided a final selection of stimuli active in human, rat or both species to be used for the main experiment.

Stimulus selection by *in silico* analysis and literature review. The following criteria were considered for the initial selection of potential candidate stimuli, including 1) stimuli that modulate the activity of transcription factors/regulators; 2) classical stimuli known to target specific pathways; and 3) stimuli with heterogeneous downstream effects. Computational and manual curation approaches were undertaken to achieve an appropriate selection (Fig. 5).

Stimuli that modulate the activity of transcription factors/regulators. Transcription factors/regulators directly regulate the transcription of target genes. Querying of databases containing biological knowledge, such as the Ingenuity database (Ingenuity® Systems, www.ingenuity.com), enabled the identification of compounds that could modulate the activity of transcription factors/regulators expressed in the tissues/cells (lung, lung cells, lung cell lines, lung tissue, small airway epithelial cells, airway epithelium and airway epithelial cells) and organisms of interest (human, mouse and rat). Overall, 710 compounds were identified that could modulate the activity of 182 transcription factors/regulators.

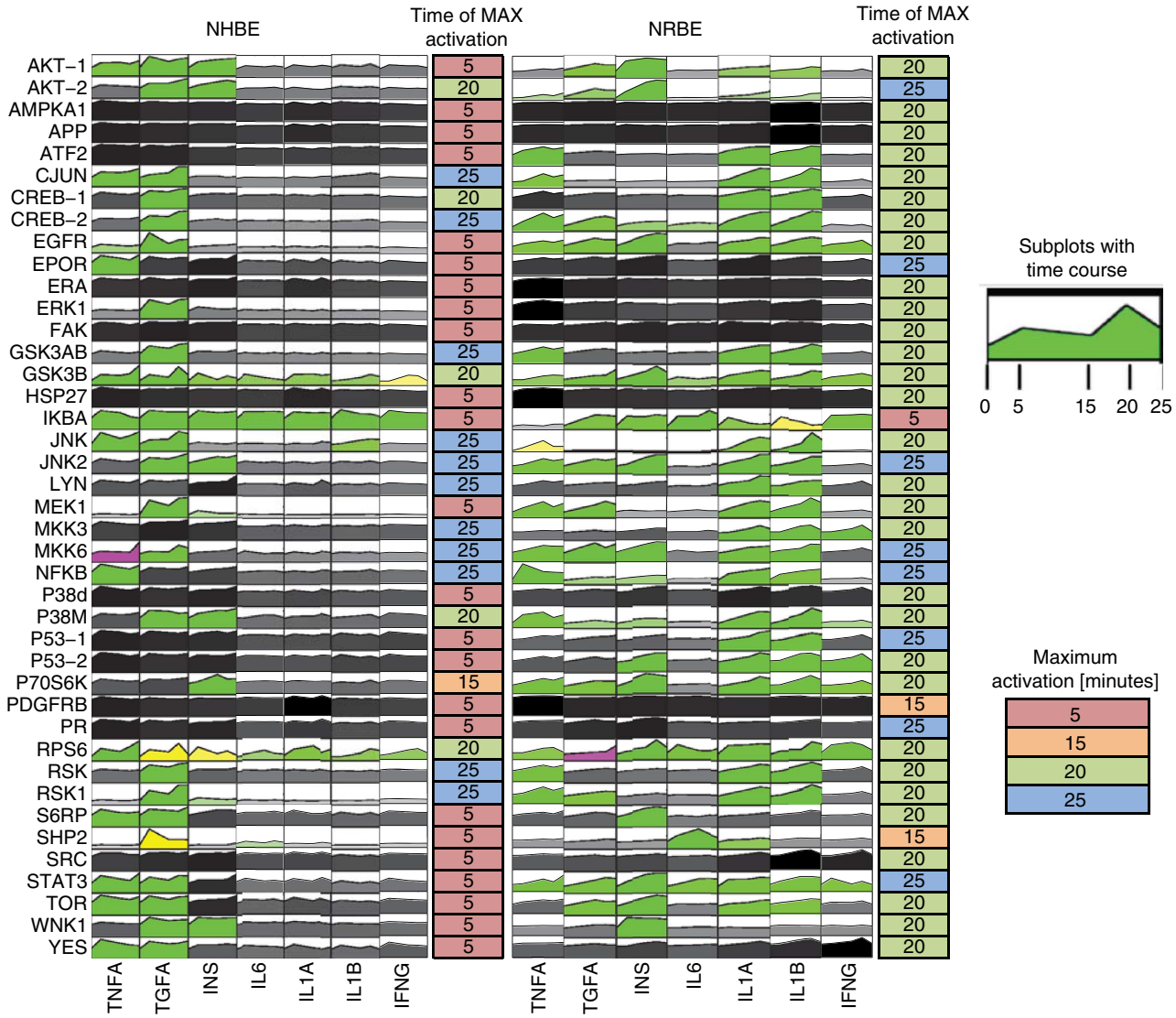

Figure 4. The determination of optimal time points for phosphoproteomics measurements in NHBE and NRBE.Human and rat bronchial epithelial cells were treated with seven stimuli at five different time points (0, 5, 15, 20 and 25 min). The time course of the raw data (fluorescent intensity: FI) for each phosphoprotein was plotted in subplots using a modified version of DataRail. The solid fill colours (yellow, green, purple, grey/black) of the time course correspond to different signal behaviour over time according to the DataRail colouring scheme. Yellow colour corresponds to transient activity (FI increases and then decreases), green colour corresponds to sustained activity (FI increases and remains active), purple colour corresponds to late activity (FI starts stable and then increases) and grey/black to no change (FI increase/decrease compare to basal level at 0 time point less than 50% across all time points - the darker the grey colour the bigger the average FI). In the majority of experiments, maximum phosphoprotein activation in NHBE cells was found at 5 (red) and 25 (blue) minutes, whereas NRBE cells were maximally activated at 20 (green) and 25 (blue) minutes. Thus, 5 and 25 min were selected as the optimal time points for both cell types.

Compounds that modulate the activity of many different transcription regulators (e.g., beta-estradiol) were prioritized and retained in the initial stimulus list.

Stimuli known to target specific pathways. Prototypical stimuli that have been extensively used as proxy tools to perturb (activate or inhibit) specific pathways were identified from the literature and selected (e.g., rapamycin, an inhibitor of the mTor pathway; lipopolysaccharide, an activator of the NFκB signaling pathway; and tunicamycin, an inducer of the unfolded protein response). Some cytokines and growth factors were also selected because they target very specific pathways.

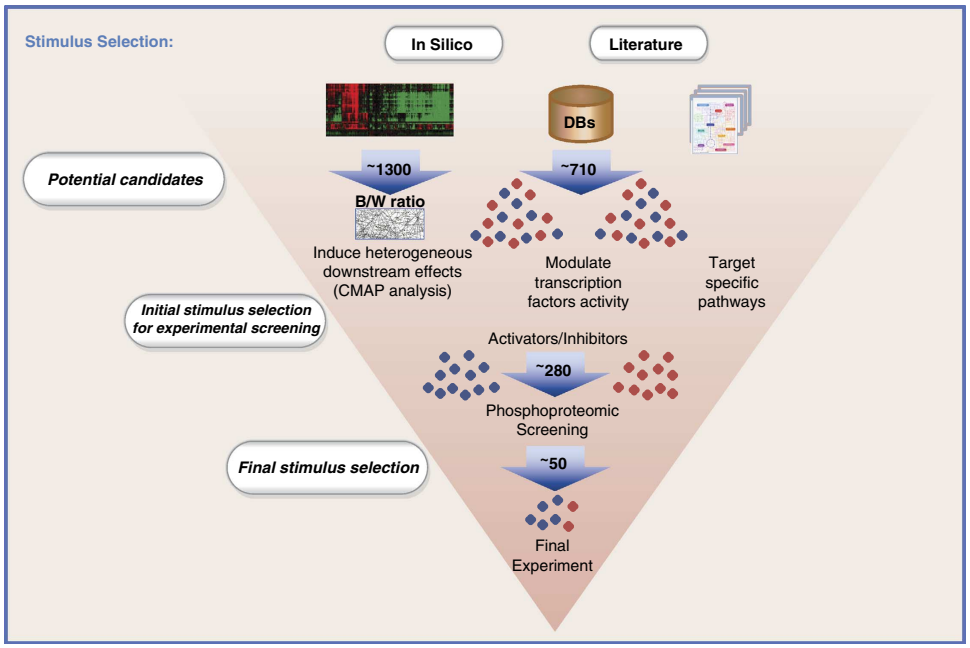

Figure 5. The process of selection of the stimuli used to generate the dataset for the Species Translation Challenge. The selection processes involve various steps, including *in silico* analysis, literature review, and phosphoproteomics screening.

Stimuli with heterogeneous downstream effects. Analysis of the connectivity map (CMAP) large-scale expression compendium enabled the selection of compounds that induce heterogeneous downstream effects[17]. The CMAP dataset is a collection of genome-wide gene expression profiles that represent the transcriptional responses of five different human cell lines (HL60, MCF7, PC3, SKMEL5 and ssMCF7) to 1,309 different compounds (small active molecules) or control vehicles following 6 h of exposure (12 h for a specific subset of compounds)[17]. Interestingly, Lorio and colleagues have constructed a 'drug network' partitioned into communities using the CMAP dataset[18]. Drugs within a community were clustered together on the basis of similar regulation patterns of gene expression, suggesting an analogous mode of actions. This 'drug network' was leveraged to identify drugs/compounds with heterogeneous downstream effects by computing a between- *versus* within-community (B/W) average distance ratio. Compounds with the highest ratios were prioritized during the screening of the stimuli. Including the review of the scientific literature, a total set of 270 stimuli were selected for *in vitro* testing as described below.

In vitro stimulus screening

In vitro screening was performed to identify a subset of stimuli that could elicit responses in NHBE and NRBE cells, which then would be used for the main experiment (Fig. 5). This *in vitro* screening was performed by measuring the phosphorylation levels of proteins from the lysates of NHBE and NRBE cells that were exposed to 270 compounds for 5 min. For each of the 270 compounds, the concentration was manually curated from the literature or through a semi-automated literature-mining approach (Supplementary Table 2 deposited in the figshare public repository). Briefly, the Google search engine was used to query the names and common aliases of the 270 compounds found in the HUGO database (for cytokines and growth factors) followed by the concentration units (i.e., 'mg/ml', 'ng/ml', 'mM', 'μM', 'nM'). The concentrations were automatically extracted from the top 100 results and plotted on a histogram. A rounded value 20% above the median of the histogram was chosen as the final concentration.

Final stimulus selection for the main experiment

Out of 270 compounds that were originally used in the screening of phosphoproteomics, the most potent compounds were selected for the main experiment. These compounds, including activators/inhibitors, were chosen based on (i) the number of phosphorylation signals that were affected, (ii) the strength of their responses (maximum fold increase of the activated signals), (iii) the diversity of the affected pathways, and (iv) and downstream gene expression changes (see paragraph 'Stimuli with heterogeneous downstream effect') (Supplementary Table 2 deposited in the figshare public repository). With respect to these criteria for the final selection of stimuli, the following analysis was performed. The screening phase included one biological replicate for each stimulus/protein/species, therefore, a statistical analysis could

not be performed and an alternative approach was followed. For each protein, a fold change of the signal was calculated comparing the phosphorylation signals measured for compound-treated cells and for unstimulated cells (control). Subsequently, a number of thresholds ranging from 1.25 to 2.0-fold (i.e. 1.25, 1.3, 1.35, 1.4, 1.45, 1.5, 1.55, 1.60, 1.65, 1.7, 1.75, 1.8, 1.85, 1.9, 1.95 and 2.0-fold) were used to binarize signals as active or non-active. For example, a threshold of 2.0 implies that an increase of 2-fold or higher was required for the signal to be considered activated. For each threshold, the number of signals considered as up- and down-regulated was calculated for each stimulus across all phosphoproteins. This number was used as a score to sort the stimuli from the most to the less potent. The most potent stimuli were prioritized also with respect to the other criteria (mentioned above) for the final selection of stimuli (Supplementary Table 4).

III-Main experiment

The cell culture and experimental conditions established for NHBE and NRBE cells during the optimization phase were applied in the main experiment. Both NHBE and NRBE cells were seeded in 96-well plates on day one and incubated overnight at 37 °C in 5% CO_2. The cells were starved in bronchial epithelial cell basal medium (Lonza) for 4 h and then exposed, in parallel, to 52 different stimuli or to a control medium (DME), which is the standard culture medium for these cells. The cells were collected and lysed at different time points (5 and 25 min for phosphoproteins measurement, 6 h for gene expression measurement) and supernatants were collected at 24 h for cytokine release measurement in the supernatants. The cells were exposed to each stimulus in triplicate, or in 5-plicates and 6-plicates for the DME controls. To avoid spatial confounding effects, the stimuli and DME controls were randomly distributed throughout the 96-well plate.

Datasets for the main experiment were generated from two sets of independent experiments: 75% of the stimuli (40 compounds) were tested first; whereas, the remaining 25% (12 compounds) were tested in a second experiment. For each experiment DME controls were included. The final STC compendium contains a collection of phosphoproteomics, transcriptomics, and cytokine data.

References

1. Meyer, P. *et al.* Verification of systems biology research in the age of collaborative competition. *Nat. Biotechnol.* **29,** 811–815 (2011).
2. Meyer, P. *et al.* Industrial methodology for process verification in research (IMPROVER): toward systems biology verification. *Bioinformatics* **28,** 1193–1201 (2012).
3. Bhattacharya, S., Zhang, Q., Carmichael, P. L., Boekelheide, K. & Andersen, M. E. Toxicity testing in the 21 century: defining new risk assessment approaches based on perturbation of intracellular toxicity pathways. *PloS One* **6,** e20887 (2011).
4. Huber, P. J. *Robust Statistics*. (John Wiley & Sons, Inc., 1981).
5. Wu, Z., Irizarry, R. A., Gentleman, R., Murillo, F. M. & Spencer, F. A model based background adjustment for oligonucleotide expression arrays. *J. Am. Stat. Assoc.* **99,** 909–917 (2004).
6. Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4,** 249–264 (2003).
7. Rayner, T. F. *et al.* A simple spreadsheet-based, MIAME-supportive format for microarray data: MAGE-TAB. *BMC Bioinformatics* **7,** 489 (2006).
8. Bartelt, R. R., Cruz-Orcutt, N., Collins, M. & Houtman, J. C. Comparison of T cell receptor-induced proximal signaling and downstream functions in immortalized and primary T cells. *PloS One* **4,** e5430 (2009).
9. Boerma, M. *et al.* Comparative expression profiling in primary and immortalized endothelial cells: changes in gene expression in response to hydroxy methylglutaryl-coenzyme A reductase inhibition. *Blood Coagul. Fibrin.* **17,** 173–180 (2006).
10. Czekanska, E. M., Stoddart, M. J., Ralphs, J. R., Richards, R. G. & Hayes, J. S. A phenotypic comparison of osteoblast cell lines versus human primary osteoblasts for biomaterials testing. *J. Biomed. Mater. Res. A* (2013).
11. Hou, A., Voorhoeve, P. M., Lan, W., Tin, M. & Tong, L. Comparison of gene expression profiles in primary and immortalized human pterygium fibroblast cells. *Exp. Cell Res.* **319,** 2781–2789 (2013).
12. Pan, C., Kumar, C., Bohl, S., Klingmueller, U. & Mann, M. Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions. *Mol. Cell. Proteomics* **8,** 443–450 (2009).
13. Pezzulo, A. A. *et al.* The air-liquid interface and use of primary cell cultures are important to recapitulate the transcriptional profile of in vivo airway epithelia. *Am. J. Physiol. - Lung C* **300,** L25–L31 (2011).
14. Clarke, D. C., Morris, M. K. & Lauffenburger, D. A. Normalization and statistical analysis of multiplexed bead-based immunoassay data using mixed-effects modeling. *Mol. Cell. Proteomics* **12,** 245–262 (2013).
15. Saez-Rodriguez, J. *et al.* Flexible informatics for linking experimental data to mathematical models via DataRail. *Bioinformatics* **24,** 840–847 (2008).
16. Alexopoulos, L. G., Saez-Rodriguez, J., Cosgrove, B. D., Lauffenburger, D. A. & Sorger, P.K. Networks inferred from biochemical data reveal profound differences in toll-like receptor and inflammatory signaling between normal and transformed hepatocytes. *Mol. Cell. Proteomics* **9,** 1849–1865 (2010).
17. Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313,** 1929–1935 (2006).
18. Iorio, F. *et al.* Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc. Natl Acad. Sci. USA* **107,** 14621–14626 (2010).

Data Citations

1. Poussin, C., Mathis, C., Alexopoulos, L. G., Messinis, D. E., Dulize, R. H. J., Belcastro, V., Melas, I. N., Sakellaropoulos, T., Rhrissorrakrai, K., Bilal, E., Meyer, P., Talikka, M., Boué, S., Norel, R., Rice, J. J., Stolovitzky, G., Ivanov, N. V., Peitsch, M. C., Hoeng, J., Ansari, S. ArrayExpress E-MTAB-2091 (2014).
2. Poussin, C., Mathis, C., Alexopoulos, L. G., Messinis, D. E., Dulize, R. H. J., Belcastro, V., Melas, I. N., Sakellaropoulos, T., Rhrissorrakrai, K., Bilal, E., Meyer, P., Talikka, M., Boué, S., Norel, R., Rice, J. J., Stolovitzky, G., Ivanov, N. V., Peitsch, M. C., Hoeng, J., Ansari, S. Figshare http://dx.doi.org/10.6084/m9.figshare.960097 (2014).

Acknowledgements

We acknowledge Dr Sam Ansari for preparing and submitting the data to public repositories. We also thank Lionel Schilli, Jean Binder and Anouk Ertan for their contribution in the management of the project, Dr Anita Iskandar for scientific review of the manuscript, and Dr Justyna Szostak for the revision of the stimulus table.

Author Contributions

CP, CM LGA, DEM wrote the manuscript. CP, CM, LGA, VB, KR, EB, PM, RN, JR, GS, JH designed the challenge. CP, CM, LGA, JH designed and followed up the experiment. LGA, DEM, RD, NI performed the experiment and/or generated the data. CP, CM, LGA, VB, MT, SB contributed to stimulus identification and selection. CP, VB, INM, TS, KR, EB, PM contributed to data analysis. JH and MCP contributed to the manuscript and supported the project.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/sdata

Competing financial interests: The authors declare no competing financial interests.

Funding source: This work was funded by Philip Morris international R&D.

How to cite this article: Poussin, C. *et al.* The species translation challenge—A systems biology perspective on human and rat bronchial epithelial cells. *Sci. Data* 1:140009 doi: 10.1038/sdata.2014.9 (2014).