Received: 21 December 2020 | Revised: 15 March 2021 | Accepted: 22 March 2021

DOI: 10.1002/psp4.12626

ARTICLE

Integrative chemogenomic analysis identifies small molecules that partially rescue ΔF508-CFTR for cystic fibrosis

Rachel A. Hodos[1,2] | Matthew D. Strub[3,4] | Shyam Ramachandran[5] | Ella A. Meleshkevitch[6] | Dmitri Y. Boudko[6] | Robert J. Bridges[6] | Joel T. Dudley[1] | Paul B. McCray Jr.[3,4]

[1]Institute for Next Generation Healthcare, Mount Sinai School of Medicine, New York, NY, USA

[2]Courant Institute for Mathematical Sciences, New York University, New York, NY, USA

[3]Department of Pediatrics, University of Iowa, Carver College of Medicine, Iowa City, IA, USA

[4]Interdisciplinary Graduate Program in Genetics, University of Iowa, Iowa City, IA, USA

[5]Editas Medicine, Cambridge, MA, USA

[6]Department of Physiology and Biophysics, Rosalind Franklin University, North Chicago, IL, USA

Correspondence
Paul B. McCray, Department of Pediatrics, University of Iowa, 169 Newton Rd., Iowa City, IA 52242, USA.
Email: paul-mccray@uiowa.edu

Joel T. Dudley, Institute for Next Generation Healthcare, Mount Sinai School of Medicine, 770 Lexington Avenue, New York, NY 10065, USA.
Email: joel.dudley@mssm.edu

Present address
Rachel A. Hodos, BenevolentAI, New York, NY, USA

Funding information
National Institutes of Health RO1 HL118000, P01 HL51670, and P01 HL091842, the Center for Gene Therapy of Cystic Fibrosis P30 DK54759, and

Abstract

Rare diseases affect 10% of the first-world population, yet over 95% lack even a single pharmaceutical treatment. In the present age of information, we need ways to leverage our vast data and knowledge to streamline therapeutic development and lessen this gap. Here, we develop and implement an innovative informatic approach to identify therapeutic molecules, using the Connectivity Map and LINCS L1000 databases and disease-associated transcriptional signatures and pathways. We apply this to cystic fibrosis (CF), the most common genetic disease in people of northern European ancestry leading to chronic lung disease and reduced lifespan. We selected and tested 120 small molecules in a CF cell line, finding 8 with activity, and confirmed 3 in primary CF airway epithelia. Although chemically diverse, the transcriptional profiles of the hits suggest a common mechanism associated with the unfolded protein response and/or TNFα signaling. This study highlights the power of informatics to help identify new therapies and reveal mechanistic insights while moving beyond target-centric drug discovery.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

Current therapeutic options to treat the basic cystic fibrosis (CF) defect for the most common mutation, ΔF508, have limited efficacy and do not benefit all individuals. Whereas the gold standard elexacaftor-tezacaftor-ivacaftor treatment increases FEV_1 by 13.8 points, the corrected function is still well below even carrier levels and ~ 30% of patients experience less than a 5% increase. Likewise, CF-related lung disease still persists in individuals taking ivacaftor alone. Therefore, a substantial treatment gap remains and improved correctors are needed to maximize pulmonary health. Although many in vitro strategies have been identified that can partially ameliorate the CF defect, the challenge remains to translate these into therapies and to understand the mechanisms of such strategies.

WHAT QUESTION DID THIS STUDY ADDRESS?

Integrating gene expression and pathway information from a variety of sources, we aimed to identify novel molecules to rescue the trafficking defect of ΔF508-CFTR.

Rachel A. Hodos and Matthew D. Strub contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *CPT: Pharmacometrics & Systems Pharmacology* published by Wiley Periodicals LLC on behalf of American Society for Clinical Pharmacology and Therapeutics.

the Cystic Fibrosis Foundation. We acknowledge the support of the In Vitro Models and Cell Culture Core, the Roy J. Carver Charitable Trust (P.B.M.) and Training Grant T32GM008629 (M.D.S.).

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

We identified a number of novel molecules with efficacy to rescue ΔF508-CFTR in a CF cell line and primary cells. Post hoc analysis of the small molecule-associated expression profiles suggests a common mechanism across our set of chemically diverse molecules, involving TNFα signaling and/or induction of the unfolded protein response.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

This work provides a positive example of an integrative connectivity mapping approach to identify novel compounds with relevant biological activity.

INTRODUCTION

Cystic fibrosis (CF) is a recessive Mendelian disease affecting roughly 1 in 2000 people of European ancestry, causing a severe and progressive lung disease characterized by persistent bacterial infection, inflammation, and bronchiectasis. Although the genetic basis of the disease, mutations in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, was identified in 1989,[1] few new therapeutic options have advanced to the clinic based on this discovery. The most common mutation, ΔF508, results in a misfolded protein that is degraded before trafficking to the plasma membrane where it normally functions as an anion channel. However, if ΔF508-CFTR is able to reach the plasma membrane, it can retain channel function.[2] Hence, one goal in CF research is to rescue the trafficking and function of this mutant protein. Toward this end, several drug combinations (identified by high throughput screening) have recently been approved by the US Food and Drug Administration (FDA) for ΔF508 homozygous patients but show only modest efficacy.[3,4] Although elexacaftor-tezacaftor-ivacaftor treatment in patients with one ΔF508 allele and a minimal-function mutation on the second allele resulted in an increased percentage of predicted forced expiratory volume in 1 s (FEV_1) of 13.8 points, the improved lung function was still well below even carrier levels.[4,5] Furthermore, ~ 30% of ΔF508 homozygotes receiving elexacaftor-tezacaftor-ivacaftor experienced less than a 5% increase in FEV_1 and some patients do not tolerate the medications.[6] Likewise, although ivacaftor alone has been shown to improve FEV_1 and decrease the rate of pulmonary exacerbations in patients with gating mutations, CF-related lung disease still persists in these patients.[7–9] Whereas other related drugs are being developed, a treatment gap remains and improved correction is required to maximize the pulmonary health of patients. Many in vitro interventions have also been identified that partially rescue function of ΔF508-CFTR, such as low-temperature treatment[2]; knockdown of *AHSA1*,[10] *SYVN1*,[11] *NEDD8*,[11] or *SIN3A*[12]; and overexpression of miR-138[12] or miR-16.[13] In parallel, since the 1989 discovery of CFTR's role in CF, much biological knowledge has accumulated about its protein structure, biosynthesis, and trafficking. However, the challenge remains to translate these interventions and insights into effective therapies and understand their mechanisms of activity.

One strategy that offers promise to help bridge this gap is called connectivity mapping.[14] Connectivity mapping compares disease-relevant gene expression signatures with a database of small molecule-induced signatures to identify small molecules that either mimic an established rescue intervention or reverse a disease signature. To support such work, large public compendia of chemogenomic data have been assembled (namely, the Connectivity Map or "CMap"[14] and the L1000 chemical perturbation data from the Library of Integrated Network-based Cellular Signatures, herein "LINCS L1000"[15]) that catalog transcriptional responses of human cell lines treated with thousands of drugs and other small molecule compounds. In this work, we develop an integrative computational pipeline to select small molecules for experimental evaluation, extending the principle of connectivity mapping to include both transcriptomic and pathway-based information. Similar approaches have been reported (e.g., applying connectivity mapping in the context of signaling pathways to identify candidate small molecules to treat idiopathic pulmonary fibrosis[16,17] or integrating multiple disease signatures characterizing lung cancer to prioritize and study small molecules with common mechanisms).[18] The pipeline described herein builds on our recent work[19] by integrating three core gene expression signatures associated with CF disease and in vitro rescue along with expert-curated pathways[20] relevant to CFTR biosynthesis and trafficking.

This work is not the first to apply a connectivity mapping strategy to identify novel CFTR correctors, and, importantly, prior work has provided evidence to support such a strategy. For example, our group previously applied a connectivity mapping approach using transcriptional signatures of miR-138 overexpression and *SIN3A* knockdown in Calu-3 epithelial cells to query the CMap database.[21] We tested 27 small molecules and identified 4 that partially restore the trafficking and function of ΔF508-CFTR. In another study, Pesce et al. queried CMap using a signature characterizing

low-temperature incubation, again seeking small molecules that mimic this rescue signature.[22] They identified several glucocorticoids that demonstrated rescue in CFBE cells, a widely used human airway epithelial cell line expressing ΔF508-CFTR,[23] but not in primary CF airway epithelial cells. Malcomson et al. also used connectivity mapping to identify licensed drugs with TNFAIP3-induced anti-inflammatory effects. Ikarugamycin and quercetin were shown to normalize the inflammatory response in CF airway epithelial cells.[24] Finally, a study by Zhang et al. found that a number of cardiac glycosides (previously identified to rescue ΔF508-CFTR via high-throughput screening) induce highly similar transcriptional patterns to that of low-temperature rescue of CFTR. Hence, although their work does not apply connectivity mapping, it further supports the notion that transcriptional connectivity among cellular perturbations can be associated with CFTR rescue.[25]

The present study expands on such work in several ways. First, we cast a wider net by querying signatures from 13,000 molecules from the LINCS L1000 dataset in addition to the 1309 from CMap used in past studies. Second, the CF transcriptional signatures that we use are derived from meta-analyses of multiple studies, and hence are expected to be of higher quality than signatures derived from individual, smaller-scale studies.[19] Third, we develop an integrative scoring approach to rank small molecules for predicted efficacy using multiple CF-relevant queries, including both pathways and transcriptional signatures and show that our integrative approach outperforms comparable approaches based on a single data source. Finally, we present an approach to computationally validate our scoring strategy using previously identified small molecules as a validation set.

We start by briefly describing our small molecule prioritization pipeline and then present some in silico validation of the resulting rankings based on a set of 15 molecules previously identified to partially rescue ΔF508-CFTR. Then, we present results of selected molecules assayed in CFBE cells and subsequent evaluation in primary cultures of CF airway epithelia. Finally, we describe post hoc chemogenomic analysis to shed insight on potential mechanisms.

METHODS

Small molecule scoring and selection

Chemogenomic signature processing

Chemogenomic profiles from both CMap and LINCS L1000 were downloaded and processed for connectivity mapping. First, raw data from CMap were downloaded from https://portals.broadinstitute.org/cmap/ and log-transformed fold changes (FCs) were computed for each of the 6100 signatures representing 1309 molecules profiled in 5 human cancer cell lines (MCF7, PC3, HL60, SKMEL5, and ssMCF7). Additionally, LINCS L1000 z-scores[15] were downloaded using the *lincscloud* API. Only signatures labeled as "gold" (indicating signal strength and consistency between replicates) were used, covering 13,000 unique molecules. We further processed the signatures in both databases to generate a consensus signature per small molecule/cell line combination, using the Prototype Ranked List method, as described in Iorio et al.,[26] which combines gene rankings via the geometric mean. Quantile-normalized z-scores (for LINCS L1000) and FCs (for CMap) corresponding to each gene rank were also computed for use in the *XSum* score, as described in File S1.[27]

Preparation of CF signature- and pathway-based gene sets

Three meta-analyses of CF disease and rescue datasets were performed as described in ref. 19. This resulted in three core signatures, namely: (1) disease, characterizing CF versus wild type (WT) samples from both human nasal and bronchial brushings, as well as tissues dissected from newborn pig airways; (2) low-temperature, capturing transcriptional changes associated with incubating CFBE41o- cells at 27°C for 24 h compared with cells kept at 37°C; and (3) RNAi, a meta-analysis from signatures associated with RNAi-based interventions, including knockdown of *SIN3A*, *SYVN1*, and *NEDD8*, and overexpression of miR-138. From each of these 3 signatures, the top and bottom K differentially expressed genes (DEGs; ranked by FC) were extracted, where K was 50, 100, or 150 for the disease signature, and 100, 200, or 300 for both the low temperature and RNAi rescue signatures (multiple values of K were used in order to avoid being overly dependent on a single choice of K, and the smaller number of DEGs selected for the disease signature is due to the limited number of significant DEGs in this signature). DEGs were called based on two filters: (1) adjusted $p < 0.10$, 0.01, and 0.01, respectively, for the disease, low temperature, and RNAi signatures; and (2) gene representation in all or most experiments that informed each meta-analysis (6, 3, and 6experiments, respectively). This resulted in nine signature-based gene sets, corresponding to three values of K for each of the three signatures (see Figure 1a). To provide an additional input to the scoring pipeline, a set of nine expert-curated pathways related to CFTR processing, trafficking, and degradation were selected from the CF MetaMiner platform.[20] All signature- and pathway-based gene sets used in this work are listed in Table S3.

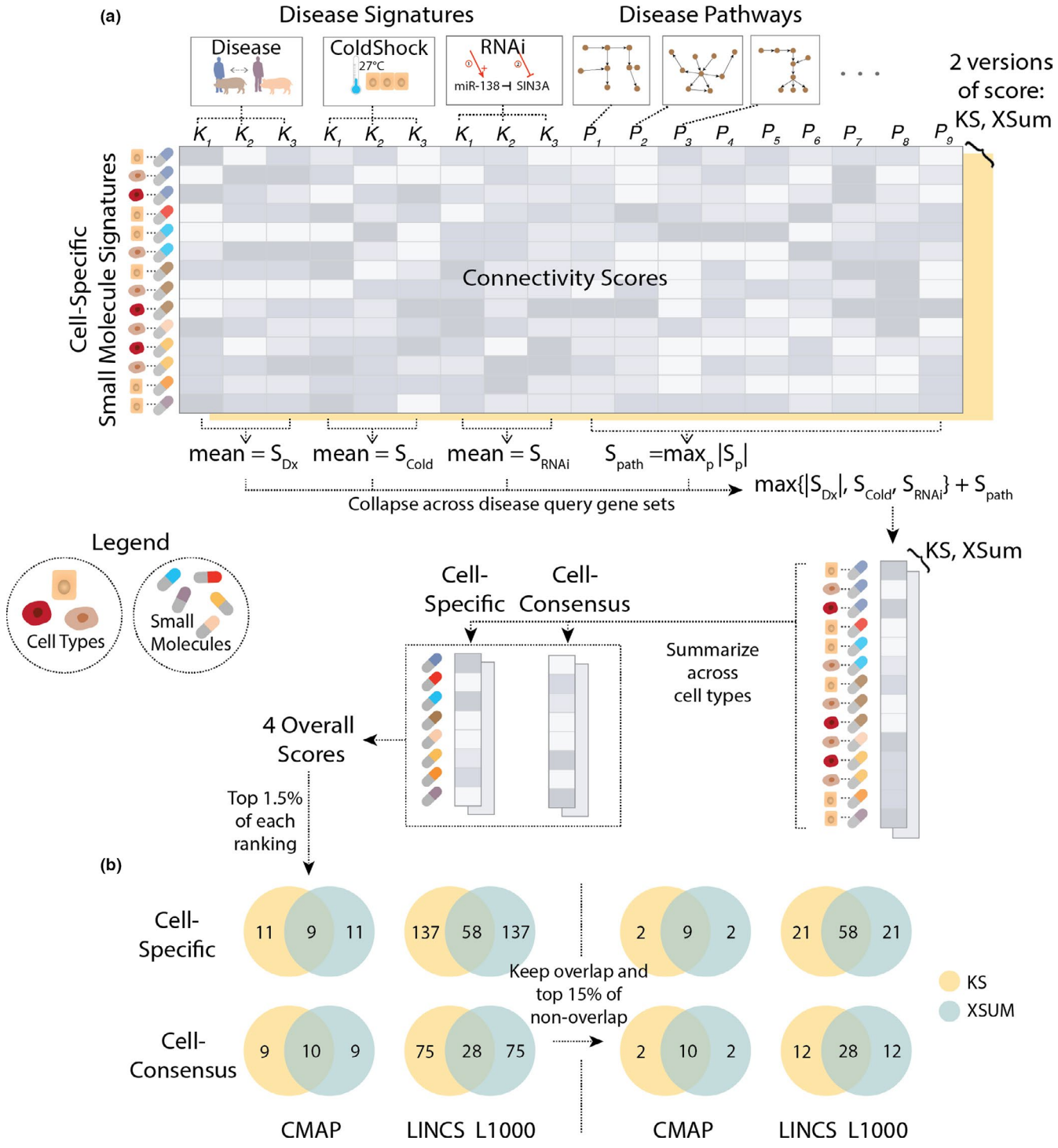

FIGURE 1 (a) Integrative small molecule prioritization pipeline for CF therapeutic discovery. Molecules were prioritized using an aggregate score derived from three signatures from a prior meta-analysis[19] (and using three numbers of DEGs for each signature, denoted by K_i for *i* in 1–3), in combination with nine pathways related to CFTR processing and trafficking (denoted by P_i for i in 1–9). Scores for each cell-specific signature are first collapsed across the 18 CF query gene sets, and then collapsed across cell lines tested per small molecule. Two different connectivity mapping metrics and two different methods for combining across cell lines, yields four distinct overall scores for each molecule. (b) Small molecule selection strategy. Eight small molecule rankings were generated (i.e., the 4 summary scores applied to each of the 2 databases) and are represented by the eight colored circles in the left half of the figure. The top 1.5% of molecules were identified from each of the 8 rankings, where the number of molecules in the overlap between the corresponding *KS*- and *XSum*-based rankings are shown in the figure. All molecules in these overlaps were selected. Molecules that were only identified by one of *KS* or *XSum* were further filtered to only keep the top 15% (resulting in the numbers on the righthand side). Altogether this yielded 150 molecules, of which we tested 120 (see text). CF, cystic fibrosis; CMap, Connectivity Map; DEG, differentially expressed gene; LINCS, Library of Integrated Network-based Cellular Signatures

Small molecule scoring strategy

A brief summary of our small molecule scoring approach is presented here; for further details see File S1. First, for each of the signature- and pathway-based gene sets, connectivity scores were computed to quantify the degree and the direction in which each small molecule alters the expression of the genes in the query gene set. In the case of the two rescue signatures (*RNAi* and low temperature), molecules with high, positive connectivity scores were sought, aiming to mimic these rescue phenotypes. For the disease signature, small molecules were sought that either mimic or reverse the signature by taking the absolute value of the connectivity score. Finally, in the case of the nine CFTR pathways, we allowed either positive or negative pathway regulation by taking the absolute value of these connectivity scores, as we did not want to impose naïve biases about directionality in the case of complex pathway modulation.

As mentioned in the results, scores were then summarized across the 18 CF query gene sets and then across cell lines to yield a unified score per small molecule predicting its relevance to CFTR rescue. Based on two different connectivity score metrics (*KS* and *XSum*) and two different ways to summarize across cell lines (*CellSpecific* and *CellConsensus*), this resulted in four distinct summary scores (i.e., *KS-CellSpecific*, *KS-CellConsensus*, *XSum-CellSpecific*, and *XSum-CellConsensus*) for each of the two databases, LINCS and CMap.

In silico validation of small molecule scores

To test the validity of the various scores generated, we compiled a list of 15 molecules represented in the CMap and/or LINCS L1000 chemogenomic databases with prior evidence for their ability to partially restore the trafficking and/or function of ΔF508-CFTR,[21,25,28–32] although the "gold standard" correctors lumacaftor and elexacaftor are not cataloged in CMap or LINCS. We evaluated whether this validation set (call it *C*) was ranked more highly than a random selection of molecules of the same size. This was computed using the GSEA-based enrichment score *ES(R, C)* where *R* is the corresponding molecule ranking. Significance of scores was estimated using 100,000 permutations of the rankings to a null distribution of scores. Further details are included in File S1.

Small molecule selection

We derived a scheme for selecting small molecules for experimental validation among these eight rankings (4 scores times 2 databases). First, the top-scoring 1.5% of small molecules were identified for each of the eight rankings (the specific percentile threshold was chosen so that the final list would yield our predefined capacity of 100–150 small molecules.) We found that *KS*- and *Xsum*-based rankings yielded substantial overlap, and we decided to keep these and further filter the remaining molecules to only keep the top 15% from each individual category (see Figure 1b). This resulted in a set of 150 unique molecules, from which 120 were tested (see File S4 for the list of molecules tested). The remaining 30 small molecules were either not available for purchase or had been previously tested.

Functional screen of small molecule activity by conductance assay

Effects of selected small molecules were tested using a robotic Transepithelial Current Clamp (TECC)-24 assay[33] with either CFBE41o-[23] or primary human CF airway epithelial cells grown on permeable membranes of Transwell plates (Corning), representing a planar array of 24 Ussing chambers with a 6 mm diameter insert of the epithelial monolayer. Cells were pretreated 24 h before experiments with log-titrated and 0.2% DMSO-normalized small molecules with and without 3 μM C18, used as a positive control (1, 3, and 10 μM final concentration in 250 μl top +750 μl bottom solution, at 37°C). The 10 μM C18 and 0.2% DMS0 were used as positive and negative controls, respectively. C18 is a structural analog of the FDA-approved Lumacaftor. Small molecules are tested alone to measure independent efficacy and in combination with C18 to test for synergy or additive effects between the two small molecules. For primary human CF airway epithelia, baseline sodium conductance was suppressed by addition of 3 μM benzamil apically at the 7 min mark of the experiment. Chloride conductance was stimulated with the addition of 10 μM forskolin and 1 μM VX-770 apically and basolaterally at the 20 min mark. Chloride conductance was then inhibited by the basolateral addition of 20 μM bumetanide at 80 min. A similar protocol is used in CFBE41o- cells, with the only differences being the addition of benzamil occurs before the experiment and 20 μM inhibitor, Inh-172, is substituted for bumetanide. For primary human CF airway epithelia, the efficacy was quantified by the ratio of area under the curve (AUC) of chloride current calculated for the 60 min interval between additions of 10 μM forskolin and 1 μM VX-770 (20 min) and 20 μM bumetanide (80 min). For CFBE41o- cells, the efficacy was quantified by the AUC ratio of chloride conductance calculated for the 60 min interval between additions of 10 μM forskolin and 1 μM VX-770 (20 min) and 20 μM Inh-172 (80 min).

Statistical analysis of functional screen results

AUC ratios were computed for small molecule-treated replicates relative to the mean AUC of control experiments performed on the same plate. Experiments with evidence of toxicity were removed by thresholding baseline transepithelial resistance less than 400 G·cm.^{2} One-sample, two-sided *t*-tests were then performed to evaluate whether AUC ratios were different from 1. Benjamini-Hochberg[34] multiple hypothesis adjustment was applied to the results of each small molecule. The AUC thresholds were established based on experience from other small molecules screened for rescue of ΔF508-CFTR and validated positive control small molecules. The more stringent threshold of AUC greater than 1.3 was used for the initial screen because significance could not be estimated. For post hoc analysis, "inactive" was defined as follows: (1) the mean AUC was in the range (0.85–1.15) for all 4 experiments at 1 μM and 3 μM (± C18); (2) the mean AUC was below 1.15 for both experiments at 10 μM (± C18).

RESULTS

Eighteen disease-relevant gene sets were compiled and used to query 1309 molecules from CMap and 13,000 from LINCS L1000, quantifying the degree to which each small molecule alters the expression of genes in each disease gene set (see Figure 1a). The 18 CF gene sets included 9 expert-curated pathways related to CFTR processing and trafficking, as well as 9 lists of DEGs extracted from 3 signatures derived from meta-analyses of (1) CF versus WT pig and human airways, (2) low-temperature rescue of ΔF508-CFTR, and (3) RNAi-based rescue of ΔF508-CFTR. Scores were summarized across the 18 gene sets and then across cell lines to yield an overall score per molecule predicting its relevance to CFTR rescue. Based on two different connectivity score metrics (*KS* and *XSum*) and two different ways to summarize across cell lines (*CellSpecific* and *CellConsensus*), this resulted in four distinct summary scores (i.e., *KS-CellSpecific*, *KS-CellConsensus*, *XSum-CellSpecific*, and *XSum-CellConsensus*) for each of the two databases, LINCS and CMap. See Methods and File S1 for details.

In silico evaluation of small molecule scoring

Small molecule ranking significantly prioritizes known CFTR correctors

As an initial test of our scoring strategy, a list of 15 small molecules (13 of which are represented in LINCS L1000) with prior evidence of ΔF508-CFTR correction was compiled and used as a validation set (see Methods and File S1). For each of the summary scores as well as the individual pathway and signature scores, we tested whether the rankings of the known correctors were significantly higher than those of randomly selected molecules. Results are shown in Figure 2a. Figure 2a shows the most significant ranking, which was for one of the integrative scores—*XSum-CellSpecific*, where the validation set achieved an enrichment score (ES) of 0.55 in the range (−1 to 1; $p = 1.7e\text{-}4$). Figure 2b summarizes these results for the eight overall rankings (in the first two columns) in addition to the individual signature- and pathway-based rankings. A few observations can be made from this figure. First, we see that the integrative scoring strategies were generally better than or competitive with the nonintegrative (signature and pathway) approaches, and this was particularly true for the scores using the *CellSpecific* approach to summarize across cell lines. The second observation is that the signature-based scores tended to have better performance than the pathway-based scores. Finally, we see that ranking molecules by the absolute value of the disease signature connectivity score (i.e., allowing molecules that either mimic or reverse this signature) gave better performance than using the signed score ("disease abs" and "disease raw" columns, respectively), leading us to use the former score within our integrative small molecule scores.

Small molecule selections are consistent between chemogenomic databases

As an additional sanity check, we tested whether the molecules selected from each of the two databases had a significant overlap. Indeed, we found that among the 974 molecules represented in both databases, 17 were selected from CMap, 25 from LINCS L1000, and 7 were in the intersection (Fisher's exact $p = 4.9e\text{-}8$).

Experimental validation

Based on these scores, 120 molecules were selected (see Methods for details) for an initial screen using a transepithelial conductance assay to test for rescue of ΔF508-CFTR anion channel function in the CFBE^{23} cell line.

Eight molecules significantly increase CFTR-dependent chloride conductance in CF cell line

Of the 120 small molecules screened, 38 had a mean AUC ratio of at least 1.3 at one or more concentrations, alone or with the ΔF508-CFTR corrector compound C18, suggesting potential efficacy (however, the sample size did not allow us to make judgments on significance). Of these, 19 were

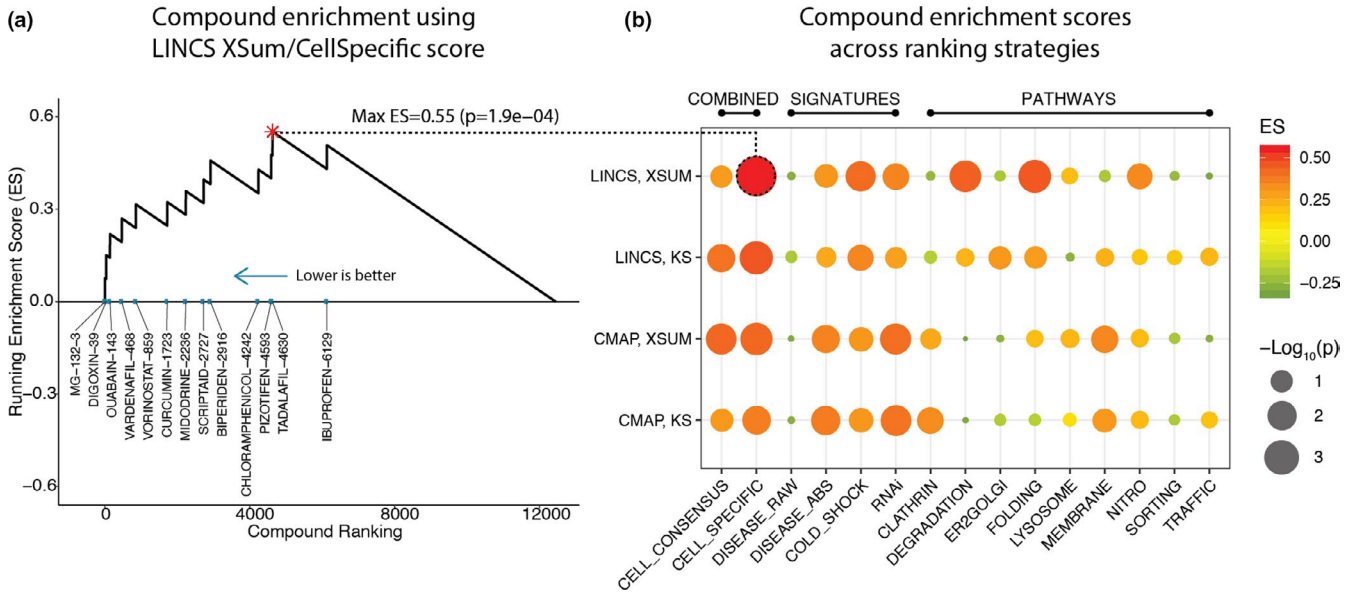

FIGURE 2 In silico validation of integrative scoring strategy. (a) Ranking and corresponding enrichment of the molecules from the validation set that are represented in the L1000 data, based on the *XSum-CellSpecific* score. Numbers next to each molecule name correspond to the ranks (lower number means higher prioritization). (b) Overview of analogous computations performed for a variety of score types, including those from our integrative approach (left 2 columns) as well as the individual signature-based scores (next 4 columns), and pathway-based scores (remaining columns). Pathway information is included in Table S3. Signature scores indicate the mean across the three values of *K*. "disease_raw" indicates the signed score, whereas "disease_abs" indicates the ranking resulting from the absolute value of this score. CMap, Connectivity Map; LINCS, Library of Integrated Network-based Cellular Signatures

selected (using a triage strategy based on various factors in the results, such as number of significant experiments and maximum efficacy across the 6 experiments) for validation at an additional $n = 3$, again testing at 3 concentrations (1, 3, and 10 μM) both with and without corrector compound C18. Among these, eight small molecules demonstrated significant efficacy (adjusted $p < 0.05$, AUC ratio > 1.2) for at least one concentration and background (see Figure 3a). Herein, we refer to these eight small molecules as the "CFBE hits."

Results in primary human airway cells

Among the CFBE hits, five small molecules were selected for further evaluation in human primary CF airway epithelial cells, using the same conductance assay. Molecules were selected based on the combination of mean AUC and significance across concentrations and backgrounds (and strophanthidin was ruled out due to the existence of prior studies). Of these, three demonstrated significant efficacy (adjusted $p < 0.05$, AUC ratio > 1.2) in at least one concentration and background (see Figure 3b).

Investigation of potential mechanisms for hit molecules

To provide insight into potential mechanisms by which the identified molecules partially rescue the trafficking of ΔF508-CFTR, we returned to their chemogenomic profiles from the LINCS L1000 data (i.e., the profiles that were used to prioritize them initially for experimentation). We focus first on all 8 CFBE hits and compare them to a set of 13 molecules for which we felt the most confident to label as inactive (see Methods). A comparative enrichment analysis of the associated small molecule-induced transcriptional signatures is shown in Figure 4. The pathways shown were selected based on the criteria that they each rank in the top 10 enrichments for at least 2 of the 8 CFBE hits. First, we observe that there is marked similarity in the enrichments across hit molecules and also in some of the inactive molecules. We also observe that multiple pathways highlighted are associated with cellular stress, including UV response, hypoxia, apoptosis, and P53 signaling. Two pathways that stand out from the results are "PERK-mediated UPR," which had particularly high fold enrichments, and "TNFα signaling via NF-κB," which was the most significant enrichment for many of the small molecules. Comparison of the fold ESs between hits and inactive molecules revealed that eight of these pathways were significantly different (adjusted $p < 0.05$) between the two groups, as shown in Figure 4b. Finally, a closer look at individual enrichment results for the three primary hits (see Table S2) reveals a similar story. Most notably, upregulation of "TNFα signaling via NF-κB," was by far the most significant enrichment for all three molecules, with fold enrichments greater than or equal to 9.66 and adjusted p less than or equal to 1e-29 in all cases.

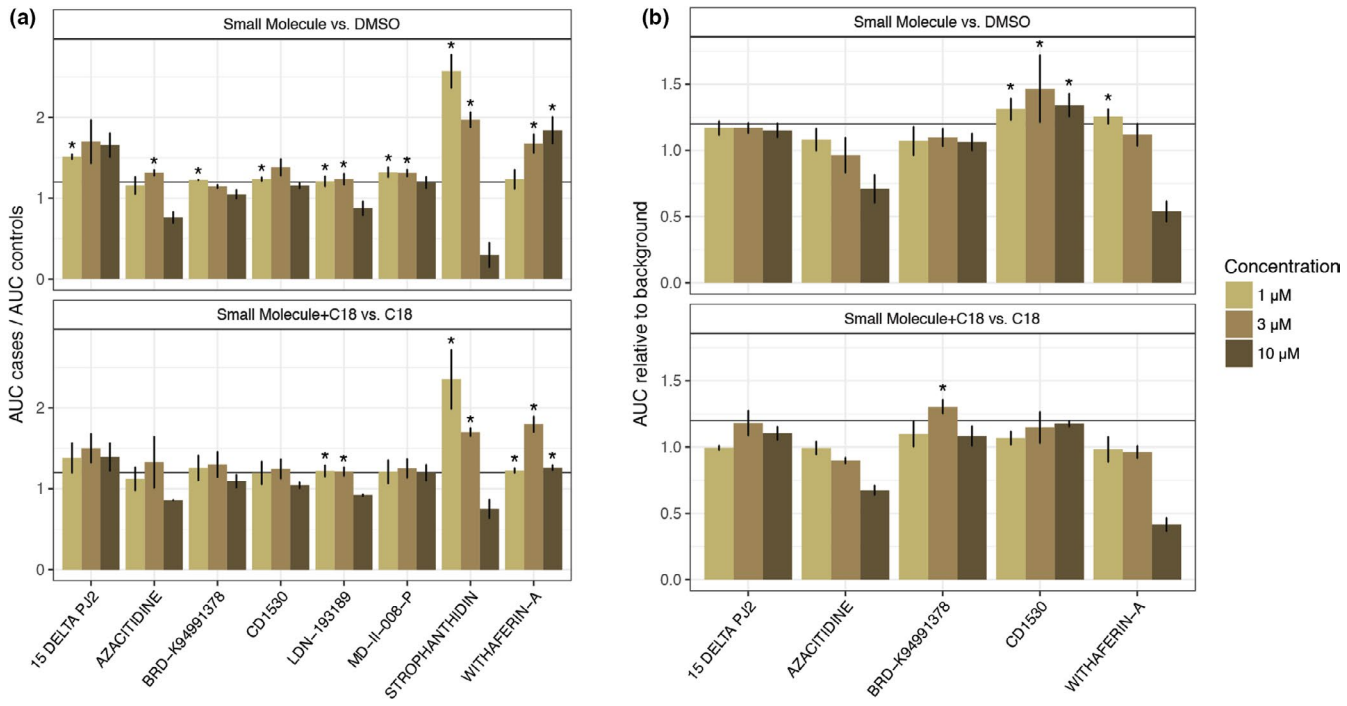

FIGURE 3 (a) Eight molecules significantly increase chloride conductance in CFBE cells. The ratio of the AUC of small molecule-treated cells compared to the AUC of control-treated cells are displayed for each of three concentrations. For the top panel, the controls (and hence the denominator) are AUCs from DMSO, whereas the bottom panel denominator comes from corrector compound C18-treated cells. (b) Three small molecules significantly increase chloride conductance in primary CF airway epithelial cells. The ratio of the AUC of small molecule-treated cells compared to the AUC of control-treated cells are displayed for each of three concentrations. For the top panel, the controls (and hence the denominator) are AUCs from DMSO, while the bottom panel denominator comes from C18-treated cells. In all figures, the horizontal bar at 1.2 indicates our minimum threshold to be considered a hit (1.0 corresponds to no effect), and (*) indicates both an adjusted p less than 0.05 and AUC ratio greater than 1.2. AUC, area under the curve, CF, cystic fibrosis

DISCUSSION

In this work, we describe an integrative bioinformatic pipeline to prioritize small molecules with predicted efficacy in restoring function to mutant ΔF508-CFTR. The pipeline was first validated computationally and then experimentally in both CFBE cells and primary CF cells. We identified 8 of 120 molecules (6.7%) with modest but significant efficacy in CFBE cells, and 3 of these with evidence of efficacy in primary CF epithelial cells.

Analysis of the molecules' chemogenomic profiles led us to identify eight pathways that were significantly differentially regulated between hits versus inactive molecules. In particular, this highlights PERK-mediated UPR and/or TNFα signaling as candidate pathways that may be a common mechanism among the CFBE hits. The connections between the UPR and CFTR rescue are well-recognized,[12,35–37] and the activation of UPR is perhaps not that surprising given the prominence of UPR in both the *RNAi* and low temperature signatures, as discussed in ref. 19. Regarding TNFα signaling, there is recent evidence that modulation of this pathway can positively affect ΔF508-CFTR trafficking and function.[38]

CD1530 had the most robust effects in primary cells, demonstrating significant improvements in chloride conductance at all three concentrations compared with DMSO. CD1530 is a retinoic acid receptor-gamma (RARγ) agonist. Notably, RARγ has been previously associated with TNFα/NF-κB signaling.[39,40] Withaferin-A, also called Ashwagandha or Indian winter cherry, is a natural compound traditionally used in Ayurvedic medicine for a variety of purposes. Withaferin-A has multiple bioactivities, including Hsp90 inhibition and induction of endoplasmic reticulum (ER) stress through proteasome inhibition.[41] Interestingly, Withaferin-A also exhibits antimicrobial activity against multiple bacterial strains, including *P. aeruginosa*, the most common bacterial pathogen affecting individuals with CF.[42] BRD-K94991378 (also called BRD1378) was initially identified in a screen for inducers of reactive oxygen species (Pubchem Bioassay ID 624156). An increase in reactive oxygen species can cause ER stress and subsequently UPR[43,44] and hence may explain the rescue observed here.

The number of effective molecules observed in the chloride conductance assay was generally greater in CFBE cells compared with primary CF cells, which is expected for this assay. This observation makes particular sense here in light of the

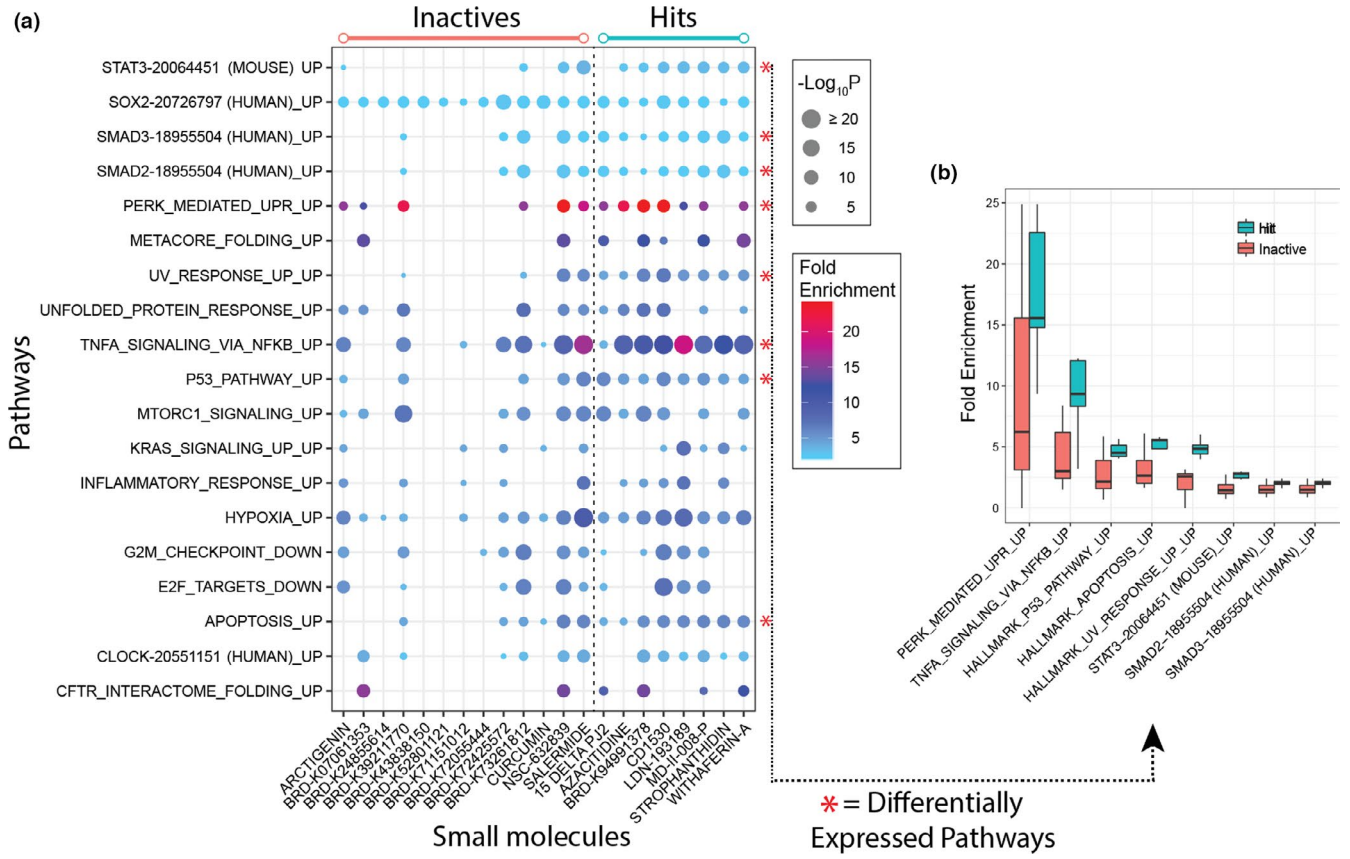

FIGURE 4 Enrichment analysis of chemogenomic profiles from the eight CFBE hits compared with inactive small molecules. (a) Dot plot showing significance and fold enrichment of pathways with significant (adjusted $p < 0.05$) enrichment in at least two of the CFBE hits. A red asterisk (*) next to the pathway name indicates significant difference in fold enrichment scores between hits and inactives (adjusted $p < 0.05$). (b) Boxplots showing fold enrichment scores for pathways from panel a with significant differences

evidence for activation of UPR. As discussed above, the UPR has been shown to downregulate expression of endogenous ΔF508-CFTR (relevant in primary cells). However, the CFBE cell line used in the primary screen stably expresses ΔF508-CFTR cDNA under the control of a cytomegalovirus promoter, which to the best of our knowledge is not affected by UPR.

Statistical evaluation of our rankings revealed several findings. First, we found that the two chemogenomic databases yielded small molecule selections with significant overlap, demonstrating some degree of consistency between databases. Second, our integrative approach showed generally better performance than scoring strategies based on individual signatures or pathways, suggesting that data integration may be a fruitful path forward in future connectivity mapping endeavors. Finally, we found that when using the CF disease signature, seeking both positive and negative connections to small molecule signatures performed better than simply seeking negative connections. We emphasize, however, that these results are for one specific disease context and are based on a relatively small number of molecules in the validation set, which may or may not share a common mechanism.

Our small molecule scoring approach has some limitations. First, we did not consider all possible connectivity mapping metrics (e.g., Zhang et al.[45]), which may have led to an even better ranking. Second, whereas the signature and pathway integration approach used was straightforward, it did not take into account pathway crosstalk. However, we believe that identifying the maximal connectivity across multiple pathways is a rational approach that would rank highly those small molecules which induce complex effects on multiple disease-relevant pathways, as long as at least one pathway has strong transcriptional connectivity.

In summary, we used an innovative informatic approach incorporating disease-associated transcriptional signatures and pathways to query the two primary chemogenomic databases. This approach identified small molecules with therapeutic activity in cystic fibrosis epithelia, partially rescuing ΔF508-CFTR function. This work paves the way for novel, integrative connectivity mapping approaches and sound validation strategies for other diseases, demonstrating its power for lead identification and mechanistic insight.

ACKNOWLEDGEMENTS

Thanks to Alessandro Laganá and Christine Becker for their feedback on analysis of small molecule mechanisms. We thank Jennifer Bartlett, Miguel Ortiz, and Laura

Marquez-Loza for critically reviewing the manuscript. We acknowledge the support of the In Vitro Models and Cell Culture Core.

CONFLICT OF INTEREST

Joel T. Dudley owns equity in NuMedii Inc., Ayasdi Inc., and LAM Therapeutics. All other authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS

R.A.H., M.D.S., J.T.D., and P.B.M. wrote the manuscript. R.A.H., M.D.S., E.A.M., D.Y.B., R.J.B., J.T.D., S.R., and P.B.M. designed the research. R.A.H., E.A.M., S.R., and D.Y.B. performed the research. R.A.H., M.D.S., D.Y.B., S.R., and P.B.M. analyzed the data.

DATA AVAILABILITY

The code has been placed in a public repository at https://github.com/rhodos/cfdrug. The data required to run the code is located at https://drive.google.com/file/d/1QqSKSmbhC1j291pd-OENByAOrYccxd5a/view and has been linked in the README.

REFERENCES

1. Riordan JR, Kerem B, Alon N, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science*. 1989;245:1066-1073.
2. Denning GM, Anderson MP, Amara JF, et al. Processing of mutant cystic fibrosis transmembrane conductance regulator is temperature-sensitive. *Nature*. 1992;358:761-764.
3. Wainwright CE, Elborn JS, Ramsey BW, et al. Lumacaftor-ivacaftor in patients with cystic fibrosis homozygous for Phe508del CFTR. *N Engl J Med*. 2015;373:220-231.
4. Keating D, Marigowda G, Burr L, et al. VX-445-tezacaftor-ivacaftor in patients with cystic fibrosis and one or two Phe508del alleles. *N Engl J Med*. 2018;379:1612-1620.
5. Douros K, Loukou I, Doudounakis S, Tzetis M, Priftis KN, Kanavakis E. Asthma and pulmonary function abnormalities in heterozygotes for cystic fibrosis transmembrane regulator gene mutations. *Int J Clin Exp Med*. 2008;1:345-349.
6. Heijerman HGM, McKone EF, Downey DG, et al. Efficacy and safety of the elexacaftor plus tezacaftor plus ivacaftor combination regimen in people with cystic fibrosis homozygous for the F508del mutation: a double-blind, randomised, phase 3 trial. *Lancet*. 2019;394:1940-1948.
7. Bessonova L, Volkova N, Higgins M, et al. Data from the US and UK cystic fibrosis registries support disease modification by CFTR modulation with ivacaftor. *Thorax*. 2018;73:731-740.
8. Volkova N, Moy K, Evans J, et al. Disease progression in patients with cystic fibrosis treated with ivacaftor: data from national US and UK registries. *J Cyst Fibros*. 2020;19:68-79.
9. Sawicki GS, McKone EF, Pasta DJ, et al. Sustained benefit from ivacaftor demonstrated by combining clinical trial and cystic fibrosis patient registry data. *Am J Respir Crit Care Med*. 2015;192:836-842.
10. Wang X, Venable J, LaPointe P, et al. Hsp90 cochaperone Aha1 downregulation rescues misfolding of CFTR in cystic fibrosis. *Cell*. 2006;127:803-815.
11. Ramachandran S, Osterhaus SR, Parekh KR, Jacobi AM, Behlke MA, McCray PB. SYVN1, NEDD8, and FBXO2 regulate ΔF508-CFTR ubiquitin-mediated proteasomal degradation. *J Biol Chem*. 2016;M116:754283.
12. Ramachandran S, Karp PH, Jiang P, et al. A microRNA network regulates expression and biosynthesis of wild-type and ΔF508 mutant cystic fibrosis transmembrane conductance regulator. *Proc Natl Acad Sci USA*. 2012;109:13362-13367.
13. Kumar P, Bhattacharyya S, Peters KW, et al. MiR-16 rescues ΔF508-CFTR function in native cystic fibrosis epithelial cells. *Gene Ther*. 2015;22:908.
14. Lamb J, Crawford ED, Peck D, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science*. 2006;313:1929-1935.
15. Subramanian A, Narayan R, Corsello SM, et al. A next generation Connectivity Map: L1000 platform and the first 1,000,000 profiles. *Cell*. 2017;171(6):1437-1452.
16. Wang Y, Yella JK, Ghandikota S, et al. Pan-transcriptome-based candidate therapeutic discovery for idiopathic pulmonary fibrosis. *Ther Adv Respir Dis* 2020;14:1753466620971143.
17. Peyvandipour A, Saberian N, Shafi A, Donato M, Draghici S. A novel computational approach for drug repurposing using systems biology. *Bioinformatics*. 2018;34:2817-2825.
18. Fortney K, Griesman J, Kotlyar M, et al. Prioritizing therapeutics for lung cancer: an integrative meta-analysis of cancer gene signatures and chemogenomic data. *PLoS Comput Biol*. 2015;11:e1004068.
19. Hodos RA, Strub MD, Ramachandran S, et al. Integrative genomic meta-analysis reveals novel molecular insights into cystic fibrosis and DeltaF508-CFTR rescue. *Sci Rep*. 2020;10:20553.
20. Wright JM, Nikolsky Y, Serebryiskaya T, Wetmore DR. MetaMiner (CF): a disease-oriented bioinformatics analysis environment. *Methods Mol Biol*. 2009;563:353–367.
21. Ramachandran S, Osterhaus SR, Karp PH, Welsh MJ, McCray PB Jr. A genomic signature approach to rescue ΔF508-cystic fibrosis transmembrane conductance regulator biosynthesis and function. *Am J Resp Cell Mol Biol*. 2014;51:354-362.
22. Pesce E, Gorrieri G, Sirci F, et al. Evaluation of a systems biology approach to identify pharmacological correctors of the mutant CFTR chloride channel. *J Cyst Fibros*. 2016;15:425-435.
23. Kunzelmann K, Schwiebert EM, Zeitlin PL, et al. An immortalized cystic fibrosis tracheal epithelial cell line homozygous for the ΔF508-CFTR mutation. *Am J Resp Cell Mol Biol*. 1993;8:522-529.
24. Malcomson B, Wilson H, Veglia E, et al. Connectivity mapping (ssCMap) to predict A20-inducing drugs and their antiinflammatory action in cystic fibrosis. *Proc Natl Acad Sci USA*. 2016;113:E3725-3734.
25. Zhang D, Ciciriello F, Anjos SM, et al. Ouabain mimics low temperature rescue of ΔF508-CFTR in cystic fibrosis epithelial cells. *Front Pharmacol*. 2012;3:176.
26. Iorio F, Tagliaferri R, Bernardo DD. Identifying network of drug mode of action by gene expression profiling. *J Comput Biol*. 2009;16:241-251.
27. Cheng J, Yang L, Kumar V, Agarwal P. Systematic evaluation of connectivity map for disease indications. *Genome Med*. 2014;6:540.

28. Wilke M, Bot A, Jorna H, Scholte BJ, De Jonge HR. Rescue of murine ΔF508-CFTR activity in native intestine by low temperature and proteasome inhibitors. *PLoS One*. 2012;7:e52070.
29. Liu Z, Borlak J, Tong W. Deciphering miRNA transcription factor feed-forward loops to identify drug repurposing candidates for cystic fibrosis. *Genome Med*. 2014;6:94.
30. Hutt DM, Herman D, Rodrigues APC, et al. Reduced histone deacetylase 7 activity restores function to misfolded CFTR in cystic fibrosis. *Nat Chem Biol*. 2010;6:25-33.
31. Carlile GW, Robert R, Zhang D, et al. Correctors of protein trafficking defects identified by a novel high-throughput screening assay. *ChemBioChem*. 2007;8:1012-1020.
32. Antigny F, Norez C, Cantereau A, Becq F, Vandebrouck C. Abnormal spatial diffusion of Ca^{2+} in ΔF508-CFTR airway epithelial cells. *Respir Res*. 2008;9:70.
33. Mendoza J, Schmidt A, Li Q, et al. Requirements for efficient correction of ΔF508-CFTR revealed by analyses of evolved sequences. *Cell*. 2012;148:164-174.
34. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc Ser B (Methodol)*. 1995;57(1):289-300.
35. Gee HY, Noh SH, Tang BL, Kim KH, Lee MG. Rescue of ΔF508-CFTR trafficking via a GRASP-dependent unconventional secretion pathway. *Cell*. 2011;146:746-760.
36. Gomes-Alves P, Couto F, Pesquita C, Coelho AV, Penque D. Rescue of ΔF508-CFTR by RXR motif inactivation triggers proteome modulation associated with the unfolded protein response. *Biochim Biophys Acta*. 2010;1804:856-865.
37. Singh OV, Pollard HB, Zeitlin PL. Chemical rescue of ΔF508-CFTR mimics genetic repair in cystic fibrosis bronchial epithelial cells. *Mol Cell Proteomics*. 2008;7:1099-1110.
38. Bitam S, Pranke I, Hollenhorst M, et al. An unexpected effect of TNF-α on ΔF508-CFTR maturation and function. *F1000 Research*. 2015;4:218.
39. Xu Q, Jitkaew S, Choksi S, et al. The cytoplasmic nuclear receptor RARγ controls RIP1 initiated cell death when cIAP activity is inhibited. *Nat Commun*. 2017;8:425.
40. Huang G-L, Luo QI, Rui G, et al. Oncogenic activity of retinoic acid receptor γ is exhibited through activation of the Akt/NF-κB and Wnt/β-catenin pathways in cholangiocarcinoma. *Mol Cell Biol*. 2013;33:3416-3425.
41. Berghe WV, Sabbe L, Kaileh M, Haegeman G, Heyninck K. Molecular insight in the multifunctional activities of withaferin A. *Biochem Pharmacol*. 2012;84:1282-1291.
42. Sundaram S, Dwivedi P, Purwar S. In vitro evaluation of antibacterial activities of crude extracts of withania somnifera (Ashwagandha) to bacterial pathogens. *Asian J Biotechnol*. 2011;3:194-199.
43. Santos CX, Tanaka LY, Wosniak J Jr, Laurindo FR. Mechanisms and implications of reactive oxygen species generation during the unfolded protein response: roles of endoplasmic reticulum oxidoreductases, mitochondrial electron transport, and NADPH oxidase. *Antioxid Redox Signal*. 2009;11:2409-2427.
44. Ali I, Shah SZA, Jin YI, et al. Reactive oxygen species-mediated unfolded protein response pathways in preimplantation embryos. *J Vet Sci*. 2017;18:1-9.
45. Zhang SD, Gant TW. A simple and robust method for connecting small-molecule drugs using gene-expression signatures. *BMC Bioinform*. 2008;9:258.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Hodos RA, Strub MD, Ramachandran S, et al. Integrative chemogenomic analysis identifies small molecules that partially rescue ΔF508-CFTR for cystic fibrosis. *CPT Pharmacometrics Syst. Pharmacol.* 2021;10:500–510. https://doi.org/10.1002/psp4.12626