Genomics 106 (2015) 268–277

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Transcriptome meta-analysis reveals common differential and global gene expression profiles in cystic fibrosis and other respiratory disorders and identifies CFTR regulators

Luka A. Clarke [a,*], Hugo M. Botelho [a], Lisete Sousa [b], Andre O. Falcao [c], Margarida D. Amaral [a]

[a] University of Lisboa, Faculty of Sciences, BioISI– Biosystems & Integrative Sciences Institute, Campo Grande, Lisboa, Portugal
[b] University of Lisboa, Faculty of Sciences, DEIO and CEAUL, Portugal
[c] University of Lisboa, Faculty of Sciences, Department of Informatics, Portugal

ARTICLE INFO

Article history:
Received 1 June 2015
Received in revised form 13 July 2015
Accepted 14 July 2015
Available online 29 July 2015

Keywords:
Cystic fibrosis
respiratory disease
differentiation, transcriptome
meta-analysis

ABSTRACT

A meta-analysis of 13 independent microarray data sets was performed and gene expression profiles from cystic fibrosis (CF), similar disorders (COPD: chronic obstructive pulmonary disease, IPF: idiopathic pulmonary fibrosis, asthma), environmental conditions (smoking, epithelial injury), related cellular processes (epithelial differentiation/regeneration), and non-respiratory "control" conditions (schizophrenia, dieting), were compared. Similarity among differentially expressed (DE) gene lists was assessed using a permutation test, and a clustergram was constructed, identifying common gene markers. Global gene expression values were standardized using a novel approach, revealing that similarities between independent data sets run deeper than shared DE genes. Correlation of gene expression values identified putative gene regulators of the CF transmembrane conductance regulator (CFTR) gene, of potential therapeutic significance. Our study provides a novel perspective on CF epithelial gene expression in the context of other lung disorders and conditions, and highlights the contribution of differentiation/EMT and injury to gene signatures of respiratory disease.

1. Introduction

A legacy of standardized microarray data reporting [11] is an ever-growing archive of transcriptomic data sets which are available for reanalysis and comparison, thereby enabling integration of data from diverse sources that, by themselves, might not provide a full picture of underlying biological processes. Recent meta-analyses of cystic fibrosis (CF)-related gene expression [13,20] suggest that identification of underlying similarities or specific gene signatures in closely related studies can be thwarted by variability in source materials, distinct methodologies and different microarray platforms. These limitations echo studies in cancer, where the plasticity of molecular signatures is a well described phenomenon [45] that has delayed the implementation of their diagnostic/prognostic potential [55]. Nevertheless, comparison of global gene expression data in related tissues under different conditions might provide valuable insights into the shared pathological pathways and regulatory networks dysregulated in individual diseases.

CF, an autosomal recessive genetic disorder caused primarily by mutations in the gene encoding the CF transmembrane conductance regulator (CFTR) protein [43], is characterized by a progressive lung disease brought about by dehydration of the epithelial airway surface liquid (ASL) and a failure of mucociliary clearance, leading to recurrent bacterial infections and a chronic inflammatory response [16]. Some clinical features of CF are shared by obstructive lung diseases such as chronic obstructive pulmonary disorder (COPD: [42]) and asthma [39], with airway surface dehydration [33] and chronic inflammation [3] implicated in the pathology of all three of these disorders. The interstitial lung disease idiopathic pulmonary fibrosis (IPF) is also a chronic inflammatory condition with a fibrotic component, like CF [15]. Furthermore, recent studies have shown that a CF-like disorder can be induced by exposure to cigarette smoke, which causes CFTR protein to be internalized from the apical membrane of airway epithelium [14]. This "acquired" CF-like condition can be an aggravating or confounding factor in the diseases mentioned above [41]. It is therefore clear that while CF itself is brought about by absence of functional CFTR protein, many of its symptoms are generalized for a variety of lung conditions, since compensatory changes in gene expression downstream of the loss of the functional CFTR chloride channel are masked by secondary changes related to the progression of CF-related lung disease. Meta-analysis of a group of individual microarray studies might allow us to assess the affinities between respiratory disorders at the level of the transcriptome and the respective regulatory networks.

In this study we assessed the similarity of independent gene expression microarray data sets from lung diseases and related conditions,

* Corresponding author.

http://dx.doi.org/10.1016/j.ygeno.2015.07.005
0888-7543/

using our own data from CF nasal epithelium [13] as a starting point, and including data from 12 other studies, including non-respiratory conditions as "negative controls". We examined the similarity between lists of differentially expressed (DE) genes using a permutation test, and identified a cluster of such genes common to independent data sets. We then used a novel data standardization approach and multidimensional scaling to assess the similarity of global gene expression and the overall relative distances between data sets and their constituent samples. Finally, by correlating the standardized expression of the CFTR gene with that of all other genes common to all samples we were able to identify potential regulators of F508del-CFTR expression, and perform validation using siRNA knockdown. Our data suggest that a significant part of the changes in gene expression observed in the CF airway epithelium are common to other respiratory diseases and represent shared processes like epithelial–mesenchymal transition (EMT), injury, and inflammation.

2. Results

2.1. Data

The 13 microarray data sets chosen for reanalysis are summarized in Table 1. Besides our own nasal epithelial cell study [13], we chose 2 other CF data sets, one measuring gene expression in bronchial epithelial brushings [38], and one in fetal tracheal cell lines [56]. We added other studies measuring gene expression in chronic lung diseases, namely chronic obstructive pulmonary disorder (COPD: [7]), idiopathic pulmonary fibrosis (IPF: [35]) and asthma [29], with the first two of these studies using whole lung tissues, and the last airway epithelial brushings. The effect of smoking in human bronchial epithelial brushings was represented by two studies [31,50], included because of recent findings that cigarette smoke exposure can induce a CF like disorder [14]. Data sets representing airway epithelial cell differentiation following exposure of bronchial epithelial cell lines to an air–liquid interface [40,46] and repair of airway epithelial injury [22] were included to investigate the contribution of these components in airway disease transcriptomes, and because regeneration of epithelia in CF has been found to be "delayed and abnormal" [19]. Two studies on unrelated disorders were chosen to function as "negative controls", i.e., to detect potential similarities between data sets with little functional relevance to each other. These investigated the effect of schizophrenia on gene expression in brain tissue [28] and the effect of dieting on gene expression in adipose tissue [37]. The thirteen studies used four different microarray platforms, of which three were from Affymetrix (HsAirway_a520108F, HG-U133A and HG-U133_Plus2) and one from Illumina (HumanRef-8 v1).

2.2. Comparison of reanalyzed gene lists

Differentially expressed genes obtained following reanalysis of 13 multiple independent studies (Table 1) were identified (Supplementary File 1) and a clustergram of genes appearing in 5 or more separate lists (excluding the two "negative control" studies) was constructed (Fig. 1), with the direction of differential gene expression categorized as either "CF-like" (for disease phenotypes and undifferentiated/injured cells) or "Control-like" (for control phenotypes and differentiated cells). The clustergram includes 74 DE genes which may represent generalized markers of chronic inflammatory lung disease, including but not exclusively CF. Twenty seven of these genes were more often up-regulated in "CF-like" phenotypes, and significant ($p < 0.05$) numbers of them were associated with the inflammatory response ($n = 9$), wounding ($n = 10$), defense response ($n = 10$) and regulation of cell proliferation ($n = 10$) GO terms (DAVID: [25]), were expressed in the extracellular space ($n = 11$), or possessed cytokine activity ($n = 7$). Forty-three clustergram genes were more highly expressed in "Control-like" phenotypes (i.e., down-regulated in disease), and included significant ($p < 0.05$) association of genes with antigen presentation ($n = 5$), immune response ($n = 8$), and MHC Class II protein complex ($n = 4$) GO terms. Three genes present on the clustergram, including S100A8, an inflammation marker, formerly known as one half of the "CF antigen" [5] were up-regulated equally in "CF-like" and "Control-like" conditions. Many of the disease markers identified in this comparison are hallmarks of epithelial differentiation [40,46] or injury [22], as indicated by their inclusion in gene lists from those studies, and a striking number of up-regulated genes, e.g., in CF epithelium, are associated with the undifferentiated phenotype.

2.3. Gene list overlap permutation analysis

The probabilities that observed DE gene list overlaps between independent microarray studies were greater than would be expected by chance (with "expected" overlaps being calculated in a permutation test using 1000 randomly generated pairs of gene lists per comparison) and were calculated for all studies against our CF nasal cell data set [13], and are presented in Table 2. Examination of the data reveals a tendency for DE gene lists from other relevant studies to show a significant overlap with the CF gene list, for similar phenotypes: for example, the observed overlaps between [13]_UP and up-regulated gene lists from CF ([38]: 21 genes; [56]: 11 genes), COPD ([7]: 8 genes), IPF ([35]: 8 genes), smoking ([31]: 13 genes), injury ([22]: 12 genes) and the two gene lists representing the undifferentiated epithelium ([46]: 18 genes; [40]: 16 genes) are significantly greater than expected at the highest level of significance. These data are also mirrored by overlaps of down-regulated genes: namely, there are observed overlaps at the highest possible level of significance between [13]_DOWN and down-regulated gene lists from CF ([38]: 15 genes), smoking ([50]: 9 genes; [31]: 17 genes), Injury ([22]: 60 genes), and gene lists representing the differentiated epithelium ([46]: 50 genes; [40]: 153 genes), these latter figures suggesting a major role for injury and differentiation genes in determining the transcriptomic profile of the CF epithelium. The full results of the permutation test for all other

Table 1
Summary of datasets included in present study.Abbreviations of first author initials (second column) are used to designate studies in Figs. 1–3.

Study reference	Abbrev.	Data accession	Array	Tissue	Group *N* and comparison
Clarke et al. [13]	C	GSE40445	HsAirway	Human native nasal epithelial brushings	5 vs. 5 (CF/non-CF)
Ogilvie et al. [38]	O	E-MTAB-360	Illumina	Human native bronchial epithelial brushings	8 vs. 16 (CF/non-CF)
Verhaeghe et al. [56]	V	E-MEXP-980	U133 Plus 2	Human fetal tracheal cell lines	3 vs. 3 (CFT-2/NT-T)
Bhattacharya et al. [7]	B	E-GEOD-8581	U133 Plus 2	Lung tissue	13 vs. 19 (COPD/control)
Spira et al. [50]	S	GSE994	U133A	Human bronchial airway brushings	19 vs. 7 (current/never smokers)
Leopold et al. [31]	L	GSE16696	U133 Plus 2	Human bronchial airway brushings	20 vs. 16 (smokers/non-smokers)
Kicic et al. [29]	K	GSE18965	U133A	AEC bronchial brushed and cultured	8 vs. 6 (asthma/control)
Meltzer et al. [35]	M	E-GEOD-24206	U133 Plus 2	Whole lung tissue	17 vs. 4 (IPF/control)
Heguy et al. [22]	H	GSE5372	U133Plus 2	Bronchial brushings	7 vs. 9 (day 7 injury/day 0 resting)
Ross et al. [46]	R	E-GEOD-5264	U133 Plus 2	Differentiating primary HBEs cultured at ALI	3 vs. 3 (day 28 ALI/day 0)
Pezzulo et al. [40]	P	GSE20502	HsAirway	Primary airway/bronchial cultures	7 vs. 5 (air exposed/submerged)
Iwamoto et al. [28]	I	E-GEOD-12649	U133A	Human prefrontal cortex	14 vs. 19 (schizophrenia/normal)
Nookaew et al. [37]	N	GSE35710	U133A	Human subcutaneous adipose tissue	8 vs. 8 (after/before dieting, males only)

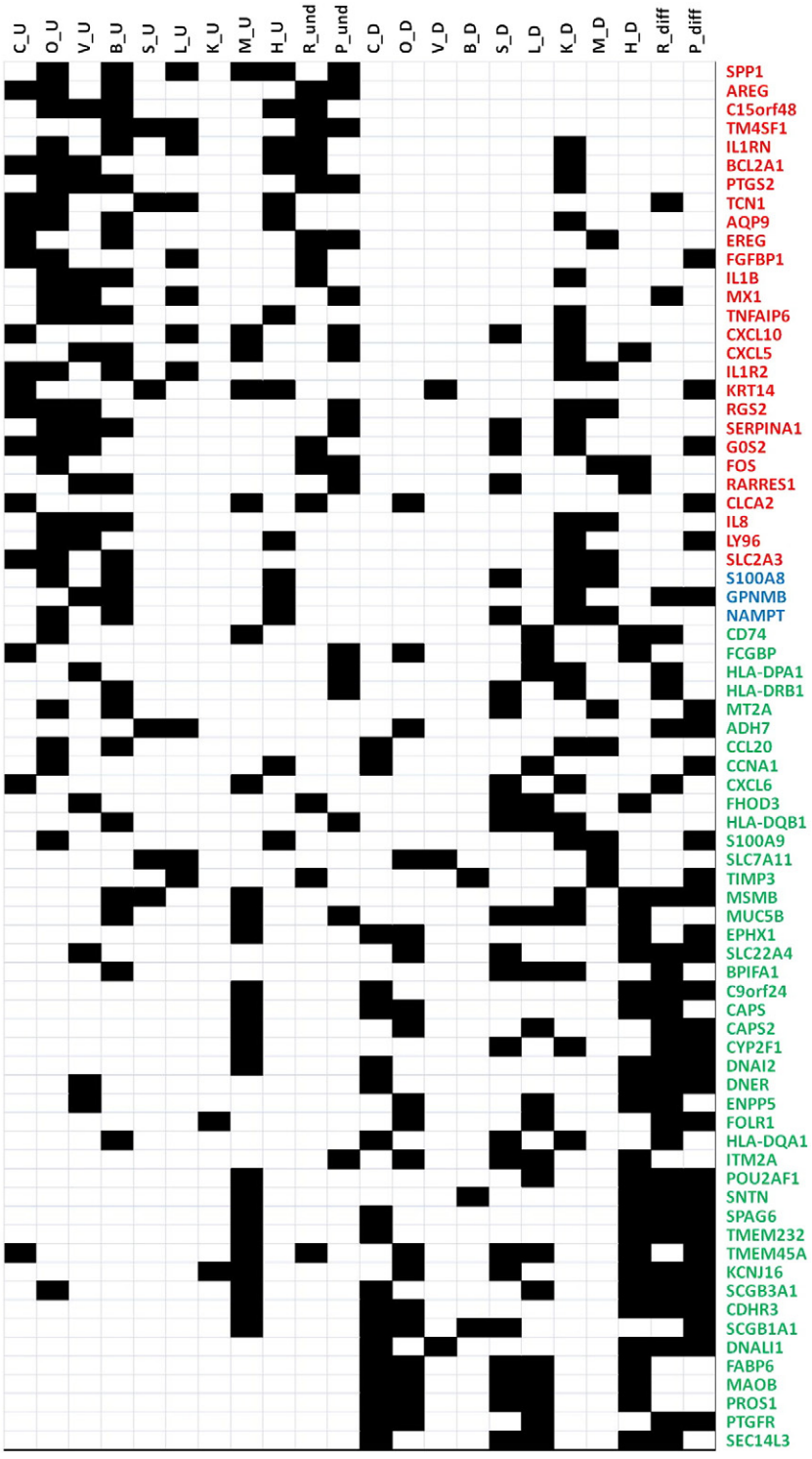

Fig. 1. Clustergram of shared DE genes. Genes found to be up- or down-regulated in at least 5 independent gene lists are clustered according to direction of regulation (U=up-regulated; D= down-regulated) with respect to "CF-like" (left hand side of diagram) or "Control-like" (right hand side of diagram) phenotypes. Studies are designated by initial of first author (see Table 1). For differentiation studies (R, P) undifferentiated cells are considered to be "CF-like" and differentiated cells "Control-like". Genes are clustered according to whether they are predominantly up-regulated in "CF-like" phenotypes (27 genes: upper section), "Control-like" phenotypes (43 genes: lower section), or is equally represented (3 genes: central section between lines).

possible comparisons are presented in Supplementary File 2, and summarized in Table 3, revealing that certain gene lists (e.g., [35]_UP (IPF) and [50]_DOWN (Smoking), which have highly significant overlaps with 16 other gene lists of any direction) were much richer in shared regulated genes than others (e.g., [29]_UP, which has no overlaps at the highest level of significance with any other gene list).

Interestingly, the data sets we chose as "negative controls", where the phenotypes and tissues under study were judged to be irrelevant to respiratory epithelial biology and therefore less likely to produce significantly greater than random overlaps [28,37], did overlap significantly with a number of experimental data sets, thereby providing a caveat for such assumptions. Nevertheless, Table 3 shows that only the asthma

Table 2
Permutation analysis of overlap between DE gene lists from 12 studies and CF nasal cell data set [13].Observed overlaps (i.e., number of identical shared genes, collapsed from multiple probesets if relevant) between gene lists shown at left (named as first author of studies described in Table 1, plus direction of differential expression) and up- or down-regulated gene lists from [13]: numbers in brackets indicate sizes of gene lists. N (perm. > obs.) indicates number of times observed overlap was exceeded in 1000 random permutations of the same overlap. P values are adjusted for multiple testing, and the maximum significance level (N (perm. > obs.) = 0) highlighted in black.

	Clarke UP (152)			Clarke DOWN (333)			
Study, direction, (*n*)	obs. overlap	N (perm. > obs.)	p val. A dj	obs. overlap	N (perm. > obs.)	p val. adj	Factor
Cystic Fibrosis							
Ogilvie UP (386)	21	0	0.0026	6	137	0.1878	CFTR [38]
Ogilvie DOWN (499)	5	41	0.0658	15	0	0.0026	
Verhaeghe UP (377)	11	0	0.0026	6	171	0.2291	CFTR [56]
Verhaeghe DOWN (340)	2	444	0.5335	5	267	0.3396	
Differentiation							
Ross UP (381) (diff)	2	510	0.6033	50	0	0.0026	DIFF [46]
Ross DOWN (333)	18	0	0.0026	3	626	0.7132	
Pezzulo UP (475) (diff)	15	0	0.0026	153	0	0.0026	DIFF [40]
Pezzulo DOWN (478)	16	0	0.0026	5	987	1.0000	
Other Respiratory							
Bhattacharya UP (191)	8	0	0.0026	7	3	0.0087	COPD [7]
Bhattacharya DOWN (180)	4	7	0.0160	1	1000	1.0000	
Kicic UP (143)	0	1000	1.0000	2	510	0.6033	ASTHMA [29]
Kicic DOWN (102)	14	0	0.0026	4	24	0.0426	
Meltzer UP (379)	8	0	0.0026	18	0	0.0026	IPF [35]
Meltzer DOWN (351)	11	0	0.0026	3	692	0.7714	
Spira UP (116)	4	4	0.0105	1	1000	1.0000	SMOKING [50]
Spira DOWN (104)	5	0	0.0026	9	0	0.0026	
Leopold UP (180)	13	0	0.0026	2	560	0.6525	SMOKING [31]
Leopold DOWN (304)	4	51	0.0791	17	0	0.0026	
Heguy UP (263)	12	0	0.0026	2	752	0.8323	INJURY [22]
Heguy DOWN (387)	2	531	0.6234	60	0	0.0026	
"Control"							
Iwamoto UP (103)	3	15	0.0282	5	3	0.008719	SCHITZ [28]
Iwamoto DOWN (176)	2	279	0.3519	4	177	0.235087	
Nookaew UP (99)	2	108	0.1503	2	342	0.425933	DIETING [37]
Nookaew DOWN (111)	1	1000	1.0000	3	141	0.190775	

study [29] had less total overlaps at the highest level of significance than the two "negative control" studies, mainly due to the uniqueness of the up-regulated gene list from that study. A schematic diagram of highly significant overlaps among all 13 studies is presented in Fig. 2, in which it can be seen that there is a high level of overlap in both up- and down-regulated gene lists between roughly half of the studies, namely the in vivo CF, smoking, injury and differentiation studies, with the "controls", asthma, IPF, COPD and cell line CF studies comprising a less connected half of the diagram.

2.4. Standardized global gene expression data

Standardized gene expression values for 7837 genes (including CFTR) common to all 13 re-analyzed studies were generated by a novel method (see Supplementary File 3). The average positions of samples from reanalyzed studies projected in 2D space based on MDS applied to standardized global gene expression values are shown in Fig. 3. The samples were divided into control (A) and treatment (B) groups, and divided into subgroups (a/b suffixes) if clear sample

Table 3
Summary of DE gene list overlaps (permutation test): number of overlaps in any direction which were significant at highest level (compiled from Supplementary File 2).

Study	UP	DOWN	Total
Meltzer et al. [35]	16	13	29
Pezzulo et al. [40]	14	12	26
Spira et al. [50]	5	16	21
Ogilvie et al. [38]	13	7	20
Leopold et al. [31]	9	11	20
Bhattacharya et al. [7]	15	4	19
Clarke et al. [13]	12	7	19
Ross et al. [46]	8	11	19
Verhaeghe et al. [56]	14	3	17
Heguy et al. [22]	10	7	17
Iwamoto et al. [28]	8	5	13
Nookaew et al. [37]	5	7	12
Kicic et al. [29]	0	10	10

heterogeneity was observed (e.g., samples from bronchial cells — [38] — lying in two locations are designated Oa and Ob). Arrows (C) show movement through 2D space from mean position of control samples to mean position of treated samples, for each study. Some similarities in the position and movement of samples in this space were noted. For example, global CF-related gene expression in nasal cells [13] produced similar positions, and therefore shared overall movement through 2D space between control and treated groups, with COPD [7] and epithelial injury [22], while the undifferentiated samples in the two differentiation studies [40,46] were located in the same position as the disease samples of these three studies, in agreement with the significant overlaps of differentially expressed genes shown in Table 2 and Supplementary File 2 for the disease/undifferentiated phenotypes. However, unlike the data on DE gene list overlaps, the positions shown here are based on global gene expression, with expression values of 7837 common genes given equal weight regardless of differential expression profile, and given the methodological differences between independent studies, the concurrences are therefore all the more remarkable. Similarities and differences between standardized data from each study are also shown in the global correlation of gene expression among all samples from all studies with each gene contributing equal weight (D), where the shading can be used to assess the influence of array type on the extent of global correlation between data sets.

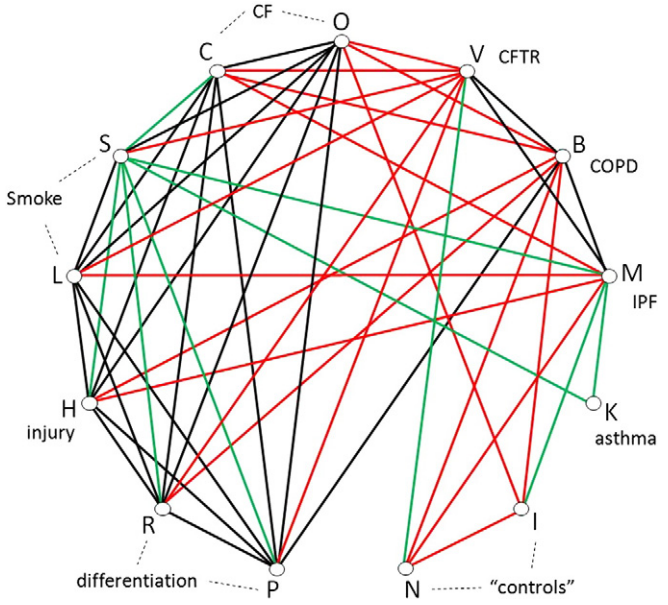

Fig. 2. Connectedness between independent microarray studies based on shared DE genes. Diagram shows whether DE gene overlap between two independent studies was significant at the highest level (lowest p value) as assessed by permutation test (black: true for both up and down-regulated gene lists; red: true for up-regulated gene lists only; green: true for down-regulated gene lists only). Studies are designated by initial of first author (see Table 1), and grouped into a more highly connected half (left) and a less connected half (right).

2.5. Correlation of standardized gene expression values

Pearson's correlation coefficients (R) for standardized expression values were calculated for all 7836 common genes vs. CFTR, in each of the 13 re-analyzed studies for all samples, and in each CF study, for F508del samples only. Weighted mean Pearson's R values (±SEM) for the 100 most highly negatively correlated genes in each case are given in Supplementary File 4. Several of the genes that negatively correlated with CFTR expression across all studies have been previously described to affect CFTR expression. The most highly negatively correlated gene, *YWHAB* (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta) positively regulates CFTR biosynthesis [32] and binds to the CFTR regulatory (R) domain [10], but is also suspected to delay airway epithelial remodeling via positive regulation of *MMP1* (matrix metalloproteinase 1) expression [2]. The next gene, *AP2M1* (adaptor-related protein complex 2, mu 1 subunit), has been identified in a high-content functional screen as a corrector of CFTR [53]. The presence of two previously identified positive CFTR regulators at the very top of our list of negative *CFTR* correlators suggests that their enhanced expression might be triggered by reduced *CFTR* expression as a compensatory mechanism. The list of genes negatively correlated with *F508del-CFTR* expression values only was also rich in genes with possible links to CFTR function or to localizations or processes that might implicate them in CFTR trafficking (see Supplementary File 4). *SNAP23* (synaptosomal-associated protein, 23 kDa), for example, is part of CFTR's SNARE interactome, and regulates CFTR gating [52], whereas *FAS* (Fas cell surface death receptor), the next gene on the list, is a known modulator of CF disease severity [30], which has been found to be crucial to regulation of CFTR-dependent apoptosis and autophagy following epithelial injury induced by cigarette smoke [8].

2.6. siRNA knockdown of potential CFTR regulators

We tested the effect of siRNA knockdown on F508del-CFTR trafficking for 73 genes whose expression was negatively correlated with either F508del-CFTR only (n = 38) or with CFTR across all studies (n = 38: 3 common genes, see Supplementary File 4). For nine of the 73 genes tested, siRNA knockdown produced appreciable increases in CFTR traffic to the plasma membrane (median Z score > 1: see Supplementary File 4 and Table 4) in our assay. For most products of these genes, a direct role in CFTR traffic regulation is hard to ascribe, although some are involved in maintaining an undifferentiated phenotype. The protein tyrosine phosphatases PTP4A1/2 have both been implicated in cell motility and invasiveness signaling [44], while NMT1 (N-myristoyltransferase 1) and MGA (MAX dimerization protein) are both involved in cell proliferation [17,26]. There is no evidence for a functional link between GNAQ (guanine nucleotide binding protein (G protein) q polypeptide) and CFTR, but PSEN1 (presenilin 1) is a member of the gamma secretase complex, and therefore may have a role in proteolysis [18]. RCN2 (reticulocalbin 2), an ER calcium binding protein, has been described as a member of the CFTR interactome [58]. Rho target PKN2 is crucial to the formation of apical cell junctions in differentiating bronchial epithelium [57], and SNX6 (sorting nexin 6) is a retromer component which regulates protein transport from the endosome to the trans-Golgi network (TGN: [24,36]).

3. Discussion

This study presents the meta-analysis of 13 microarray data sets, of which 11 are related to respiratory disease and epithelial differentiation, and functionally relevant to the study of cystic fibrosis (CF)-related gene

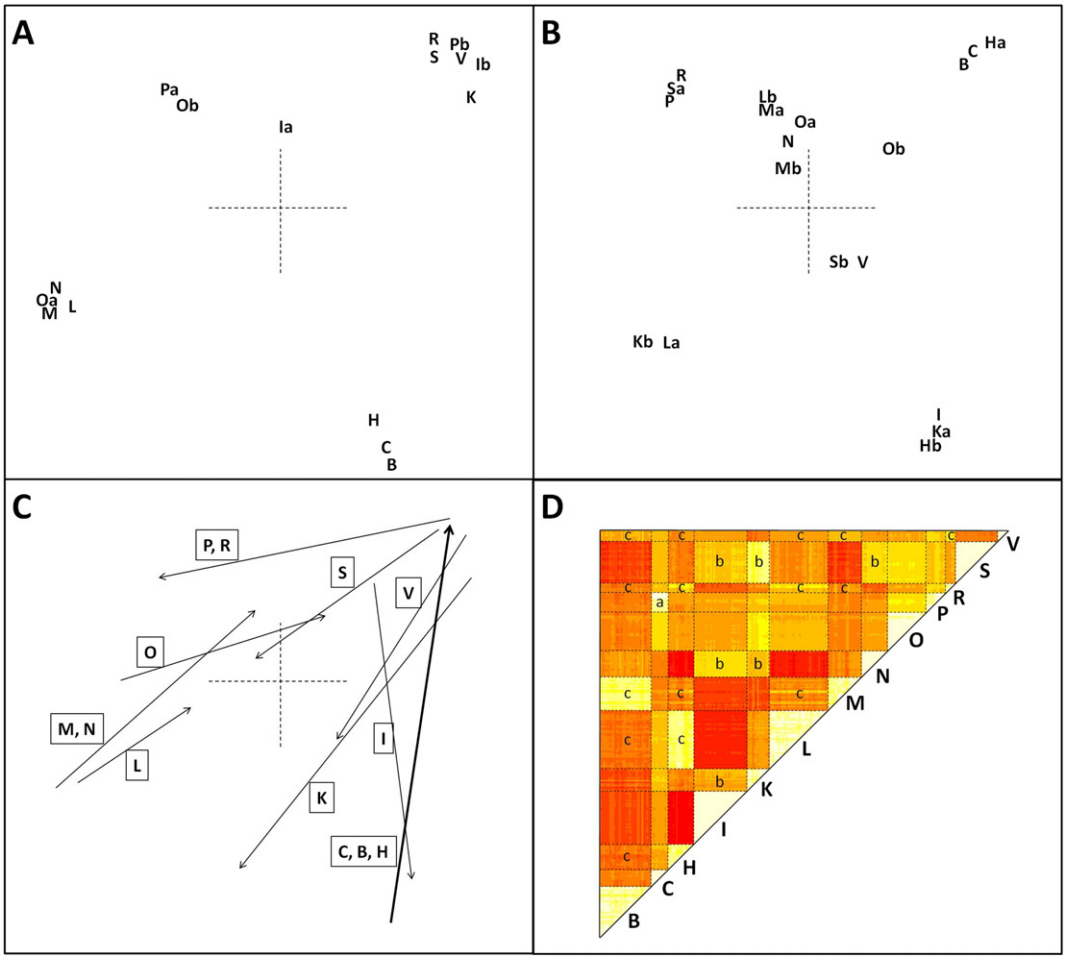

Fig. 3. Similarity of standardized global gene expression in independent microarray studies. Projection in 2D space of standardized values of expression for 7837 common genes following multidimensional scaling (MDS: A–Table 1 for abbreviations) samples were divided into control (A) and treatment (B groups), and divided into subgroups (a/b suffixes) where clear sample cluster heterogeneity was observed. Arrows (C) show movement through 2D space between overall mean positions of control and experimental samples. Effect of use of same microarray platform (a: HsAirway; b: U133A; c: U133 Plus 2) on global correlation of gene expression values among all samples is also assessed (D, where red = high, yellow = intermediate and white = low correlation). Study O used a unique microarray platform (Illumina).

expression. The other two studies were chosen for their presumed lack of a pathological basis for similarity with the other data sets and were thus used as "negative controls". Our analysis has enabled the comparison of independent microarray studies with respect to their differential gene expression and standardized global gene expression, and allowed us to identify some potential novel gene regulators of CFTR.

3.1. Clustergram — shared markers and common processes

The identity of the DE genes shared by different studies and the degree to which certain diseases/processes share them can best be appreciated in the clustergram (Fig. 1), where the genes most frequently appearing on independent gene lists are displayed. GO term enrichment analysis associated the genes up-regulated in CF-like phenotypes with functional categories highly relevant to CF lung pathology (e.g., inflammation, defense, wounding, proliferation, cytokine activity), and identified down-regulation of antigen presentation and immune response genes, as previously described for CF [20]. The general trend is for up-regulated genes to be shared by CF [13,38] and COPD [7], perhaps reflecting the pro-inflammatory airway phenotype secondary to CFTR dysfunction, as several of these DE genes are also present in injured [22] or undifferentiated cells [40,46]. On the

Table 4
Genes for which siRNA knockdown produced an appreciable increase in CFTR traffic to the plasma membrane in CFBE cells (median Z score >1: see Supplementary File 4).

Gene	Correlation list	Median Z-score	Function
PSEN1 (presenilin 1)	F508del-CFTR	2.68	Proteolysis
PTP4A2 (protein tyrosine phosphatase 4A2)	All samples	2.36	EMT
NMT1 (N-myristoyltransferase 1)	All samples	1.93	Proliferation
PKN2 (protein kinase N2)	F508del-CFTR	1.38	Apical junction formation
SNX6 (sorting nexin 6)	F508del-CFTR/All	1.28	Retromer component
RCN2 (reticulocalbin 2)	F508del-CFTR	1.24	Lung development
MGA (MAX dimerization protein)	F508del-CFTR	1.23	Transcription factor
GNAQ (guanine nucleotide binding protein (G protein), q polypeptide)	F508del-CFTR	1.20	Phospholipase C beta activation
PTP4A1 (protein tyrosine phosphatase 4A1)	F508del-CFTR	1.00	EMT

Official gene symbols, full names and broad functional area given, as defined by NCBI (http://www.ncbi.nlm.nih.gov/gene/).

other hand, a prominent set of genes down-regulated in injured and undifferentiated cells corresponds to genes which are down-regulated either in CF in vivo [13,38], or in smoking studies [31,50], suggesting the possibility that they represent genes repressed by a lack of functional CFTR protein, although the pattern of the in vitro CF study [56] does not support this trend. The clustergram also clearly identifies conditions which do not fit these tendencies: for example, many genes up-regulated in IPF [35] and down-regulated in asthma [29] invert the direction of expression seen in other conditions, namely in CF and COPD. From our data, it is not clear whether such inversions would be perpetuated in other studies of the same diseases, and therefore represent robust differences in gene expression profiles between CF/COPD and asthma or IPF.

Our data provide a novel perspective by revealing the extent to which genes or their functional groups, some of which have previously been identified as candidate marker genes for CF, are CF-specific (e.g., *RGS2*: regulator of G-protein signaling 2, *G0S2*: G0/G1 switch 2), or else are revealed as general markers of respiratory inflammation (e.g., *IL1R2*: interleukin 1 receptor, type II, *IL8*: interleukin 8). Up-regulated expression of four genes (*SPP1*: secreted phosphoprotein 1, *AREG*: amphiregulin, *C15ORF48*: chromosome 15 open reading frame 48, and *TM4SF1*: transmembrane 4L six family member 1) was found to be associated with CF-like phenotypes only, and may provide interesting clues for further investigation of a potential differentiation defect in the CF epithelium [19]. Indeed, all four of these genes have been associated with cancer lethality, progression or invasiveness [48,49,51,61], and this finding is interesting in the light of CFTR's own role in suppressing tumor progression [60]. Of the six genes associated exclusively with "Control-like" gene lists (i.e., down-regulated in disease), *SEC14L3* (SEC14-like 3 (*Saccharomyces cerevisiae*)) stands out for playing an important role in lung development [23], and is known to have an inverse relationship with allergic airway inflammation [47], whereas *DNALI1* (dynein, axonemal, light intermediate chain 1) is an essential component of motile respiratory cilia [62]. The functional significance of the other genes exclusively down-regulated in our clustergram (*PTGFR*: prostaglandin F receptor, *PROS1*: protein S (alpha), *MAOB*: monoamine oxidase B, *FABP6*: fatty acid binding protein 6, ileal) is less clear, although as a group these six genes are down-regulated in in vivo CF, smoking and epithelial injury, thereby potentially representing a gene signature linking those conditions. Finally, the DE gene which appears in the greatest number of independent gene lists (3 "CF-like", 5 "Control-like"), *TMEM45a*, has been described as a marker for keratinization [21], and its regulation in several of the studies analyzed here also suggests a role in airway epithelium differentiation.

We conclude that the differential gene expression pattern in CF is most closely related to that of COPD, smoking, injury and (de-)differentiation. Despite heterogeneity between studies, our clustergram presents a coherent molecular signature for various airway conditions that associates airway inflammation, de-differentiation and injury with up-regulation of specific genes, while hinting that loss of functional CFTR may underlie the down-regulation of others.

3.2. Permutation analysis: DE gene overlap significance

We have attempted to determine the extent to which gene lists from independent studies with phenotypic relevance to one another overlap more than predicted by randomly generated gene lists, using a permutation test. Not all overlaps between "similar" gene lists were significantly greater than random (see Table 2 and Supplementary File 2): for example, up-regulated gene lists from CF nasal cells [13] and COPD lung [7] overlapped with the highest level of significance, but down-regulated gene lists from these two studies did not. Nevertheless, many similarities were detected between independent studies which presumably signify shared biological processes. The inclusion of data sets representing epithelial injury [22] and differentiation [40,46] allowed us to identify genes involved in those processes as being associated with DE gene expression in airway disorders. In particular, the number of DE genes shared between these studies and in vivo CF data sets [13,38] is striking, and reveals a significant injury and de-differentiation or epithelial–mesenchymal transition (EMT) signature in CF epithelium, associating CFTR dysfunction with perturbed differentiation pathways [19]. This is further highlighted in Fig. 2, where we show that the in vivo CF studies are highly associated (with respect to DE gene list similarity) with smoking, injury and de-differentiation.

The two "control" studies [28,37] also overlapped significantly with other studies in certain cases (see Table 3 and Supplementary File 2), showing that many biological processes share common pathways and DE genes despite being superficially unrelated, and that these common processes may compromise attempts to identify lists of DE marker genes unique to any particular process or disease. Indeed, both "control" studies originally reported changes in gene expression, and specifically down-regulation of mitochondrial genes, both in adipose tissue following dieting [37] and in brain tissue of schizophrenic individuals [28]. Since CF is also associated with reduced mitochondrial function [54], some DE gene overlap between the "control" studies and the CF studies can therefore be expected, and the same may be true for other lung diseases [1].

3.3. Standardized global gene expression

The standardized data based on expression values of 7837 common genes across all studies and microarray platforms are shown in MDS plots in 2D-space giving an approximation of "closeness" among data sets split between the phenotypes compared (Fig. 3). It can be noted that our CF nasal cell study [13] has the same position in both control and experimental samples as COPD [7] and epithelial injury [22], suggesting that the differences between control and experimental phenotypes in these three studies therefore represent a similar global shift in the transcriptome. It is also noticeable that certain groups of samples cluster together in the MDS plot. For example, the control (undifferentiated) samples of the differentiation studies [40,46] lie in a similar position to the experimental CF [13], injury [22] and COPD [7] samples, which makes sense in the context of the permutation test results for the DE gene overlaps between those studies. Other superimpositions and overall movements are harder to interpret, including the mismatch in position between the in vivo CF studies [13,38], but it must be borne in mind that all 7837 genes were given the same weight in this analysis (see Supplementary File 3). It is therefore notable that the majority of control and experimental samples can easily be discriminated by this global measure of expression, implying that disease-related gene expression represents a global "tilt" of the transcriptome, and not just changes in a few hundred DE genes. Thus, the agreements between certain studies for similarity data based on DE gene list overlaps on the one hand and standardized global gene expression values on the other are remarkable. It is also worth noting that the influence of array type, which could have been a strong determinant of sample position, was minimized by our standardization (see Fig. 3D). Despite this, the differing and even inverted positions of some of the studies in Fig. 3 may also represent systematic differences in overall gene expression values that derive from materials and methodologies specific to independent studies, rather than reflecting underlying biological differences between conditions or airway diseases.

3.4. Identification of potential CFTR regulators

In a preliminary siRNA knockdown assay we have identified some potential negative regulators of CFTR traffic (Table 4) whose expression was highly inversely correlated with *CFTR* expression, either across all studies, or in the subset of samples expressing F508del-CFTR (Supplementary File 4). The correlation of gene expression values with those of CFTR across many experimental and control samples in several independent studies is likely to be fraught by the effects of variability, but

the presence of several proven CFTR regulators and interactome members on our lists partly validates the approach. Potential CFTR gene regulators identified by siRNA knockdown in a CFBE model of F508del-CFTR traffic rescue include the retromer component *SNX6* (sorting nexin 6: [59]), which might cross the path of wild type CFTR in the early endosome, during recycling, although negative correlation with F508del-CFTR was also found. Another negative CFTR correlator, *PSEN1* (presenilin 1), is implicated, as a potential gamma secretase component, in a protease network regulating expression of CXCR1 surface expression on neutrophils, and their IL8-mediated recruitment [4], and therefore may be of therapeutic interest in CF. *RCN2* (reticulocalbin 2) encodes a calcium binding protein in the ER lumen, where F508del-CFTR, with which it is negatively correlated, mediates the chloride conductance that is responsible for the enhanced Ca^{2+}-dependent chloride (Cl^-) conductance (CaCC) seen in CF epithelia [34]. Reduced expression of this CFTR interacting protein [58] might therefore lead to enhanced CFTR expression via interference in the CaCC feedback loop. Confirmation of these and other such interactions between gene expression of our proposed potential regulators and maturation and traffic of the F508del-CFTR protein will require much further work, and the lists we present are rich in untested potential leads, worthy of future investigation.

3.5. Conclusions

We have used archived gene expression data to extend our knowledge about the relationships between the transcriptomes of related respiratory conditions, and have thereby supplied a novel perspective on airway pathophysiology. Our meta-analysis of 13 microarray data sets has allowed us to perform a comparative analysis of CF-related differential gene expression in the context of other respiratory diseases and conditions, including smoking, and the allied processes of differentiation and injury. We have found genes associated with de-differentiation and epithelial injury to be strongly associated with the CF gene expression profile, and we have confirmed that exposure to cigarette smoke, with its down-regulation of CFTR expression, causes a gene expression profile with many affinities to that of CF. Furthermore, we show that the affinities between certain independent studies (e.g., CF, COPD and injury) hold true not only for DE genes but also for global gene expression values, suggesting that similarities between related conditions run deeper than just a few perturbed pathways. Using standardized gene expression correlation we have also been able to propose potential regulators of CFTR expression which may be of future interest in the design of novel strategies for rescue of mutant CFTR protein.

4. Materials and methods

4.1. Microarray data reanalysis

Affymetrix Genechip expression data (see Table 1) were quantile normalized in RMA Express [27] and the Ilumina data set [38] was reanalyzed with alternative tools using similar methods as previously described [13]. Prior to normalization, quality control (QC) parameters (GAPDH ratios, $\log_2$PM distributions and RLE/NUSE plots) were examined to determine sample inclusion, independent of QC decisions made by the original authors. Samples used in reanalysis are presented in Supplementary File 1. Normalized values were then analyzed using the Rank Products method ([12]; Bioconductor Package RankProd). Variable cutoffs with a starting point of pfp < 0.01, p < 0.0001, and FC > 2 were used to choose lists of differentially expressed (DE) genes from each study not exceeding 500 genes in length. The criteria used and the DE genes chosen for each study are given in Supplementary File 1. A clustergram was constructed based on genes appearing most frequently in lists of both up- and down-regulated genes. For the clustergram, phenotypes for which a gene was preferentially expressed were categorized as either "CF-like" (e.g., genes up-regulated in CF, COPD, Asthma, IPF, smoking, injury studies) or "Control-like" (e.g., genes down-regulated in those studies). For the differentiation studies [40,46], the undifferentiated and differentiated phenotypes were considered to be "CF-like" and "Control-like", respectively, so genes up-regulated in cells differentiating at the air–liquid interface (ALI) were grouped with genes down-regulated in CF and related diseases, and vice versa.

4.2. GO term enrichment analysis

We submitted gene lists to the online bioinformatic resource DAVID [25] for enrichment analysis of GO terms, making use of the following categories: BP_Fat (biological process), CC-Fat (cellular component) and MF_Fat (molecular function), against the default *Homo sapiens* background. Benjamini–Hochberg-corrected p values were used to determine significance [6].

4.3. Overlap permutation statistical analysis

In order to test the null hypothesis that the number of differentially expressed genes shared between different studies was no different from the number that would be shared by random selections of genes from the same microarrays, we developed a permutation-based methodology. Gene lists (up or down with respect to the particular phenotype relevant to each study) were collapsed to single gene symbols and the total number of genes appearing in any two compared lists ("overlap") was counted for all lists from all studies and for both directions of differential expression. The observed overlap for each pair of lists was compared with that in 1000 randomly generated gene list pairs (also collapsed to single gene symbols), with the sizes of each observed list and the relevant microarray backgrounds being taken into account for each comparison. Multiple testing adjusted p values were calculated using the Benjamini and Hochberg procedure [6].

4.4. Standardization of microarray data sets and comparison of global gene expression data

For comparison of global values of gene expression between all 13 re-analyzed studies, data sets were reduced to 7837 common genes (including CFTR) represented by varying numbers of probes on each of the four microarray platforms used in the 13 studies (Affymetrix HsAirway, HG-U133A, HG-U133A plus2, and Illumina HumanRef-8 v1). The RMA-normalized expression values for these common genes were standardized by a novel method described in detail in Supplementary File 3. We then conducted a two-dimensional projection of cases using multidimensional scaling, in order to assess global similarities between data sets. Standardized expression values as calculated above for all 7836 shared genes (excluding CFTR) were then correlated (Pearson's R) with the corresponding expression values for CFTR itself (see Supplementary File 4). This was done for all 13 studies (n = 252 samples including both experimental and control groups), and also for a subset of F508del expressing samples from the three CF studies (n = 16 samples). Representative genes of interest showing a strong negative correlation to CFTR or F508del CFTR expression were selected for siRNA knockdown assays to test their effect on F508del-CFTR traffic correction in a CFBE epithelial cell model.

4.5. siRNA knockdowns and CFTR trafficking assay

An assay to test the effect of siRNA knockdown on trafficking of F508del-CFTR traffic to the plasma membrane was performed for 73 genes (identified in Supplementary File 4) in a CF bronchial epithelial (CFBE) cell model as previously described [9]. The chosen genes were equally divided between those found to have a high mean negative

correlation with CFTR expression in all studies, and in F508del-CFTR expressing samples only.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ygeno.2015.07.005.

Author contributions

LAC designed and coordinated the study. LAC & HMB undertook experimental work. LS & AOF developed and implemented statistical methods for meta-analysis. LAC, HMB & MDA analyzed and interpreted data. LAC drafted the manuscript, and all authors contributed to the writing of the manuscript.

Disclosure declaration

The authors declare no competing interests.

Acknowledgments

This work was supported by grants from FCT/MCTES/PIDDAC (Portugal): PTDC/SAU-GMG/122299/2010 (to MDA) and center grants UID/MULTI/04046/2013 (to BioISI) and PEstOE/MAT/UI0006/2011 (to CEAUL), as well as by post-doctoral fellowship SFRH/BPD/93017/2013 (to HMB) from FCT/MCTES (Portugal).

References

[1] B. Aravamudan, M.A. Thompson, C.M. Pabelick, Y.S. Prakash, Mitochondria in lung diseases, Expert Rev. Respir. Med. 7 (2013) 631–646.
[2] N. Asdaghi, R.T. Kilani, A. Hosseini-Tabatabaei, S.O. Odemuyiwa, T.L. Hackett, D.A. Knight, A. Ghahary, R. Moqbel, Extracellular 14-3-3 from human lung epithelial cells enhances MMP-1 expression, Mol. Cell. Biochem. 360 (2012) 261–270.
[3] R. Athanazio, Airway disease: similarities and differences between asthma, COPD and bronchiectasis, Clinics (Sao Paulo) 67 (2012) 1335–1343.
[4] M. Bakele, A.S. Lotz-Havla, A. Jakowetz, M. Carevic, V. Marcos, A.C. Muntau, S.W. Gersting, D. Hartl, An interactive network of elastase, secretases, and PAR-2 protein regulates CXCR1 receptor surface expression on neutrophils, J. Biol. Chem. 289 (2014) 20516–20525.
[5] C. Barthe, C. Figarella, J. Carrère, O. Guy-Crotte, Identification of 'cystic fibrosis protein' as a complex of two calcium-binding proteins present in human cells of myeloid origin, Biochim. Biophys. Acta 1096 (1991) 175–177.
[6] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc. Ser. B Methodol. 57 (1995) 289–300.
[7] S. Bhattacharya, S. Srisuma, D.L. Demeo, S.D. Shapiro, R. Bueno, E.K. Silverman, J.J. Reilly, T.J. Mariani, Molecular biomarkers for quantitative and discrete COPD phenotypes, Am. J. Respir. Cell Mol. Biol. 40 (2009) 359–367.
[8] M. Bodas, T. Min, N. Vij, Critical role of CFTR-dependent lipid rafts in cigarette smoke-induced lung epithelial injury, Am. J. Physiol. Lung Cell. Mol. Physiol. 300 (2011) L811–L820.
[9] H.M. Botelho, I. Uliyakina, N.T. Awatade, M.C. Proença, C. Tischer, L. Sirianant, K. Kunzelmann, R. Pepperkok, M.D. Amaral, Protein traffic disorders: an effective high-throughput fluorescence microscopy pipeline for drug discovery, Sci. Rep. 5 (2015) 9038.
[10] Z. Bozoky, M. Krzeminski, R. Muhandiram, J.R. Birtley, A. Al-Zahrani, P.J. Thomas, R.A. Frizzell, R.C. Ford, J.D. Forman-Kay, Regulatory R region of the CFTR chloride channel is a dynamic integrator of phospho-dependent intra- and intermolecular interactions, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) E4427–E4436.
[11] A. Brazma, P. Hingamp, J. Quackenbush, G. Sherlock, P. Spellman, C. Stoeckert, J. Aach, W. Ansorge, C.A. Ball, H.C. Causton, et al., Minimum information about a microarray experiment (MIAME)—toward standards for microarray data, Nat. Genet. 29 (2001) 365–371.
[12] R. Breitling, P. Armengaud, A. Amtmann, P. Herzyk, Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments, FEBS Lett. 573 (2004) 83–92.
[13] L.A. Clarke, L. Sousa, C. Barreto, M.D. Amaral, Changes in transcriptome of native nasal epithelium expressing F508del-CFTR and intersecting data from comparable studies, Respir. Res. 14 (2013).
[14] L.A. Clunes, C.M. Davies, R.D. Coakley, A.A. Aleksandrov, A.G. Henderson, K.L. Zeman, E.N. Worthington, M. Gentzsch, S.M. Kreda, D. Cholon, et al., Cigarette smoke exposure induces CFTR internalization and insolubility, leading to airway surface liquid dehydration, FASEB J. 26 (2012) 533–545.
[15] H.R. Collard, B.B. Moore, K.R. Flaherty, K.K. Brown, R.J. Kaner, T.E. King, J.A. Lasky, J.E. Loyd, I. Noth, M.A. Olman, et al., Acute exacerbations of idiopathic pulmonary fibrosis, Am. J. Respir. Crit. Care Med. 176 (2007) 636–643.
[16] J.C. Davies, E.W. Alton, A. Bush, Cystic fibrosis, BMJ 335 (2007) 1255–1259.
[17] C.E. Ducker, J.J. Upson, K.J. French, C.D. Smith, Two N-myristoyltransferase isozymes play unique roles in protein myristoylation, proliferation, and apoptosis, Mol. Cancer Res. 3 (2005) 463–476.
[18] P.C. Fraering, W. Ye, J.M. Strub, G. Dolios, M.J. LaVoie, B.L. Ostaszewski, A. van Dorsselaer, R. Wang, D.J. Selkoe, M.S. Wolfe, Purification and characterization of the human gamma-secretase complex, Biochemistry 43 (2004) 9774–9789.
[19] R. Hajj, P. Lesimple, B. Nawrocki-Raby, P. Birembaut, E. Puchelle, C. Coraux, Human airway surface epithelial regeneration is delayed and abnormal in cystic fibrosis, J. Pathol. 211 (2007) 340–350.
[20] T.H. Hampton, B.A. Stanton, A novel approach to analyze gene expression data demonstrates that the DeltaF508 mutation in CFTR downregulates the antigen presentation pathway, Am. J. Physiol. Lung Cell. Mol. Physiol. 298 (2010) L473–L482.
[21] A. Hayez, J. Malaisse, E. Roegiers, M. Reynier, C. Renard, M. Haftek, V. Geenen, G. Serre, M. Simon, C.L. de Rouvroit, et al., High TMEM45A expression is correlated to epidermal keratinization, Exp. Dermatol. 23 (2014) 339–344.
[22] A. Heguy, B.G. Harvey, P.L. Leopold, I. Dolgalev, T. Raman, R.G. Crystal, Responses of the human airway epithelium transcriptome to in vivo injury, Physiol. Genomics 29 (2007) 139–148.
[23] D. Hishikawa, H. Shindou, T. Harayama, R. Ogasawara, A. Suwabe, T. Shimizu, Identification of Sec14-like 3 as a novel lipid-packing sensor in the lung, FASEB J. 27 (2013) 5131–5140.
[24] Z. Hong, Y. Yang, C. Zhang, Y. Niu, K. Li, X. Zhao, J.J. Liu, The retromer component SNX6 interacts with dynactin p150(Glued) and mediates endosome-to-TGN transport, Cell Res. 19 (2009) 1334–1349.
[25] dW Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nat. Protoc. 4 (2009) 44–57.
[26] P.J. Hurlin, E. Steingrìmsson, N.G. Copeland, N.A. Jenkins, R.N. Eisenman, Mga, a dual-specificity transcription factor that interacts with Max and contains a T-domain DNA-binding motif, EMBO J. 18 (1999) 7019–7028.
[27] R.A. Irizarry, B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf, T.P. Speed, Exploration, normalization, and summaries of high density oligonucleotide array probe level data, Biostatistics 4 (2003) 249–264.
[28] K. Iwamoto, M. Bundo, T. Kato, Altered expression of mitochondria-related genes in postmortem brains of patients with bipolar disorder or schizophrenia, as revealed by large-scale DNA microarray analysis, Hum. Mol. Genet. 14 (2005) 241–253.
[29] A. Kicic, T.S. Hallstrand, E.N. Sutanto, P.T. Stevens, M.S. Kobor, C. Taplin, P.D. Paré, R.P. Beyer, S.M. Stick, D.A. Knight, Decreased fibronectin production significantly contributes to dysregulated repair of asthmatic epithelium, Am. J. Respir. Crit. Care Med. 181 (2010) 889–898.
[30] V. Kumar, T. Becker, S. Jansen, A. van Barneveld, K. Boztug, S. Wölfl, B. Tümmler, F. Stanke, Expression levels of FAS are regulated through an evolutionary conserved element in intron 2, which modulates cystic fibrosis disease severity, Genes Immun. 9 (2008) 689–696.
[31] P.L. Leopold, M.J. O'Mahony, X.J. Lian, A.E. Tilley, B.G. Harvey, R.G. Crystal, Smoking is associated with shortened airway cilia, PLoS One 4 (2009) e8157.
[32] X. Liang, A.C. Da Paula, Z. Bozóky, H. Zhang, C.A. Bertrand, K.W. Peters, J.D. Forman-Kay, R.A. Frizzell, Phosphorylation-dependent 14-3-3 protein interactions regulate CFTR biogenesis, Mol. Biol. Cell 23 (2012) 996–1009.
[33] M.A. Mall, D. Hartl, CFTR: cystic fibrosis and beyond, Eur. Respir. J. 44 (2014) 1042–1054.
[34] J.R. Martins, P. Kongsuphol, E. Sammels, S. Dahimène, F. Aldehni, L.A. Clarke, R. Schreiber, H. de Smedt, M.D. Amaral, K. Kunzelmann, F508del-CFTR increases intracellular $Ca(2+)$ signaling that causes enhanced calcium-dependent $Cl(-)$ conductance in cystic fibrosis, Biochim. Biophys. Acta 1812 (2011) 1385–1392.
[35] E.B. Meltzer, W.T. Barry, T.A. D'Amico, R.D. Davis, S.S. Lin, M.W. Onaitis, L.D. Morrison, T.A. Sporn, M.P. Steele, P.W. Noble, Bayesian probit regression model for the diagnosis of pulmonary fibrosis: proof-of-principle, BMC Med. Genomics 4 (2011) 70.
[36] Y. Niu, C. Zhang, Z. Sun, Z. Hong, K. Li, D. Sun, Y. Yang, C. Tian, W. Gong, J.J. Liu, PtdIns(4)P regulates retromer–motor interaction to facilitate dynein–cargo dissociation at the trans-Golgi network, Nat. Cell Biol. 15 (2013) 417–429.
[37] I. Nookaew, P.A. Svensson, P. Jacobson, M. Jernås, M. Taube, I. Larsson, J.C. Andersson-Assarsson, L. Sjöström, P. Froguel, A. Walley, et al., Adipose tissue resting energy expenditure and expression of genes involved in mitochondrial function are higher in women than in men, J. Clin. Endocrinol. Metab. 98 (2013) E370–E378.
[38] V. Ogilvie, M. Passmore, L. Hyndman, L. Jones, B. Stevenson, A. Wilson, H. Davidson, R.R. Kitchen, R.D. Gray, P. Shah, et al., Differential global gene expression in cystic fibrosis nasal and bronchial epithelium, Genomics 98 (2011) 327–336.
[39] S. Papiris, A. Kotanidou, K. Malagari, C. Roussos, Clinical review: severe asthma, Crit. Care 6 (2002) 30–44.
[40] A.A. Pezzulo, T.D. Starner, T.E. Scheetz, G.L. Traver, A.E. Tilley, B.G. Harvey, R.G. Crystal, P.B. McCray, J. Zabner, 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 Cell. Mol. Physiol. 300 (2011) L25–L31.
[41] A. Rab, S.M. Rowe, S.V. Raju, Z. Bebok, S. Matalon, J.F. Collawn, Cigarette smoke and CFTR: implications in the pathogenesis of COPD, Am. J. Physiol. Lung Cell. Mol. Physiol. 305 (2013) L530–L541.
[42] S.I. Rennard, COPD: overview of definitions, epidemiology, and factors influencing its development, Chest 113 (1998) 235S–241S.
[43] J.R. Riordan, J.M. Rommens, B. Kerem, N. Alon, R. Rozmahel, Z. Grzelczak, J. Zielenski, S. Lok, N. Plavsic, J.L. Chou, Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA, Science 245 (1989) 1066–1073.
[44] P. Rios, X. Li, M. Köhn, Molecular mechanisms of the PRL phosphatases, FEBS J. 280 (2013) 505–524.

[45] P. Roepman, P. Kemmeren, L.F. Wessels, P.J. Slootweg, F.C. Holstege, Multiple robust signatures for detecting lymph node metastasis in head and neck cancer, Cancer Res. 66 (2006) 2361–2366.
[46] A.J. Ross, L.A. Dailey, L.E. Brighton, R.B. Devlin, Transcriptional profiling of mucociliary differentiation in human airway epithelial cells, Am. J. Respir. Cell Mol. Biol. 37 (2007) 169–185.
[47] L. Shan, T. Kawakami, S. Asano, S. Noritake, D. Yoshimoto, K. Yamashita, H. Kikkawa, M. Kinoshita, S. Matsubara, Inverse relationship between Sec14l3 mRNA/protein expression and allergic airway inflammation, Eur. J. Pharmacol. 616 (2009) 293–300.
[48] M. Shipitsin, C. Small, E. Giladi, S. Siddiqui, S. Choudhury, S. Hussain, Y.E. Huang, H. Chang, D.L. Rimm, D.M. Berman, et al., Automated quantitative multiplex immunofluorescence in situ imaging identifies phospho-S6 and phospho-PRAS40 as predictive protein biomarkers for prostate cancer lethality, Proteome Sci. 12 (2014) 40.
[49] W.K. So, Q. Fan, M.T. Lau, X. Qiu, J.C. Cheng, P.C. Leung, Amphiregulin induces human ovarian cancer cell invasion by down-regulating E-cadherin expression, FEBS Lett. 588 (2014) 3998–4007.
[50] A. Spira, J. Beane, V. Shah, G. Liu, F. Schembri, X. Yang, J. Palma, J.S. Brody, Effects of cigarette smoke on the human airway epithelial cell transcriptome, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 10143–10148.
[51] A. Su, S. Ra, X. Li, J. Zhou, S. Binder, Differentiating cutaneous squamous cell carcinoma and pseudoepitheliomatous hyperplasia by multiplex qRT-PCR, Mod. Pathol. 26 (2013) 1433–1437.
[52] B.L. Tang, H.Y. Gee, M.G. Lee, The cystic fibrosis transmembrane conductance regulator's expanding SNARE interactome, Traffic 12 (2011) 364–371.
[53] A.M. Trzcinska-Daneluti, D. Ly, L. Huynh, C. Jiang, C. Fladd, D. Rotin, High-content functional screen to identify proteins that correct F508del-CFTR function, Mol. Cell. Proteomics 8 (2009) 780–790.
[54] A.G. Valdivieso, M. Clauzure, M.C. Marín, G.L. Taminelli, M.M. Massip Copiz, F. Sánchez, G. Schulman, M.L. Teiber, T.A. Santa-Coloma, The mitochondrial complex I activity is reduced in cells with impaired cystic fibrosis transmembrane conductance regulator (CFTR) function, PLoS One 7 (2012) e48059.
[55] D. Venet, J.E. Dumont, V. Detours, Most random gene expression signatures are significantly associated with breast cancer outcome, PLoS Comput. Biol. 7 (2011) e1002240.
[56] C. Verhaeghe, C. Remouchamps, B. Hennuy, A. Vanderplasschen, A. Chariot, S.P. Tabruyn, C. Oury, V. Bours, Role of IKK and ERK pathways in intrinsic inflammation of cystic fibrosis airways, Biochem. Pharmacol. 73 (2007) 1982–1994.
[57] S.W. Wallace, A. Magalhaes, A. Hall, The Rho target PRK2 regulates apical junction formation in human bronchial epithelial cells, Mol. Cell. Biol. 31 (2011) 81–91.
[58] X. Wang, J. Venable, P. LaPointe, D.M. Hutt, A.V. Koulov, J. Coppinger, C. Gurkan, W. Kellner, J. Matteson, H. Plutner, et al., Hsp90 cochaperone Aha1 downregulation rescues misfolding of CFTR in cystic fibrosis, Cell 127 (2006) 803–815.
[59] T. Wassmer, N. Attar, M. Harterink, J.R. van Weering, C.J. Traer, J. Oakley, B. Goud, D.J. Stephens, P. Verkade, H.C. Korswagen, et al., The retromer coat complex coordinates endosomal sorting and dynein-mediated transport, with carrier recognition by the trans-Golgi network, Dev. Cell 17 (2009) 110–122.
[60] C. Xie, X.H. Jiang, J.T. Zhang, T.T. Sun, J.D. Dong, A.J. Sanders, R.Y. Diao, Y. Wang, K.L. Fok, L.L. Tsang, et al., CFTR suppresses tumor progression through miR-193b targeting urokinase plasminogen activator (uPA) in prostate cancer, Oncogene 32 (2013) 2282–2291 (2291.e2281-2287).
[61] L. Xu, Q. Li, D. Xu, Q. Wang, Y. An, Q. Du, J. Zhang, Y. Zhu, Y. Miao, hsa-miR-141 downregulates TM4SF1 to inhibit pancreatic cancer cell invasion and migration, Int. J. Oncol. 44 (2014) 459–466.
[62] M.A. Zariwala, H.Y. Gee, M. Kurkowiak, D.A. Al-Mutairi, M.W. Leigh, T.W. Hurd, R. Hjeij, S.D. Dell, M. Chaki, G.W. Dougherty, et al., ZMYND10 is mutated in primary ciliary dyskinesia and interacts with LRRC6, Am. J. Hum. Genet. 93 (2013) 336–345.