

From CFTR to a CF signalling network: a systems biology approach to study Cystic Fibrosis

Matthieu Najm^{1,2,3*}, Loredana Martignetti^{1,2,3},
Matthieu Cornet^{1,2,3,4}, Mairead Kelly-Aubert^{4,5},
Isabelle Sermet^{4,5,6}, Laurence Calzone^{1,2,3*†}, Véronique Stoven^{1,2,3*†}

¹Center for Computational Biology (CBIO), Mines Paris-PSL, Paris, 75006, France.

²Institut Curie, Université PSL, Paris, 75005, France.

³INSERM U900, Paris, France.

⁴Institut Necker Enfants Malades, INSERM U1151, Paris, 75005, France.

⁵Université Paris Cité, Paris, France.

⁶Centre de Référence Maladies Rares, Mucoviscidose et Maladies Apparentées, Hôpital Necker Enfants Malades AP-HP Centre Paris Cité, Paris, France.

*Corresponding author(s). E-mail(s): matthieu.najm@minesparis.psl.eu; Laurence.Calzone@curie.fr; veronique.stoven@mines-paristech.fr;

†These authors contributed equally to this work.

Abstract

Background: Cystic Fibrosis (CF) is a monogenic disease caused by mutations in the gene coding the Cystic Fibrosis Transmembrane Regulator (CFTR) protein, but its overall physio-pathology cannot be solely explained by the loss of the CFTR chloride channel function. Indeed, CFTR belongs to a yet not fully deciphered network of proteins participating in various signalling pathways.

Methods: We propose a systems biology approach to study how the absence of the CFTR protein at the membrane leads to perturbation of these pathways, resulting in a panel of deleterious CF cellular phenotypes.

Results: Based on publicly available transcriptomic datasets, we built and analyzed a CF network that recapitulates signalling dysregulations. The CF network topology and its resulting phenotype was found to be consistent with CF pathology.

Conclusion: Analysis of the network topology highlighted a few proteins that may initiate the propagation of dysregulations, those that trigger CF cellular phenotypes, and suggested several candidate therapeutic targets. Although our research is focused on CF, the global approach proposed in the present paper could also be followed to study other rare monogenic diseases.

Keywords: Cystic Fibrosis, CF cellular phenotypes, CF signalling network, network topology, therapeutic target

1 Background

Cystic fibrosis (CF) is the most common life-limiting autosomal disease in the Caucasian population, affecting about 162.000 patients worldwide, of which 105.000 are diagnosed (Guo et al., 2022). It is caused by mutations in the *CFTR* gene encoding for the cystic fibrosis transmembrane conductance regulator (CFTR) protein, a chloride ion channel expressed at the apical membrane of polarized epithelial cells (Seibert et al., 1997). More than 2000 mutations in *CFTR* have been reported, but the deletion of the F508 amino-acid (F508del) is present in 70% of the mutated alleles in the Caucasian population, and most of the mutations lead to compromised transepithelial anion conductance (Veit et al., 2016). Various organs are affected in CF, but the most severe symptoms are in the lungs, where the defective chloride transport leads to the dehydration of surface mucus, chronic bacterial infection, and inflammation, causing lung tissue damage and ultimately, respiratory insufficiency.

However, CF symptoms not only result from the loss of CFTR-mediated anion conductance, but also from perturbations of other CFTR-dependent biological functions (Hanssens et al., 2021). Indeed, CFTR belongs to a protein-protein interactions (PPI) network (Pereira et al., 2021; Farinha and Gentzsch, 2021), and the absence of CFTR may perturb its direct or indirect interactors, and propagate dysregulations towards various biological pathways in which these interactors play a role. In agreement with this idea, studies on *CFTR* -/- knockout mice (Crites et al., 2015), *CFTR*

-/- knockout piglets (Fleurot et al., 2022), and cell lines in which CFTR is inactivated by the CRISPR/Cas9 technology (Hao et al., 2020) have reported that the absence of CFTR affects cell signalling and transcriptional regulation. These dysregulations may explain various and apparently unrelated cellular phenotypes, including uncontrolled pro-inflammatory response (Jacquot et al., 2008), unbalanced oxidative stress with increased reactive oxygen species (Jeanson et al., 2012), impaired epithelial regeneration (Conese and Di Gioia, 2021), or perturbation of cell junctions and cytoskeleton (Pankonien et al., 2022).

To explain these seamlessly unrelated phenotypes, we propose to use a systems biology approach for CF, where the two aims are (1) to explore how the absence of CFTR can be functionally related to the signalling dysregulations that ultimately lead to CF cellular phenotypes; and (2) to suggest new therapeutic targets that may modulate these phenotypes.

Indeed, systems biology approaches provide tools for building network models to reason on complex systems. Subsequent topological analysis or dynamic mathematical models performed on these networks allow to study how different biological components of the networks interact to produce phenotypic properties, which is relevant to the questions at hand.

Systems biology approaches have seldom been implemented in monogenic diseases, but have been widely used in cancer, often referred to as a network disease (Hornberg et al., 2006), where intricate processes contribute to the emergence of unexpected and often non-intuitive phenotypes. Very few contributions have been devoted to systems biology approaches of CF. Previous studies have focused on the construction of the CFTR interactome that distinguishes PPI networks involving wt-CFTR and those involving the most frequent mutant F508del-CFTR (Pankow et al., 2015; Pereira et al., 2021). The latter led to the construction of a navigable knowledge map, the CyFi-MAP, that integrates all proteins known to be involved in the processing, maturation,

retention and degradation of wt-CFTR and F508del-CFTR. Although the CyFi-MAP represents a key contribution for the problem of rescuing F508del-CFTR, this map does not tackle the questions of interest in the present paper. Other studies highlighted links between CFTR and signalling pathways involved in the disease (see (Pankonien et al., 2022) for a review), but they did not provide a global view of how CFTR is linked to dysregulated molecular mechanisms and to CF phenotypes. Recently, transcriptomic data have been produced to identify differentially expressed genes in CF. These genes were connected within a PPI network, based on information available in PPI databases (Trivedi et al., 2023). Although this network comprises genes that are consistent with current knowledge in CF, it does not contain CFTR, which prevents understanding the functional link between CFTR and the differentially expressed genes, or with CF cellular phenotypes.

To overcome the limitation of previous studies, in the systems biology approach proposed here, we build a comprehensive signalling network, called the CF network in the following, that recapitulates CF pathway dysregulations, using transcriptomic data available for CF and control patients and information available in biological pathway databases. As detailed below, we connected CFTR to this network based on PPI information. Analysis of the CF network topology allows to formulate hypotheses on key proteins and molecular mechanisms that functionally link CFTR to major CF cellular phenotypes, and to highlight potential targets that may counteract these phenotypes.

2 Results

2.1 Global approach to building the CF network

In systems biology, various networks can be built to represent different types of biological information, such as gene regulatory networks, genetic interaction networks, signal transduction networks, metabolic networks, PPI networks, or disease networks.

111 There is no universal technique that can be followed to construct networks, and the
 112 choice of their representation needs to be adapted to the question of interest. In the
 113 present study, we wish to establish a link between the absence of CFTR and the over-
 114 all signalling dysregulations leading to the cellular phenotypes that characterize CF.
 115 Therefore, we chose to build a CF network focusing on the signalling pathways that
 116 are perturbed in the disease, and where dysregulations in one pathway may affect
 117 other pathways. In order to avoid potential bias in the CF literature, we adopted a
 118 data-driven approach based on publicly available transcriptomic studies. We are aware
 119 that some CF phenotypes might arise from biological events that are not detectable
 120 in the transcriptome of CF cells, but we considered that gene expression data had
 121 the potential to capture some of the major molecular dysregulations present in CF
 122 cells. Our study relies on a meta-analysis of public transcriptomic datasets for CF
 123 respiratory epithelial cells and their Non-Cystic Fibrosis (NCF) control counterparts,
 124 allowing the identification of the signalling pathways dysregulated in CF. Based on
 125 information available in pathway databases, these dysregulated pathways share many
 126 common proteins, which allowed to connect them into a network. As detailed below,
 127 CFTR was absent from this network, because it did not belong to any of the differ-
 128 entially expressed signalling pathways. However, we observed that several proteins of
 129 the network were also present in the CFTR PPI interactome, either as direct interac-
 130 tors of CFTR, or as indirect interactors of CFTR via a single intermediate protein.
 131 This important result was consistent with the assumption that CFTR direct interac-
 132 tors may be perturbed in CF, and initiate the propagation of dysregulations within
 133 the CF network.

134 The figure 1 summarizes the global approach followed in the present study.

135 2.2 Selection of publicly available transcriptomics data

136 Many transcriptomic studies have been performed in CF over the last 15 years ([Ideozu](#)
137 [et al., 2019](#)). However, these data suffer from a few limitations that are obstacles to
138 improve our understanding of CF. First, they consider a wide range of cell types,
139 including native nasal or bronchial cells, primary cultures of these cells, whole blood,
140 peripheral mononuclear cells, leukocytes, or immortalized cell lines. Therefore, com-
141 parison between studies to identify common key molecular determinants can lead to
142 inconsistent results. Then, compared to studies on more common diseases such as
143 cancer, most of CF transcriptomic studies have very few samples per condition (dis-
144 ease and control), decreasing the statistical power of these datasets when analyzed
145 alone. Finally, these studies rely on various experimental biological models and tran-
146 scriptomic technologies which rarely lead to consistent results between studies ([Clarke](#)
147 [et al., 2013](#)), particularly when the analyses are performed at the gene level.

148 To try and overcome these limitations, we focused on studies considering only sam-
149 ples from human Airway Epithelial Cells (hAEC hereafter), i.e., bronchial, tracheal, or
150 nasal cells. Indeed, functional modifications in these cells are expected to reflect some
151 of the most severe symptoms in the lung. We included studies of cell lines or primary
152 cultures, in order to gather a statistically significant number of samples, because as
153 shown in Table 1, each dataset comprises a very limited number of samples. We also
154 focused on studies on the F508del mutation, for which most data are available. We
155 discarded two studies (([Virella-Lowell et al., 2004](#)) and ([Rehman et al., 2021](#))) that
156 provide transcriptomic data for other mutations, because the corresponding cells could
157 display disparities with respect to F508del cells. We retrieved from the literature all
158 the CF transcriptomic studies with publicly available data that matched these criteria
159 (see Methods section), which led to 10 CF transcriptomic datasets shown in Table 1.

The studies are still heterogeneous in terms of tissue sample (bronchial, tracheal or nasal), cell culture type (cell-line or primary culture) and transcriptomic technology (micro-array or RNA-Seq). However, we kept the 10 studies in order to improve statistical significance, because the numbers of samples per condition are very small in all studies: the median number of samples was 5 for disease (CF) and control (NCF) conditions.

2.3 Meta-analysis of transcriptomic studies at the level of biological pathways

The most straightforward way to analyse transcriptomic data is to identify Differentially Expressed Genes (DEGs), and to search for biological pathways enriched in these DEGs. This approach failed in the present meta-analysis, because the number of DEGs common to at least 3 out of 7 studies was too small to be enriched in any pathway, even though many reference pathway databases were considered (the Hallmark gene sets from the the MSigDB Database (Liberzon et al., 2015), the Pathway Interaction Database (PID) (Schaefer et al., 2009), the KEGG database (Kanehisa et al., 2021)). In fact, it has become clear that, in complex diseases, identification of pathway dysregulations based on DEGs is not optimal and does not provide robust results (Wang et al., 2010).

Therefore, the meta-analysis was conducted at the pathway level. Many methods have been proposed to capture pathway dysregulations when they do not appear clearly based on enrichment from lists of DEGs (Martignetti et al., 2016; Landais and Vallot, 2023; Schubert et al., 2018; Vaske et al., 2010). In the present study, we used the Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) approach. GSEA was performed separately on each dataset identified as over-activated or under-activated signalling pathways in hAEC CF cells, based on the complete expression matrix of CF and NCF samples, and taking into account the expression level of all

186 genes belonging to the same pathway. We used pathway definitions provided by the
187 KEGG pathway database (Kanehisa et al., 2021), because this database provides
188 graphical pathway representations that also include phenotypes, which helped the
189 analysis of the CF network, as detailed in Section 2.5. We tested 131 KEGG biological
190 pathways, and Differentially Expressed Pathways (DEPs, hereafter) were identified
191 according to a adjusted p-value lower or equal to 0.25, as detailed in Section 3.4. The
192 number of up- and down-regulated pathways for each dataset is provided in Table 2

193 The analysis of DEPs showed that 15 of the 134 biological pathways tested were
194 differentially expressed in at least 3 studies. However, a closer analysis highlighted dis-
195 crepancies between studies. As shown in the heatmap presenting the GSEA Normalised
196 Enrichment Score (NES) (Figure 2), for these 15 common DEPs, the 10 datasets can
197 be gathered into 2 subgroups: subgroup 1 comprising 7 datasets in which common
198 DEPs tend to be up-regulated, while they tend to be down-regulated in subgroup 2
199 comprising the 3 other datasets. This appeared somewhat puzzling. Our hypothesis
200 is that datasets belonging to subgroups 1 or 2 arise from studies in which the differ-
201 entiation media used for the primary cultures did not favor the same cell type, and
202 therefore, should not be analyzed together.

203 This was confirmed by the Saint-Criq (UNC) and Saint-Criq (SC) datasets (see
204 Table 1), belonging respectively to subgroups 1 and 2, where it was shown that the
205 UNC and SC differentiation media (two common differentiation media used on CF and
206 non-CF epithelia) significantly impact cell lineage in primary cultures of CF hAEC,
207 and consequently, the resulting transcriptomic profiles (Saint-Criq et al., 2020). In
208 this study, it was shown that the UNC medium promoted differentiation into club
209 and goblet cells, while the SC medium favored the growth of ionocytes and ciliated
210 cells. Consistent with this result, the Ling transcriptomic dataset, which belongs to
211 subgroup 2, was also obtained from primary cultures of CF and NCF airway epithelia
212 that were differentiated into ciliated pseudo-stratified airway cells (Ling et al., 2020).

213 Datasets from subgroup 2 appeared in contradiction with the main CF phenotypes.
 214 In particular, the *TNF- α signalling pathway* or *NF- κ B signalling pathway* are down-
 215 regulated in this subgroup, although the over-activation of these pathways is a well-
 216 known feature of CF disease. Therefore, we only considered the 7 datasets belonging
 217 to subgroup 1 for further analysis.

218 In this subgroup, the transcriptomic analysis appears to be highly consistent,
 219 since among the 15 DEPs common to at least 3 studies, 5 are up-regulated in CF
 220 vs NCF samples in 4 studies (*NOD-like receptor signalling pathway*, *Cytosolic DNA-*
 221 *sensing pathway*, *Cytokine-cytokine receptor signalling pathway*, and *Regulation of*
 222 *actin cytoskeleton*), 2 are up-regulated in CF vs NCF samples in 5 studies (*Osteo-*
 223 *clast differentiation* and *Toll-like receptor signalling pathway*), and the *IL-17 signalling*
 224 *pathway* is up-regulated in CF vs NCF in 6 studies.

225 Overall, the 15 DEPs common to at least 3 studies are in agreement with various
 226 known aspects of CF disease, which confirms that our analysis did capture relevant
 227 information about CF. In particular, besides the *TNF- α* and *NF- κ B* signalling path-
 228 ways well known to be up-regulated in CF, the *IL-17* pathway contributes to CF
 229 lung disease (Hsu et al., 2016), the differentiation of osteoclast is perturbed in CF
 230 (Dumortier et al., 2021), the Toll-like receptor signalling pathway modulates function,
 231 inflammation and infection of lung in CF (Kosamo et al., 2020; Curutiu et al., 2018;
 232 Fleurot et al., 2022), and CFTR plays a role in cell junction and actin cytoskeleton
 233 organization (Pankonien et al., 2022).

234 2.4 Building the CF network

235 The 15 individual DEPs of the KEGG database provide interesting information about
 236 what is dysregulated in CF, but a lot of these pathways are partially redundant and
 237 show a high overlap of genes and interactions, indicating that they are highly inter-
 238 twined. A dysregulation in one of these pathways will have a consequence in another

239 pathway. To study the connection between them, we propose to merge them into a
240 single network called the CF network.

241 The DEPs were extracted with the OmniPathR package (Türei et al., 2016) and
242 curated, as described in Section 3.5. The rules that were used to build and clean
243 this network are detailed in Section 3.6. The network, comprising 330 nodes and 529
244 interactions, is not fully connected: it contains one main component including 317
245 nodes connected by 515 interactions, and two small additional components that are
246 non connected to the main component, and called unconnected components hereafter
247 (See Figure 3). The overall network can be accessed as a Cytoscape session, in the
248 [sysbio-curie/CFnetwork_cytoscape](https://github.com/sysbio-curie/CFnetwork_cytoscape) github repository for further analysis.

249 *Identification of CFTR interactors in the CF network*

250 It is striking to note that CFTR does not belong to any of the 15 DEPs, and there-
251 fore, is not part of the network. In fact, CFTR is present in only 7 biological pathways
252 of the KEGG database (*ABC Transporters*, *cAMP signalling pathway*, *AMPK sig-
253 nalling pathway*, *tight junction*, *Gastric and acid secretion*, *pancreatic secretion* and
254 *bile secretion*), but these pathways did not belong to the DEPs.

255 Therefore, we searched for the presence of CF network proteins in the network of
256 proteins reported to be involved in protein-protein interactions (PPI) with wt-CFTR
257 or F508del-CFTR (Pereira et al., 2021). Indeed, according to the CyFi-MAP, 4 direct
258 interactors of wt-CFTR but not of F508del-CFTR (CSNK2A1, PRKACA, SYK and
259 TRADD) belong to the CF network. Furthermore, 4 additional proteins (EZR, SRC,
260 PLCB1/3) present in the network interact with wt-CFTR (but not with F508del-
261 CFTR) through a single intermediate protein. Figure 4 shows these 8 proteins, their
262 intermediates and their interactions with CFTR. The presence in the CF network of 8
263 first or second neighbours in the CFTR interactome is an interesting result in favour
264 of our assumption that CFTR interactors may propagate functional dysregulations
265 into the network.

2.5 Analysis of the CF network

Extensive interpretation of this large network, which contains rich but complex information, is beyond the scope of the present paper. However, we will investigate how analysis of its topology can help tackle the two questions of interest: how the absence of the CFTR protein at the membrane leads to CF cellular phenotypes, and how therapeutic targets can be suggested from this network.

2.5.1 Topological description of the CF network

The final CF network comprises 330 proteins and 529 interactions. Interestingly, CFTR interactors are present only in the main component, because according to the CyFi-MAP, it would not have been possible to link CFTR to proteins of the two small unconnected components without adding a large number of intermediate nodes. One of the two unconnected components contains 10 proteins and 9 interactions, and corresponds to cascades of the JAK/STAT signalling pathway. The other contains 3 proteins and 3 interactions, and corresponds to a cascade of the Transforming Growth Factor Beta (TGF β) signalling pathway. In the present section, we will focus on the main component of the CF network, and the two unconnected components will be discussed in Section 2.5.3.

The topological description of the main component will be organized around three types of remarkable nodes: (1) the source nodes, i.e., CFTR first or second neighbours that were used to connect CFTR to the network, as described in Section 2.4; (2) the sink nodes, i.e., the nodes from which no edge leaves in the network, and whose activation finally triggers their associated phenotypes (for example, transcription factors are typical sink nodes); (3) the hubs, i.e. the nodes with high betweenness centrality, through which the flow of information that passes is high. Figure 5 illustrates where these remarkable nodes stand within the network's topology.

Source nodes and initiation of dysregulations

According to the CyFi-MAP, 8 first or second neighbours of wt-CFTR interactors whose interactions are lost with F508del-CFTR are present in the CF network: **CSNK2A1**, **EZR**, **PLCB1**, **PLCB3**, **PRKACA**, **SRC**, **SYK** and **TRADD**. In the absence of CFTR, these 8 proteins can be viewed as source nodes that may initiate dysregulations that subsequently propagate within the network and finally reach the sink nodes (see Figure 4).

Perturbations of some of these source nodes in CF cells, or their role in CF cellular phenotypes, are sustained by various studies:

- **CSNK2A1**, also known as CK2 (casein kinase 2), is strongly overactivated in CF vs wild-type cells (Venerando et al., 2011).
- Cellular levels of **TRADD** are controlled by its lysosomal degradation in a wt-CFTR-dependant manner, and this regulation is lost with F508del-CFTR and G551D-CFTR (Wang et al., 2016).
- **SRC** was shown to be overexpressed and overactivated in CF cells (Massip Copiz and Santa Coloma, 2016).
- **PLCB3** is a known CF modifier gene, for which the loss of function S845L variant is associated with a mild progression of the pulmonary disease and a reduction of *Pseudomonas aeruginosa*-induced IL8 release. This indicates that **PLCB3** plays a role in the inflammation phenotype in CF (Rimessi et al., 2018).
- The active form of ezrin (**EZR**) is mainly located in the apical region of wild type airway epithelial cells, while in their CF counterparts, it is diffusely expressed in its inactive state in the cytosol (Favia et al., 2010; Wu and Eickelberg, 2019).
- The **SYK** and **PRKACA** kinases play key roles with respect to CFTR, since the former negatively regulates the amount of CFTR at the membrane through phosphorylation at Y512 (Mendes et al., 2011), while the latter is a well-known

regulator of the CFTR chloride channel conductance (Egan et al., 1992), but their implication as propagators of dysregulations has not been investigated yet.

Sink nodes and CF phenotypes

There are 35 sink nodes in the main component of the CF network that are reached from each of the 8 source nodes. The full list of sink nodes and their associated phenotypes are given in the Supplementary file. Among them, we can cite:

- **NFKB1**, **NFKB2**, **RELA** and **RELB** are part of the NF- κ B complex, a transcription factor that can be activated by various stimuli such as cytokines, oxidant radicals, bacterial or viral products. It controls the expression of pro-inflammatory genes, and is related to various phenotypes including inflammation and cell survival/proliferation.
- **FOS** and **JUN** are two sub-units of the AP-1 transcription factor activated by the MAPK signalling pathways, and are associated with inflammation and proliferation phenotypes.
- **CASP3** and **CASP7** caspases are the effectors of apoptosis.
- **CASP1** is a caspase known to be the effector of pyroptosis, a highly pro-inflammatory cell death mechanism.
- 10 sink nodes belong to the regulation of actin cytoskeleton pathway, including **ACTN4**, **ARPC5**, **PFN**, **MYL12B** and **VCL**. These nodes are associated to various phenotypes related to cytoskeleton, including focal adhesion, adherens junction, and actin polymerisation.
- **IRF1**, **IRF3**, **IRF5**, and **IRF7**, that are members of the IRF family of transcription factors involved in the innate immune response phenotype, and controlling expression of Type-1 interferons upon viral infection.

Importantly, the phenotypes associated to these sink nodes have already been described in the CF context. In particular:

1. The NF κ B and AP-1 transcription factors are complexes of sink nodes that mediate inflammation, the most studied phenotype of CF disease. In addition to the well-known activation of NF κ B in CF, AP-1 is one of the downstream transcription factors of the MAPK pathway that was shown to be activated in CF (Bérubé et al., 2010; Wellmerling et al., 2022), as shown in Figure 6.
2. Controversial results were reported about apoptosis in CF epithelial cells. Some studies showed defective susceptibility of CF cells to pro-apoptotic stimuli (Cannon et al., 2003; Gottlieb and Dosanjh, 1996), while others observed increased apoptosis (Chen et al., 2018; Voisin et al., 2014; Yalçın et al., 2009; Rottner et al., 2007). All agree that apoptosis is dysregulated in CF.
3. The dysregulation of actin cytoskeleton in CF is well documented, with a disorganized actin cytoskeleton, absence of actin stress fibres (Favia et al., 2010; Lasalvia et al., 2016; Burat et al., 2022), and disrupted tight junctions (De Lisle, 2014; Castellani et al., 2012).
4. Finally, most of these phenotypes are related to the innate immune response and various works indicate a dysfunction in the innate immune response of CF patients (Kosamo et al., 2020; Gillan et al., 2023; Dugger et al., 2020).

Betweenness centrality and flow of information

In a network, the betweenness centrality (BC) of a node is the number of shortest paths that pass through that node. This measure is a way of detecting the amount of influence a node has over the flow of information in a network. Nodes with high BC, referred to as hubs, may provide interesting therapeutic targets, because their inhibition may efficiently reduce the propagation of information within the network (Durón et al., 2019). Therefore, we calculated the BC for all nodes of the CF network, as detailed in the Methods section. All nodes were then ranked according to this measure, and Figure 7A displays the histogram of the BC score. Interestingly, most proteins have a BC score below 3000, and only a very limited number of proteins have a BC score above

370 6000 (ARHGEF12, IKBKE, LSP1, PIK3KC1, PYCARD, RAC1, TRAF2, TRAF3,
371 TRAF6). The list of the top 30 proteins is provided in the Supplementary File. Among
372 them, PI3KCA could be an interesting therapeutic target candidate and is discussed
373 in the next section.

374 2.5.2 Biological insights from the topological analysis

375 A simple path analysis of the network shows that several source nodes may contribute
376 collectively to the emergence of the CF phenotypes, which illustrates the complexity
377 of the disease. Indeed, while the source nodes PRKACA, EZR and CSNK2A1 are
378 upstream of a limited number of sink nodes, PLCB1/3, SRC, SYK and TRADD are
379 upstream of the 35 sink nodes, i.e., there exists a path from each of these 6 source
380 nodes to each of the 35 sink nodes (Figure 7B).

381 For example, TRADD is known to be up-regulated in CF ([Ferenc Karpati, 2000](#))
382 and to participate in the uncontrolled inflammation (See Figure 6). Interaction
383 between wt-CFTR and TRADD enhances the degradation of TRADD, which con-
384 trols the activity of this pathway, as demonstrated by Wang and colleagues ([Wang](#)
385 [et al., 2016](#)). This direct interaction is lost with F508del-CFTR, which may contribute
386 to the dysregulation of TNF- α and NF- κ B signalling pathways in CF. However, up-
387 regulation of TRADD could also contribute to the inflammation phenotype through
388 another route, by inducing over-activation of the MAPK pathway, and in particular
389 of AP-1, one of its output transcription factors. In addition, as shown in Figure 8, our
390 network suggests that other source nodes than TRADD could also initiate dysregu-
391 lation of the inflammation phenotype because they are also connected to the NF- κ B
392 sink node. Among these sources, we can cite: (1) SYK, which would be consistent
393 with its role in inflammation processes shown in other diseases ([Riccaboni et al., 2010](#);
394 [Wong et al., 2004](#)); (2) PLCB1/3, which are consistent with previous studies reporting
395 PLCB3 as a key modulator of IL8 expression in CF bronchial epithelial cells ([Bezzetti](#)

et al., 2011); (3) CSNK2A1, whose hyperactivity could contribute to activation of NF- κ B by enhancing the phosphorylation and degradation of IKBKA.

Overall, the number of source nodes and routes that may contribute to inflammation in CF illustrates the challenge posed by its modulation, in order to reduce the related clinical symptoms. Various anti-inflammatory drugs have been recently evaluated in clinical trials (Bell et al., 2020), but none of them target the source nodes of the present study. Our hypothesis is that these source nodes could be interesting candidate targets in CF. In particular, SYK has emerged as a potential target for the treatment of numerous diseases. Many inhibitors are known for this kinase, which would allow to evaluate their potential anti-inflammatory effect in CF cells. These inhibitors include one marketed drug (Fostamatinib), but other inhibitors are currently under investigation in clinical trials for a range of indications (Cooper et al., 2023). Interestingly, since SYK is connected to the 35 sink nodes, its inhibition may also contribute to the modulation of other CF phenotypes than inflammation. In particular, it could modulate CF phenotypes associated to the 35 sink nodes and mentioned in Section 2.5.1, such as dysregulations in apoptosis, cytoskeleton or innate immune response. Similarly, our network suggests that PLCB3 could be an interesting target for inflammation in CF. This is consistent with the fact that PLCB3 silencing in CF bronchial epithelial cells exposed to *Pseudomonas aeruginosa*, reduces the expression of IL-8 chemokine (Bezzetti et al., 2011). The U73122 PLC inhibitor could be an interesting pharmacological tool to further evaluate this strategy. As in the case of SYK, PLCB3 is connected to the 35 sink nodes, which means that its inhibition may also improve other CF cellular phenotypes. Consistent with this idea, it was shown that treatment with a SRC inhibitor, another of the 6 source nodes upstream of the 35 sink nodes, decreased the inflammatory changes and improved cytoskeletal defects in F508del human cholangiocytes (Fiorotto et al., 2018).

Besides source nodes, candidate therapeutic targets can be searched among hubs in the network, i.e. among the best ranked proteins according to the BC score. Besides this score, additional arguments can be invoked to highlight the best candidates. In particular, the fact that a protein is known in the literature to play a role in the disease, and that pharmacological modulators (or even better, marketed drugs) are available to allow experimental validation, are important criteria. In line with these ideas, PI3KCA appears as an interesting candidate target. Indeed, several inhibitors are known for this kinase, including the marketed drug Alpelisib, which would allow experimental tests in CF models. It has been suggested as a candidate target in CF based on its role in many signalling pathways implicated in CF lung pathogenesis (Natarajan, 2020). The fact that PI3KCA belongs to best ranked proteins with respect to the BC score (See Figure 7A) offers a quantitative argument in favor of this idea. In addition, PI3KCA is connected through the network to the 35 sink nodes, which means that its inhibition may modulate inflammation, but also other CF cellular phenotypes related to the sink nodes.

2.5.3 Analysis of the unconnected components in the CF network

As mentioned in Section 2.5.1, Figure 3 shows that the CF network comprises two small unconnected components that are part of the TGF β and JAK/STAT signalling pathways. Contrary to source nodes of the main component, dysregulation of the source nodes of these unconnected components (namely the 4 interleukins IL2, IL21, IL4 and IL6 for one component, and TGF β for the other) cannot be explained by the absence of CFTR in a direct manner, because they are not linked to CFTR within a single network. However, activation of a sink node of the main component may modulate the expression of a source node in an unconnected component, affecting the activity of this unconnected component. For example, activation of the AP-1 transcription factor (a sink node of the main component) due to activation of the MAPK pathway in the main component, regulates the expression of TGF β . This example shows how

dysregulations in one pathway may have consequences in other pathways of the CF network, even if they are not connected, again illustrating to the complexity of the disease. We propose that phenotypes arising from the two unconnected components could be defined as secondary phenotypes, as opposite to primary phenotypes arising from dysregulations of the main component (discussed in Section 2.5.1).

The JAK-STAT component mediates various cellular processes, including cell growth and apoptosis, but the role of these cascades has not been widely studied in CF. The TGF β component leads to the activation of SMAD2, a transcriptional modulator that regulates multiple cellular phenotypes, including cell proliferation, apoptosis, and differentiation. High levels of TGF β have been associated with the severity of lung disease (Dorfman et al., 2008; Sagwal et al., 2020), and this protein was proposed as a therapeutic target for CF (Kramer and Clancy, 2018). Our study suggests that therapeutic targets should be chosen among proteins closer to CFTR in the network, in particular among the source nodes of the main component (as discussed above), because they may more successfully limit the global propagation of molecular dysregulations within the overall network.

3 Methods

3.1 Datasets selection

Based on the search engines of the National Center for Biotechnology Information (NCBI) and the European Nucleotide Archive (ENA), we selected 10 datasets from 8 studies published between 2007 and 2021. The selection criteria to include CF transcriptomic datasets were the following: (1) they should correspond to human Airway Epithelial Cells (hAEC); (2) the cells should be homozygous for the most common mutation F508del; (3) the transcriptomic data should be publicly available. Therefore, studies including samples heterozygous for the F508del mutation ((Virella-Lowell et al., 2004) and (Rehman et al., 2021)), studies with no data available (Zabner et al.,

2005) and (Wright et al., 2006)) were not included. In addition, studies with less than two samples were excluded ((Bampi et al., 2020) and (Veltman et al., 2021)), as the subsequent statistical analyses require several samples per condition. The list of selected transcriptomic studies is provided in Table 1.

3.2 Biological pathways databases

We initially considered a total of 380 gene sets corresponding to 380 biological pathways: 50 Hallmark gene sets from the the Molecular Signatures Database (MSigDB) (Liberzon et al., 2015), 196 from the Pathway Interaction Database (PID) (Schaefer et al., 2009) and 134 from the KEGG database, restricted to the *Genetic Information Processing*, *Environmental Information Processing*, *Cellular Processes* and *Organismal systems* subdivision. However, most of the analyses were performed using only KEGG database. Indeed, in the Hallmark and the PID databases, gene sets are defined as gene signatures rather than as biological pathways. Thus, the genes are not necessarily connected to each other through functional interactions. Conversely, gene sets retrieved from the KEGG database correspond to biological pathways defined as genes corresponding to proteins that participate in oriented molecular cascades. They are available in the form of maps on the KEGG website. In addition, the structure of the KEGG database allows to build a network that provides mechanistic interpretation. Therefore, gene set enrichment algorithms required to build the signalling network was performed based on the KEGG database. All interactions and nodes from each biological pathway of the KEGG database were retrieved thanks to the OmnipathR R package (Türei et al., 2016).

3.3 Preprocessing of RNA-Seq data

Limma was originally developed for differential expression analysis of microarray data, which values are assumed to be normally distributed, and the variance independent of

the mean. This is not the case for log2-counts per million (log-CPM) values in RNA-Seq data: expression distributions may vary across samples and methods modelling counts assume a quadratic mean-variance relationship. Therefore, for the RNA-Seq data, 3 steps of pre-processing are necessary before applying the statistical tests (Law et al., 2018): (1) low expressed genes are filtered (i.e. genes with less than 10 read counts in at least one sample in the condition with the minimum sample size); (2) normalisation using the method of trimmed mean of M-values (TMM) is performed (Robinson and Oshlack, 2010); (3) raw counts are converted to log-CPM and the mean-variance relationship is estimated with the *voom* method.

3.4 Identification of Differentially Expressed Pathways (DEPs)

For each of the 10 transcriptomic datasets, identification of DEPs was performed using the *fgseaSimple* function of the Bioconductor package *fgsea* (Korotkevich et al., 2021), for fast preranked Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005).

The *fgseaSimple* method takes two inputs: a gene-level signed statistics and a defined list of genes known as *gene set*. The method ranks the genes in descending order based on the chosen statistics, and then computes the Enrichment Score (ES) for the gene set. The ES reflects how often members of that gene set occur at the top (e.g., upregulated) or the bottom (e.g., downregulated) of the ranked gene list. To account for differences in gene sets size, a normalisation step is performed to obtain the Normalised Enrichment Score (NES). Besides, random gene sets are generated and their NES computed. These NES are then used to create a null distribution from which the significance of the NES of the tested gene set is estimated. In our study, we used the t-statistics from the differential expression analysis comparing gene expression levels of CF sample to NCF samples as the control condition. In order to compare all the studies together, all the microarray and RNA-Seq datasets were processed using the same pipeline, involving the limma (Ritchie et al., 2015) and edgeR (Robinson et al.,

2010) packages. After removing technical outlier samples and retrieving gene symbols using the biomaRt package (Durinck et al., 2009), differential expression analysis at the gene level was performed by fitting a linear model using weighted least squares for each gene.

Gene sets with size larger than 500 were excluded for statistical testing. The p-values of the gene sets were adjusted for multiple testing error with Benjamini-Hochberg (BH) procedure. Differentially Expressed Pathways (DEP)s were considered with a corrected p-value lower or equal to 0.25. If the NES is positive, the DEP is categorized as up-regulated, and if it is negative, the DEP is categorized as down-regulated.

3.5 Up-dating Omnipath DEPs pathways

The CF network was built from DEPs among pathways in the KEGG database, as extracted with the OmniPathR package. We observed a few inconsistencies between the corresponding list of genes and interactions downloaded with OmnipathR R package, and those in the 'up to date' pathways maps, as they are displayed on KEGG website. Therefore, we updated the OmnipathR version of the KEGG pathways by adding (or removing) a few nodes or interactions, in order to map the OmnipathR pathways with their corresponding pathways in KEGG. For each modification, bibliographic references were manually checked into other databases stored in Omnipath, in particular in the high confident databases SignorDB (Lo Surdo et al., 2022), and the Human Reference Interactome (Drew et al., 2021). In addition, in a few pathways, some interactions are labelled as "indirect" in KEGG database. They involve part of signalling cascades belonging to other biological pathways, and they are not detailed in the considered pathway. For example, part of the PI3K-AKT pathway belongs to the Toll-like receptor signalling pathway but is not detailed in this pathway (See KEGG map for Toll-like receptor signalling pathway). In such cases, in order to build the

network based on complete cascades involving only direct interactions, we added the missing nodes and interactions.

All the pathways modifications and the corresponding codes used to perform these modifications are available in the following github repository: [sysbio-curie/CFnetwork](https://github.com/sysbio-curie/CFnetwork).

3.6 Network building and pruning

In the KEGG database, most of the 15 common DEPs display the same overall topology: some cell-surface receptor proteins activate one or more intra-cellular signalling cascades that in turn activate downstream transcription factors, thus triggering corresponding phenotypes. For example, the NF- κ B pathway leads to the "inflammation" or "cell survival" phenotypes. However, 2 of the common DEPs, *Cytokine-cytokine receptor interaction* and *Viral protein interaction with cytokine and cytokine receptor*, are pathways that do not consist in such functional cascades. The *Cytokine-cytokine receptor interaction* pathway consists in a list of interactions between extra-cellular signal molecules and cell-surface receptors (see KEGG database to visualise this pathway's topology). These interactions are also part of larger biological pathways that comprise their corresponding downstream cascades. This means that KEGG pathways are partially redundant (i.e. small pathways are part of larger pathways), which is also found in all commonly used pathway databases. In the case of the *Cytokine-cytokine receptor interaction* pathway, this DEP is dysregulated in the meta-analysis because some of the interactions between extracellular molecules and cell surface receptors are dysregulated, but not necessarily all of them. For example, interactions between TNF- α and its receptors, or IL17 and its receptors are dysregulated, but this information is also present in the DEPs containing the complete corresponding cascades, i.e. the *TNF- α signalling pathway* and the *IL-17 signalling pathway*. The same type of analysis also holds for the *Viral protein interaction with cytokine and cytokine receptor* DEP. Overall, from these 2 DEPs, we only retained the cell-surface receptors that

are sources of downstream dysregulated cascades in our network. Overall, 25 cell surface receptors without downstream dysregulations in our CF transcriptomic data were removed from the network.

Finally, we also removed from the pathways all the interactions corresponding to genes targeted by transcription factors, downstream of the pathways' cascades, because these target genes do not define the pathways themselves.

Network building and pruning were performed using the R packages *tidyr* v.1.2.1, and *dplyr* v.1.0.10. Transcription factors were identified using the R packages *dorothea* v.1.4.2 and *hgnc* v.0.1.2, which give access to the Dorothea (Garcia-Alonso et al., 2019) and HUGO collections (Seal et al., 2023), respectively.

3.7 Betweenness centrality score

The betweenness centrality (BC) score of node n is defined by

$$\sum_{i \neq j, i \neq n, j \neq n} p_{inj} / p_{ij}$$

where p_{ij} is the total number of shortest paths between nodes i and j while p_{inj} is the number of those shortest paths which pass through vertex n .

BC scores were computed using the *betweenness* function of the R package *igraph* v.1.3.4 (Csardi and Nepusz, 2005). This package was also used for the other network topology analyses.

3.8 Network Visualization and Figure Generation

The networks, generated as dataframes in R, were imported into Cytoscape v.3.9.0 (Shannon et al., 2003) for visualization. We designed a custom style for nodes and edges, which is available in the Cytoscape session and also saved as an independent XML file, available in the [sysbio-curie/CFnetwork_cytoscape](https://github.com/sysbio-curie/CFnetwork_cytoscape) github repository. The

hierarchical layout was used to emphasize the information flow from the source nodes to the sink nodes.

Barplots were generated using the R package ggplot2 v.3.3.6, and didactic figures were created using the open-source platform diagrams.net.

4 Discussion

Using a pathway-based meta-analysis of publicly available transcriptomic data, we built the CF network that provides a more global understanding of the molecular dysregulations in CF than the view of a CFTR-related channelopathy disease. Indeed, an important outcome of this work was to integrate data analyses to network reconstruction, while proposing a strategy to relate CFTR to proteins of the network, based on CFTR interactome. The CF network comprises a restricted number of source nodes that connect the absence of CFTR to the downstream sink nodes triggering CF cellular phenotypes. Another important contribution was to propose candidate therapeutic targets, based on the topological analysis of this network (namely, SYK, PI3KCA and PLCB1/3). The network provides a comprehensive view of how pathway interactions contribute to a given disease phenotype. It reveals unintuitive effects of targeting candidate proteins because of the complex interactions of the biological pathways in the network. Overall, the CF network can be seen as a tool to formulate hypotheses and interpret experimental observations.

Although several transcriptomic datasets were gathered, the total number of samples globally included remains modest (57 CF and 46 control samples). Additional data may refine the list of dysregulated pathways, and help to improve the proposed CF network.

To cope with the low number of samples per study, we opted for a meta-analysis combining various CF transcriptomic datasets, which highlighted that distinct differentiation media used for the primary cultures may favor different cell types, leading to

inconsistent transcriptomic profiles and potential erroneous interpretations. This may explain why previous transcriptomic comparative studies reported incoherent signs of gene dysregulation (up- versus down-) between different datasets for many genes (Clarke et al., 2013). We observed the same phenomenon at the pathway level for datasets belonging to subgroup 1 or 2 (see Section 2.3). Clustering studies based on the heatmap of common DEPs appears to be a good tool to select consistent data in future meta-analysis.

Other types of dysregulations such as aberrant phosphorylations are not detectable in transcriptomic data. Including information from other types of omic data such as proteomic, phosphoproteomic, metabolomic, or volatilomic may help to refine the CF network. In particular, in the past three years, CF airway epithelial single-cell RNAseq (sc-RNAseq) datasets have been reported (Carraro et al., 2021; Thurman et al., 2022). Such data allow the study of dysregulations at the cell type level, and could facilitate building of the CF networks for specific epithelial cell types. Furthermore, CFTR is expressed in cell types beyond airway epithelial cells. Thus, refining this network within the context of these cell types could enhance our understanding of the role of CFTR in these specific cells such as macrophages, where CFTR seems to have non-channel functions (Duan et al., 2021).

Prior knowledge gathered in the KEGG pathway database was used to identify and connect DEPs, but the proposed methodology can be followed using other pathways databases. Pathway names and definitions vary between databases, and therefore, the resulting network may slightly depend on the reference database that was used. Nevertheless, it would comprise globally the same interactions and proteins. Similarly, CFTR interactors present in the network were identified according to PPI information in the CyFi-MAP. If new CFTR interactors are identified, this information may help improve the content of the network, highlighting new source nodes or routes for the propagation of dysregulations. In particular, missing interactions, because they are

not present in pathway databases, or have not been discovered yet, may explain the presence of unconnected components. If they exist, their discovery in the future may allow to link the two unconnected small components to the main component of the network. However, the proposed notion of targeting proteins as upstream as possible in the network, or among key hubs of the network, are still an interesting concept in order to prioritize candidate therapeutic targets.

An important issue of the present paper was to explore the link between absence of the CFTR protein, and more global pathway dysregulations that lead to CF cellular phenotypes. However, the precise definition of a diseased cellular phenotype is not clearly defined yet, and we used key words provided in the KEGG database or in the Gene Cards database (Stelzer et al., 2016). The present work proposes an answer to this question in the context of systems biology studies. Associating phenotypes to the activity of outputs of the signalling cascades, referred to here as sink nodes, could be a first step towards the definition of the disease read-outs. This is of particular interest for *in vitro* evaluation of drug candidates, because we expect that drugs active in CF would reduce the activity of these sink nodes.

The methodological approach proposed in our study was settled based on transcriptomic data from hAEC cells homozygous for F508del, because publicly available data are more abundant for this most frequent mutation. Therefore, our CF network characterizes the disease caused by this mutation. It would be interesting to study to which extent the CF network would differ for other mutations. A recent paper indicates that DEGs in human bronchial epithelial cell lines bearing mutations from different classes share about 30% DEGs, while 70% of the DEGs are class specific (Santos et al., 2023). It would be interesting to study if this still holds at the level of biological pathways, as they are defined in the present work, and to study whether the resulting network is strongly modified, or not. The methodology proposed in the

present paper and based on network topology could still be applied in order to search for new, and potentially class-specific, therapeutic targets.

The candidate therapeutic targets proposed based on our CF network could also be tested on CF cellular models for other mutations, because these targets may belong to biological pathways that are also dysregulated with other mutations. If this was the case, it would help to extend the therapeutic arsenal available for CF patients who are not eligible for CFTR modulators.

In the same line, it is now clear that CF patients bearing the same mutation may present diseases of different severity. Although many factors can modulate disease severity, including environmental factors, it would be interesting to explore the contribution of patients molecular profiles. In particular, building a "personalized" CF network based on patients' transcriptomic profiles would be an interesting tool to answer this question.

Beyond CF, reduced amounts of functional CFTR have also been observed in other diseases like chronic obstructive pulmonary disease (COPD) (Saint-Criq and Gray, 2017; Simões et al., 2021), cigarette smoke (Valdivieso et al., 2018), or cancer (Duan et al., 2021; Wang et al., 2022). The network could provide a basis to explore the consequences of reduced CFTR activity in these diseases.

5 Conclusion

We presented building of **the CF network**, a signalling network gathering the molecular dysregulations caused by the absence of CFTR. We adopted a data-driven systems biology approach to retrieve CF dysregulated signalling pathways. These pathways were merged to build a signalling network, recapitulating the dysregulated cascades that flow from the source nodes (proteins connected to CFTR) to the sink nodes (proteins that trigger CF cellular phenotypes). Five of the source nodes are upstream of all the sink nodes in the CF network: PLCB1/3, TRADD, SRC, and SYK. These proteins

may collectively initiate the emergence of CF phenotypes (together with the other 3 source nodes EZR, CSNK2A1, and PRKCA), illustrating the complexity of the disease. The topological analysis of the network also highlighted nodes with a high degree of betweenness centrality, which are other important players in the propagation of the dysregulations, including PI3KCA. Among these key source nodes and nodes with high degree centrality, SYK, SRC, PLCB1/3 and PIK3CA appeared as interesting candidate therapeutic targets. Interestingly, specific inhibitors are known for these proteins, and even marketed drugs in the case of SYK and PI3KCA. They stand out as potential therapeutic candidates for drug repositioning, potentially allowing the modulation of various CF phenotypes. Finally, an important contribution of the present work is that the adopted global methodology of the CFTR context, although perfectible, did provide interesting results for CF, and can be used as a common framework for other monogenic diseases.

Additional File 1. CFnetwork_additional_file_1.pdf, a PDF document with extended results concerning the sink nodes and their corresponding cellular phenotypes:

- **Table S1:** The 35 sink nodes of the CF network and their corresponding cellular phenotypes.
- **Table S2:** The top 30 proteins in the CF network according to their betweenness centrality scores.

724 **List of abbreviations. :**

BC	Betweenness Centrality
CF	Cystic Fibrosis
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
DEG	Differentially Expressed Gene
DEP	Differentially Expressed Pathway
725 GSEA	Gene Set Enrichment Analysis
hAEC	human Airway Epithelial Cells
NCF	Non-cystic Fibrosis
NES	Normalised Enrichment Score
PPI	Protein-Protein Interaction

726 **Declarations**

727 **Ethics approval and consent to participate**

728 Not applicable

729 **Consent for publication**

730 Not applicable

731 **Availability of data and materials**

732 The codes and datasets supporting the conclusions of this article are available in the
733 following github repository: [sysbio-curie/CFnetwork](https://github.com/sysbio-curie/CFnetwork).

734 The Cytoscape session of the CF network, the TSV files of the nodes and the edges of
735 the CF network and the XML file of the custom style of the Cytoscape session required
736 to reproduce the Cytoscape session are available in the following github repository:
737 [sysbio-curie/CFnetwork_cytoscape](https://github.com/sysbio-curie/CFnetwork_cytoscape).

738 **Competing interests**

739 The authors declare that they have no competing interests.

740 **Funding**

741 Fondation pour la Recherche Médicale (FRM), Vaincre La Mucoviscidose (VLM), La
742 Fondation Dassault Systèmes, Fondation Maladies Rares, and MSD Avenir

743 **Author's contribution**

744 I.S. and V.S. initiated the project and obtained funding for the project. M.N., L.M.,
745 L.C. and V.S. contributed to the conception and design of the research. M.N. and L.M.
746 designed the transcriptomic analysis. M.N. conducted the transcriptomic analyses,
747 built the signalling network and performed network analyses under the supervision of
748 L.C. and V.S.. M.C., M.K.A. and I.S. contributed to the interpretation of the data.
749 M.N., L.C. and V.S. wrote the manuscript. L.M. I.S. and M.K.A edited the manuscript.
750 All authors read and approved the final manuscript.

751 **Acknowledgments**

752 The authors would like to thank Benoît Chevalier and Alexandre Hinzpeter for fruit-
753 ful discussions on CFTR interactome, and Marco Ruscone for fruitful discussions on
754 biological databases.

755 **References**

756 Guo J, Garratt A, Hill A. Worldwide rates of diagnosis and effective treatment for
757 cystic fibrosis. *Journal of Cystic Fibrosis*. 2022 May;21(3):456–462. <https://doi.org/10.1016/j.jcf.2022.01.009>.
758

- 759 Seibert FS, Loo TW, Clarke DM, Riordan JR. Cystic Fibrosis: Channel, Catalytic,
760 and Folding Properties of the CFTR Protein. *Journal of Bioenergetics and Biomem-*
761 *branes*. 1997 Oct;29(5):429–442. <https://doi.org/10.1023/A:1022478822214>.
- 762 Veit G, Avramescu RG, Chiang AN, Houck SA, Cai Z, Peters KW, et al. From
763 CFTR biology toward combinatorial pharmacotherapy: expanded classification of
764 cystic fibrosis mutations. *Molecular Biology of the Cell*. 2016 Feb;27(3):424–433.
765 Publisher: American Society for Cell Biology (mboc). [https://doi.org/10.1091/mbc.](https://doi.org/10.1091/mbc.e14-04-0935)
766 [e14-04-0935](https://doi.org/10.1091/mbc.e14-04-0935).
- 767 Hanssens LS, Duchateau J, Casimir GJ. CFTR Protein: Not Just a Chloride Chan-
768 nel? *Cells*. 2021 Nov;10(11):2844. Number: 11 Publisher: Multidisciplinary Digital
769 Publishing Institute. <https://doi.org/10.3390/cells10112844>.
- 770 Pereira C, Mazein A, Farinha CM, Gray MA, Kunzelmann K, Ostaszewski M, et al.
771 CyFi-MAP: an interactive pathway-based resource for cystic fibrosis. *Scientific*
772 *Reports*. 2021 Nov;11(1):22223. Number: 1 Publisher: Nature Publishing Group.
773 <https://doi.org/10.1038/s41598-021-01618-3>.
- 774 Farinha CM, Gentzsch M. Revisiting CFTR Interactions: Old Partners and New Play-
775 ers. *International Journal of Molecular Sciences*. 2021 Jan;22(24):13196. Number: 24
776 Publisher: Multidisciplinary Digital Publishing Institute. [https://doi.org/10.3390/](https://doi.org/10.3390/ijms222413196)
777 [ijms222413196](https://doi.org/10.3390/ijms222413196).
- 778 Crites KSM, Morin G, Orlando V, Patey N, Cantin C, Martel J, et al. CFTR Knock-
779 down induces proinflammatory changes in intestinal epithelial cells. *Journal of*
780 *Inflammation*. 2015 Nov;12(1):62. <https://doi.org/10.1186/s12950-015-0107-y>.
- 781 Fleurot I, López-Gálvez R, Barbry P, Guillon A, Si-Tahar M, Bähr A, et al. TLR5
782 signalling is hyper-responsive in porcine cystic fibrosis airways epithelium. *Journal*

783 of Cystic Fibrosis. 2022 Mar;21(2):e117–e121. [https://doi.org/10.1016/j.jcf.2021.](https://doi.org/10.1016/j.jcf.2021.08.002)
784 [08.002](https://doi.org/10.1016/j.jcf.2021.08.002).

785 Hao S, Roesch EA, Perez A, Weiner RL, Henderson LC, Cummings L, et al. Inactiva-
786 tion of CFTR by CRISPR/Cas9 alters transcriptional regulation of inflammatory
787 pathways and other networks. Journal of Cystic Fibrosis. 2020 Jan;19(1):34–39.
788 <https://doi.org/10.1016/j.jcf.2019.05.003>.

789 Jacquot J, Tabary O, Le Rouzic P, Clement A. Airway epithelial cell inflammatory sig-
790 nalling in cystic fibrosis. The International Journal of Biochemistry & Cell Biology.
791 2008 Jan;40(9):1703–1715. <https://doi.org/10.1016/j.biocel.2008.02.002>.

792 Jeanson L, Kelly M, Coste A, Guerrero IC, Fritsch J, Nguyen-Khoa
793 T, et al. Oxidative stress induces unfolding protein response and
794 inflammation in nasal polyposis. Allergy. 2012;67(3):403–412. eprint:
795 <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1398-9995.2011.02769.x>.
796 <https://doi.org/10.1111/j.1398-9995.2011.02769.x>.

797 Conese M, Di Gioia S. Pathophysiology of Lung Disease and Wound Repair
798 in Cystic Fibrosis. Pathophysiology. 2021 Mar;28(1):155–188. Number: 1
799 Publisher: Multidisciplinary Digital Publishing Institute. [https://doi.org/10.3390/](https://doi.org/10.3390/pathophysiology28010011)
800 [pathophysiology28010011](https://doi.org/10.3390/pathophysiology28010011).

801 Pankonien I, Quaresma MC, Rodrigues CS, Amaral MD. CFTR, Cell Junctions and
802 the Cytoskeleton. International Journal of Molecular Sciences. 2022 Jan;23(5):2688.
803 Number: 5 Publisher: Multidisciplinary Digital Publishing Institute. [https://doi.](https://doi.org/10.3390/ijms23052688)
804 [org/10.3390/ijms23052688](https://doi.org/10.3390/ijms23052688).

805 Hornberg JJ, Bruggeman FJ, Westerhoff HV, Lankelma J. Cancer: A Systems Biology
806 disease. Biosystems. 2006 Feb;83(2):81–90. [https://doi.org/10.1016/j.biosystems.](https://doi.org/10.1016/j.biosystems.2005.05.014)
807 [2005.05.014](https://doi.org/10.1016/j.biosystems.2005.05.014).

- 808 Pankow S, Bamberger C, Calzolari D, Martínez-Bartolomé S, Lavallée-Adam M, Balch
809 WE, et al. Δ F508 CFTR interactome remodelling promotes rescue of cystic fibrosis.
810 Nature. 2015 Dec;528(7583):510–516. Number: 7583 Publisher: Nature Publishing
811 Group. <https://doi.org/10.1038/nature15729>.
- 812 Trivedi TS, Bhadresha KP, Patel MP, Mankad AU, Rawal RM, Patel SK. Identifica-
813 tion of hub genes associated with human cystic fibrosis: A Meta-analysis approach.
814 Human Gene. 2023 Feb;35:201139. <https://doi.org/10.1016/j.humgen.2022.201139>.
- 815 Ideozu JE, Zhang X, McColley S, Levy H. Transcriptome Profiling and Molecular
816 Therapeutic Advances in Cystic Fibrosis: Recent Insights. Genes. 2019 Feb;10(3).
817 <https://doi.org/10.3390/genes10030180>.
- 818 Clarke LA, Sousa L, Barreto C, Amaral MD. Changes in transcriptome of native
819 nasal epithelium expressing F508del-CFTR and intersecting data from compa-
820 rable studies. Respiratory Research. 2013 Mar;14:38. [https://doi.org/10.1186/](https://doi.org/10.1186/1465-9921-14-38)
821 [1465-9921-14-38](https://doi.org/10.1186/1465-9921-14-38).
- 822 Verhaeghe C, Remouchamps C, Hennuy B, Vanderplasschen A, Chariot A, Tabruyn
823 SP, et al. Role of IKK and ERK pathways in intrinsic inflammation of cystic
824 fibrosis airways. Biochemical Pharmacology. 2007 Jun;73(12):1982–1994. <https://doi.org/10.1016/j.bcp.2007.03.019>.
825 <https://doi.org/10.1016/j.bcp.2007.03.019>.
- 826 Ogilvie V, Passmore M, Hyndman L, Jones L, Stevenson B, Wilson A, et al. Dif-
827 ferential global gene expression in cystic fibrosis nasal and bronchial epithelium.
828 Genomics. 2011 Nov;98(5):327–336. <https://doi.org/10.1016/j.ygeno.2011.06.008>.
- 829 Voisin G, Bouvet GF, Legendre P, Dagenais A, Massé C, Berthiaume Y. Oxida-
830 tive stress modulates the expression of genes involved in cell survival in
831 Δ F508 cystic fibrosis airway epithelial cells. Physiological Genomics. 2014

- 832 Jun;46(17):634–646. Publisher: American Physiological Society. <https://doi.org/10.1152/physiolgenomics.00003.2014>.
833
- 834 Balloy V, Varet H, Dillies MA, Proux C, Jagla B, Coppée JY, et al. Normal and Cystic
835 Fibrosis Human Bronchial Epithelial Cells Infected with *Pseudomonas aeruginosa*
836 Exhibit Distinct Gene Activation Patterns. PLOS ONE. 2015 Oct;10(10):e0140979.
837 Publisher: Public Library of Science. <https://doi.org/10.1371/journal.pone.0140979>.
- 838 Zoso A, Sofoluwe A, Bacchetta M, Chanson M. Transcriptomic profile of cystic
839 fibrosis airway epithelial cells undergoing repair. Scientific Data. 2019 Oct;6(1):1–
840 7. Number: 1 Publisher: Nature Publishing Group. <https://doi.org/10.1038/s41597-019-0256-6>.
841
- 842 Ling KM, Garratt LW, Gill EE, Lee AHY, Agudelo-Romero P, Sutanto EN, et al.
843 Rhinovirus Infection Drives Complex Host Airway Molecular Responses in Children
844 With Cystic Fibrosis. Frontiers in Immunology. 2020 Jul;11:1327. <https://doi.org/10.3389/fimmu.2020.01327>.
845
- 846 Saint-Criq V, Delpiano L, Casement J, Onuora JC, Lin J, Gray MA. Choice of Differen-
847 tiation Media Significantly Impacts Cell Lineage and Response to CFTR Modulators
848 in Fully Differentiated Primary Cultures of Cystic Fibrosis Human Airway Epithe-
849 lial Cells. Cells. 2020 Sep;9(9):2137. Number: 9 Publisher: Multidisciplinary Digital
850 Publishing Institute. <https://doi.org/10.3390/cells9092137>.
- 851 Virella-Lowell I, Herlihy JD, Liu B, Lopez C, Cruz P, Muller C, et al. Effects of
852 CFTR, interleukin-10, and *Pseudomonas aeruginosa* on gene expression profiles in
853 a CF bronchial epithelial cell Line. Molecular Therapy. 2004 Sep;10(3):562–573.
854 Publisher: Elsevier. <https://doi.org/10.1016/j.ymthe.2004.06.215>.

- 855 Rehman T, Karp PH, Tan P, Goodell BJ, Pezzulo AA, Thurman AL, et al.
856 Inflammatory cytokines TNF- α and IL-17 enhance the efficacy of cystic fibro-
857 sis transmembrane conductance regulator modulators. *The Journal of Clinical*
858 *Investigation*. 2021 Aug;131(16):150398. <https://doi.org/10.1172/JCI150398>.
- 859 Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The
860 Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell*
861 *systems*. 2015 Dec;1(6):417–425. <https://doi.org/10.1016/j.cels.2015.12.004>.
- 862 Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T, et al. PID: the
863 Pathway Interaction Database. *Nucleic Acids Research*. 2009 Jan;37(Database
864 issue):D674–D679. <https://doi.org/10.1093/nar/gkn653>.
- 865 Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG:
866 integrating viruses and cellular organisms. *Nucleic Acids Research*. 2021
867 Jan;49(D1):D545–D551. <https://doi.org/10.1093/nar/gkaa970>.
- 868 Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide asso-
869 ciation studies. *Nature Reviews Genetics*. 2010 Dec;11(12):843–854. Number: 12
870 Publisher: Nature Publishing Group. <https://doi.org/10.1038/nrg2884>.
- 871 Martignetti L, Calzone L, Bonnet E, Barillot E, Zinovyev A. ROMA: Representation
872 and Quantification of Module Activity from Target Expression Data. *Frontiers in*
873 *Genetics*. 2016 Feb;7. <https://doi.org/10.3389/fgene.2016.00018>.
- 874 Landais Y, Vallot C. Multi-modal quantification of pathway activity with MAYA.
875 *Nature Communications*. 2023 Mar;14(1):1668. Number: 1 Publisher: Nature
876 Publishing Group. <https://doi.org/10.1038/s41467-023-37410-2>.
- 877 Schubert M, Klinger B, Klünemann M, Sieber A, Uhlitz F, Sauer S, et al. Perturbation-
878 response genes reveal signaling footprints in cancer gene expression. *Nature*

- 879 Communications. 2018 Jan;9(1):20. Number: 1 Publisher: Nature Publishing Group.
880 <https://doi.org/10.1038/s41467-017-02391-6>.
- 881 Vaske CJ, Benz SC, Sanborn JZ, Earl D, Szeto C, Zhu J, et al. Inference of
882 patient-specific pathway activities from multi-dimensional cancer genomics data
883 using PARADIGM. Bioinformatics. 2010 Jun;26(12):i237–i245. <https://doi.org/10.1093/bioinformatics/btq182>.
884
- 885 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA,
886 et al. Gene set enrichment analysis: A knowledge-based approach for interpreting
887 genome-wide expression profiles. Proceedings of the National Academy of Sciences.
888 2005 Oct;102(43):15545–15550. Publisher: National Academy of Sciences Section:
889 Biological Sciences. <https://doi.org/10.1073/pnas.0506580102>.
- 890 Hsu D, Taylor P, Fletcher D, van Heeckeren R, Eastman J, van Heeckeren A, et al.
891 Interleukin-17 Pathophysiology and Therapeutic Intervention in Cystic Fibrosis
892 Lung Infection and Inflammation. Infection and Immunity. 2016 Aug;84(9):2410–
893 2421. Publisher: American Society for Microbiology. <https://doi.org/10.1128/iai.00284-16>.
894
- 895 Dumortier C, Danopoulos S, Velard F, Al Alam D. Bone Cells Differentiation: How
896 CFTR Mutations May Rule the Game of Stem Cells Commitment? Frontiers in
897 Cell and Developmental Biology. 2021;9.
- 898 Kosamo S, Hisert KB, Dmyterko V, Nguyen C, Black RA, Holden TD, et al. Strong
899 toll-like receptor responses in cystic fibrosis patients are associated with higher lung
900 function. Journal of Cystic Fibrosis. 2020 Jul;19(4):608–613. <https://doi.org/10.1016/j.jcf.2019.11.009>.
901

- 902 Curutiu C, Iordache F, Lazar V, Pisoschi AM, Pop A, Chifiriuc MC, et al. Impact
903 of *Pseudomonas aeruginosa* quorum sensing signaling molecules on adhesion and
904 inflammatory markers in endothelial cells. Beilstein Journal of Organic Chemistry.
905 2018 Oct;14:2580–2588. <https://doi.org/10.3762/bjoc.14.235>.
- 906 Türei D, Korcsmáros T, Saez-Rodriguez J. OmniPath: guidelines and gate-
907 way for literature-curated signaling pathway resources. Nature Methods. 2016
908 Dec;13(12):966–967. Number: 12 Publisher: Nature Publishing Group. <https://doi.org/10.1038/nmeth.4077>.
- 910 Venerando A, Pagano MA, Tosoni K, Meggio F, Cassidy D, Stobbart M, et al.
911 Understanding protein kinase CK2 mis-regulation upon F508del CFTR expres-
912 sion. Naunyn-Schmiedeberg's Archives of Pharmacology. 2011 Oct;384(4):473–488.
913 <https://doi.org/10.1007/s00210-011-0650-x>.
- 914 Wang H, Cebotaru L, Lee HW, Yang Q, Pollard BS, Pollard HB, et al. CFTR Con-
915 trols the Activity of NF- κ B by Enhancing the Degradation of TRADD. Cellular
916 Physiology and Biochemistry. 2016;40(5):1063–1078. Publisher: Karger Publishers.
917 <https://doi.org/10.1159/000453162>.
- 918 Massip Copiz MM, Santa Coloma TA. c- Src and its role in cystic fibrosis. European
919 Journal of Cell Biology. 2016 Oct;95(10):401–413. <https://doi.org/10.1016/j.ejcb.2016.08.001>.
- 921 Rimessi A, Bezzerri V, Salvatori F, Tamanini A, Nigro F, Dehecchi MC, et al.
922 PLCB3 Loss of Function Reduces *Pseudomonas aeruginosa*–Dependent IL-8 Release
923 in Cystic Fibrosis. American Journal of Respiratory Cell and Molecular Biol-
924 ogy. 2018 Oct;59(4):428–436. Publisher: American Thoracic Society - AJRCMB.
925 <https://doi.org/10.1165/rcmb.2017-0267OC>.

- 926 Favia M, Guerra L, Fanelli T, Cardone RA, Monterisi S, Di Sole F, et al. Na⁺/H⁺
927 Exchanger Regulatory Factor 1 Overexpression-dependent Increase of Cytoskeleton
928 Organization Is Fundamental in the Rescue of F508del Cystic Fibrosis Transmem-
929 brane Conductance Regulator in Human Airway CFBE41o- Cells. Molecular Biology
930 of the Cell. 2010 Jan;21(1):73–86. Publisher: American Society for Cell Biology
931 (mboc). <https://doi.org/10.1091/mbc.e09-03-0185>.
- 932 Wu Q, Eickelberg O. Ezrin in Asthma: A First Step to Early Biomarkers of
933 Airway Epithelial Dysfunction. American Journal of Respiratory and Critical
934 Care Medicine. 2019 Feb;199(4):408–410. Publisher: American Thoracic Society -
935 AJRCCM. <https://doi.org/10.1164/rccm.201810-1964ED>.
- 936 Mendes AI, Matos P, Moniz S, Luz S, Amaral MD, Farinha CM, et al. Antagonistic
937 Regulation of Cystic Fibrosis Transmembrane Conductance Regulator Cell Surface
938 Expression by Protein Kinases WNK4 and Spleen Tyrosine Kinase. Molecular and
939 Cellular Biology. 2011 Oct;31(19):4076–4086. Publisher: Taylor & Francis .eprint:
940 <https://doi.org/10.1128/MCB.05152-11>. <https://doi.org/10.1128/MCB.05152-11>.
- 941 Egan M, Flotte T, Afione S, Solow R, Zeitlin PL, Carter BJ, et al. Defective regulation
942 of outwardly rectifying Cl channels by protein kinase A corrected by insertion of
943 CFTR. Nature. 1992 Aug;358(6387):581–584. Number: 6387 Publisher: Nature
944 Publishing Group. <https://doi.org/10.1038/358581a0>.
- 945 Bérubé J, Roussel L, Nattagh L, Rousseau S. Loss of Cystic Fibrosis Transmem-
946 brane Conductance Regulator Function Enhances Activation of p38 and ERK
947 MAPKs, Increasing Interleukin-6 Synthesis in Airway Epithelial Cells Exposed to
948 Pseudomonas aeruginosa. Journal of Biological Chemistry. 2010 Jul;285(29):22299–
949 22307. Publisher: American Society for Biochemistry and Molecular Biology.
950 <https://doi.org/10.1074/jbc.M109.098566>.

- 951 Wellmerling J, Rayner RE, Chang SW, Kairis EL, Kim SH, Sharma A, et al. Tar-
952 geting the EGFR-ERK axis using the compatible solute ectoine to stabilize CFTR
953 mutant F508del. FASEB journal: official publication of the Federation of Ameri-
954 can Societies for Experimental Biology. 2022 May;36(5):e22270. [https://doi.org/10.](https://doi.org/10.1096/fj.202100458RRR)
955 [1096/fj.202100458RRR](https://doi.org/10.1096/fj.202100458RRR).
- 956 Cannon CL, Kowalski MP, Stopak KS, Pier GB. Pseudomonas aeruginosa-Induced
957 Apoptosis Is Defective in Respiratory Epithelial Cells Expressing Mutant Cystic
958 Fibrosis Transmembrane Conductance Regulator. American Journal of Respiratory
959 Cell and Molecular Biology. 2003 Aug;29(2):188–197. Publisher: American Thoracic
960 Society - AJRCMB. <https://doi.org/10.1165/rcmb.4898>.
- 961 Gottlieb RA, Dosanjh A. Mutant cystic fibrosis transmembrane conductance regu-
962 lator inhibits acidification and apoptosis in C127 cells: possible relevance to cystic
963 fibrosis. Proceedings of the National Academy of Sciences. 1996 Apr;93(8):3587–
964 3591. Publisher: Proceedings of the National Academy of Sciences. [https://doi.org/](https://doi.org/10.1073/pnas.93.8.3587)
965 [10.1073/pnas.93.8.3587](https://doi.org/10.1073/pnas.93.8.3587).
- 966 Chen Q, Pandi SPS, Kerrigan L, McElvaney NG, Greene CM, Elborn JS, et al. Cystic
967 fibrosis epithelial cells are primed for apoptosis as a result of increased Fas (CD95).
968 Journal of Cystic Fibrosis. 2018 Sep;17(5):616–623. [https://doi.org/10.1016/j.jcf.](https://doi.org/10.1016/j.jcf.2018.01.010)
969 [2018.01.010](https://doi.org/10.1016/j.jcf.2018.01.010).
- 970 Yalçın E, Talim B, Özçelik U, Doğru D, Çobanoğlu N, Pekcan S, et al. Does Defec-
971 tive Apoptosis Play A Role in Cystic Fibrosis Lung Disease? Archives of Medical
972 Research. 2009 Oct;40(7):561–564. <https://doi.org/10.1016/j.arcmed.2009.07.005>.
- 973 Rottner M, Kunzelmann C, Mergey M, Freyssinet JM, Martínez MC. Exagger-
974 ated apoptosis and NF-kappaB activation in pancreatic and tracheal cystic fibrosis
975 cells. FASEB journal: official publication of the Federation of American Societies

- 976 for Experimental Biology. 2007 Sep;21(11):2939–2948. [https://doi.org/10.1096/fj.](https://doi.org/10.1096/fj.06-7614com)
977 [06-7614com](https://doi.org/10.1096/fj.06-7614com).
- 978 Lasalvia M, Castellani S, D’Antonio P, Perna G, Carbone A, Colia AL, et al. Human
979 airway epithelial cells investigated by atomic force microscopy: A hint to cystic
980 fibrosis epithelial pathology. Experimental Cell Research. 2016 Oct;348(1):46–55.
981 <https://doi.org/10.1016/j.yexcr.2016.08.025>.
- 982 Burat B, Reynaerts A, Baiwir D, Fléron M, Gohy S, Eppe G, et al. Sweat
983 Proteomics in Cystic Fibrosis: Discovering Companion Biomarkers for Precision
984 Medicine and Therapeutic Development. Cells. 2022 Jan;11(15):2358. Number: 15
985 Publisher: Multidisciplinary Digital Publishing Institute. [https://doi.org/10.3390/](https://doi.org/10.3390/cells11152358)
986 [cells11152358](https://doi.org/10.3390/cells11152358).
- 987 De Lisle RC. Disrupted tight junctions in the small intestine of cystic fibrosis
988 mice. Cell and Tissue Research. 2014 Jan;355(1):131–142. [https://doi.org/10.1007/](https://doi.org/10.1007/s00441-013-1734-3)
989 [s00441-013-1734-3](https://doi.org/10.1007/s00441-013-1734-3).
- 990 Castellani S, Guerra L, Favia M, Di Gioia S, Casavola V, Conese M. NHERF1 and
991 CFTR restore tight junction organisation and function in cystic fibrosis airway
992 epithelial cells: role of ezrin and the RhoA/ROCK pathway. Laboratory Investiga-
993 tion. 2012 Nov;92(11):1527–1540. Number: 11 Publisher: Nature Publishing Group.
994 <https://doi.org/10.1038/labinvest.2012.123>.
- 995 Gillan JL, Chokshi M, Hardisty GR, Clohisey Hendry S, Prasca-Chamorro D, Robin-
996 son NJ, et al. CAGE sequencing reveals CFTR-dependent dysregulation of type
997 I IFN signaling in activated cystic fibrosis macrophages. Science Advances. 2023
998 May;9(21):eadg5128. Publisher: American Association for the Advancement of
999 Science. <https://doi.org/10.1126/sciadv.adg5128>.

- 1000 Dugger DT, Fung M, Zlock L, Caldera S, Sharp L, Hays SR, et al. Cystic Fibrosis
1001 Lung Transplant Recipients Have Suppressed Airway Interferon Responses during
1002 Pseudomonas Infection. Cell Reports Medicine. 2020 Jul;1(4):100055. <https://doi.org/10.1016/j.xcrm.2020.100055>.
1003
- 1004 Durón C, Pan Y, Gutmann DH, Hardin J, Radunskaya A. Variability of Betweenness
1005 Centrality and Its Effect on Identifying Essential Genes. Bulletin of Mathematical
1006 Biology. 2019 Sep;81(9):3655–3673. <https://doi.org/10.1007/s11538-018-0526-z>.
- 1007 Ferenc Karpati LH Bengt Wretlind. TNF-A and IL-8 in Consecutive Sputum Sam-
1008 ples from Cystic Fibrosis Patients During Antibiotic Treatment. Scandinavian
1009 Journal of Infectious Diseases. 2000 Jan;32(1):75–79. Publisher: Taylor & Fran-
1010 cis _eprint: <https://doi.org/10.1080/00365540050164263>. [https://doi.org/10.1080/](https://doi.org/10.1080/00365540050164263)
1011 [00365540050164263](https://doi.org/10.1080/00365540050164263).
- 1012 Riccaboni M, Bianchi I, Petrillo P. Spleen tyrosine kinases: biology, therapeutic targets
1013 and drugs. Drug Discovery Today. 2010 Jul;15(13):517–530. [https://doi.org/10.](https://doi.org/10.1016/j.drudis.2010.05.001)
1014 [1016/j.drudis.2010.05.001](https://doi.org/10.1016/j.drudis.2010.05.001).
- 1015 Wong BR, Grossbard EB, Payan DG, Masuda ES. Targeting Syk as a treat-
1016 ment for allergic and autoimmune disorders. Expert Opinion on Investiga-
1017 tional Drugs. 2004 Jul;13(7):743–762. Publisher: Taylor & Francis _eprint:
1018 <https://doi.org/10.1517/13543784.13.7.743>. [https://doi.org/10.1517/13543784.13.](https://doi.org/10.1517/13543784.13.7.743)
1019 [7.743](https://doi.org/10.1517/13543784.13.7.743).
- 1020 Bezzetti V, d'Adamo P, Rimessi A, Lanzara C, Crovella S, Nicolis E, et al. Phospho-
1021 lipase C- B3 Is a Key Modulator of IL-8 Expression in Cystic Fibrosis Bronchial
1022 Epithelial Cells. The Journal of Immunology. 2011 Apr;186(8):4946–4958. <https://doi.org/10.4049/jimmunol.1003535>.
1023

- 1024 Bell SC, Mall MA, Gutierrez H, Macek M, Madge S, Davies JC, et al. The
1025 Lancet Respiratory Medicine Commission on the Future of Care of Cystic Fibrosis.
1026 The Lancet Respiratory medicine. 2020 Jan;8(1):65–124. [https://doi.org/10.1016/](https://doi.org/10.1016/S2213-2600(19)30337-6)
1027 [S2213-2600\(19\)30337-6](https://doi.org/10.1016/S2213-2600(19)30337-6).
- 1028 Cooper N, Ghanima W, Hill QA, Nicolson PL, Markovtsov V, Kessler C. Recent
1029 advances in understanding spleen tyrosine kinase (SYK) in human biology and
1030 disease, with a focus on fostamatinib. Platelets. 2023 Dec;34(1):2131751. Pub-
1031 lisher: Taylor & Francis _eprint: <https://doi.org/10.1080/09537104.2022.2131751>.
1032 <https://doi.org/10.1080/09537104.2022.2131751>.
- 1033 Fiorotto R, Amenduni M, Mariotti V, Fabris L, Spirli C, Strazzabosco M. Src
1034 kinase inhibition reduces inflammatory and cytoskeletal changes in $\Delta F508$ human
1035 cholangiocytes and improves cystic fibrosis transmembrane conductance regulator
1036 correctors efficacy: Fiorotto, Amenduni, et al. Hepatology. 2018 Mar;67(3):972–988.
1037 Number: 3. <https://doi.org/10.1002/hep.29400>.
- 1038 Natarajan V. Is PI3K a Villain in Cystic Fibrosis? American Journal of Respiratory
1039 Cell and Molecular Biology. 2020 May;62(5):552–553. Publisher: American Thoracic
1040 Society - AJRCMB. <https://doi.org/10.1165/rcmb.2020-0029ED>.
- 1041 Dorfman R, Sandford A, Taylor C, Huang B, Frangolias D, Wang Y, et al. Complex
1042 two-gene modulation of lung disease severity in children with cystic fibrosis. The
1043 Journal of Clinical Investigation. 2008 Mar;118(3):1040–1049. Publisher: American
1044 Society for Clinical Investigation. <https://doi.org/10.1172/JCI33754>.
- 1045 Sagwal S, Chauhan A, Kaur J, Prasad R, Singh M, Singh M. Association of Serum
1046 TGF-B1 Levels with Different Clinical Phenotypes of Cystic Fibrosis Exacerbation.
1047 Lung. 2020 Apr;198(2):377–383. <https://doi.org/10.1007/s00408-020-00320-x>.

- 1048 Kramer EL, Clancy JP. TGFB as a therapeutic target in cystic fibrosis. Expert Opin-
1049 ion on Therapeutic Targets. 2018 Feb;22(2):177–189. Publisher: Taylor & Francis
1050 eprint: [https://doi.org/10.1080/](https://doi.org/10.1080/14728222.2018.1406922)
1051 [14728222.2018.1406922](https://doi.org/10.1080/14728222.2018.1406922).
- 1052 Zabner J, Scheetz TE, Almabrazi HG, Casavant TL, Huang J, Keshavjee S, et al.
1053 CFTR Δ F508 mutation has minimal effect on the gene expression profile of
1054 differentiated human airway epithelia. American Journal of Physiology-Lung Cel-
1055 lular and Molecular Physiology. 2005 Oct;289(4):L545–L553. Publisher: American
1056 Physiological Society. <https://doi.org/10.1152/ajplung.00065.2005>.
- 1057 Wright JM, Merlo CA, Reynolds JB, Zeitlin PL, Garcia JGN, Guggino WB, et al.
1058 Respiratory epithelial gene expression in patients with mild and severe cystic fibrosis
1059 lung disease. American Journal of Respiratory Cell and Molecular Biology. 2006
1060 Sep;35(3):327–336. <https://doi.org/10.1165/rcmb.2005-0359OC>.
- 1061 Bampi GB, Rauscher R, Kirchner S, Oliver KE, Bijvelds MJC, Santos LA, et al.
1062 Global assessment of the integrated stress response in CF patient-derived airway
1063 and intestinal tissues. Journal of Cystic Fibrosis. 2020 Nov;19(6):1021–1026. <https://doi.org/10.1016/j.jcf.2020.04.005>.
- 1065 Veltman M, De Sanctis JB, Stolarczyk M, Klymiuk N, Bähr A, Brouwer RW, et al.
1066 CFTR Correctors and Antioxidants Partially Normalize Lipid Imbalance but not
1067 Abnormal Basal Inflammatory Cytokine Profile in CF Bronchial Epithelial Cells.
1068 Frontiers in Physiology. 2021;12.
- 1069 Law CW, Alhamdoosh M, Su S, Dong X, Tian L, Smyth GK, et al. RNA-seq analysis
1070 is easy as 1-2-3 with limma, Glimma and edgeR. F1000Research. 2018 Dec;5:ISCB
1071 Comm J–1408. <https://doi.org/10.12688/f1000research.9005.3>.

1072 Robinson MD, Oshlack A. A scaling normalization method for differential expression
1073 analysis of RNA-seq data. *Genome Biology*. 2010 Mar;11(3):R25. [https://doi.org/](https://doi.org/10.1186/gb-2010-11-3-r25)
1074 [10.1186/gb-2010-11-3-r25](https://doi.org/10.1186/gb-2010-11-3-r25).

1075 Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A.:
1076 Fast gene set enrichment analysis. *bioRxiv*. Pages: 060012 Section: New Results.
1077 Available from: <https://www.biorxiv.org/content/10.1101/060012v3>.

1078 Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential
1079 expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids*
1080 *Research*. 2015 Apr;43(7):e47. <https://doi.org/10.1093/nar/gkv007>.

1081 Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for dif-
1082 ferential expression analysis of digital gene expression data. *Bioinformatics*. 2010
1083 Jan;26(1):139–140. <https://doi.org/10.1093/bioinformatics/btp616>.

1084 Durinck S, Spellman PT, Birney E, Huber W. Mapping Identifiers for the Integration
1085 of Genomic Datasets with the R/Bioconductor package biomaRt. *Nature protocols*.
1086 2009;4(8):1184–1191. <https://doi.org/10.1038/nprot.2009.97>.

1087 Lo Surdo P, Iannuccelli M, Contino S, Castagnoli L, Licata L, Cesareni G, et al.
1088 SIGNOR 3.0, the SIGnaling network open resource 3.0: 2022 update. *Nucleic acids*
1089 *research*. 2022 Oct;p. gkac883. <https://doi.org/10.1093/nar/gkac883>.

1090 Drew K, Wallingford JB, Marcotte EM. hu.MAP 2.0: integration of over 15,000 pro-
1091 teomic experiments builds a global compendium of human multiprotein assemblies.
1092 *Molecular Systems Biology*. 2021 May;17(5):e10016. Publisher: John Wiley & Sons,
1093 Ltd. <https://doi.org/10.15252/msb.202010016>.

- 1094 Garcia-Alonso L, Holland CH, Ibrahim MM, Turei D, Saez-Rodriguez J. Benchmark
1095 and integration of resources for the estimation of human transcription factor activ-
1096 ities. *Genome Research*. 2019 Jan;29(8):1363–1375. Company: Cold Spring Harbor
1097 Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution:
1098 Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press
1099 Publisher: Cold Spring Harbor Lab. <https://doi.org/10.1101/gr.240663.118>.
- 1100 Seal RL, Braschi B, Gray K, Jones TEM, Tweedie S, Haim-Vilmovsky L, et al.
1101 Genenames.org: the HGNC resources in 2023. *Nucleic Acids Research*. 2023
1102 Jan;51(D1):D1003–D1009. <https://doi.org/10.1093/nar/gkac888>.
- 1103 Csardi G, Nepusz T. The Igraph Software Package for Complex Network Research.
1104 *InterJournal*. 2005 Nov;Complex Systems:1695.
- 1105 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape:
1106 A Software Environment for Integrated Models of Biomolecular Interaction Net-
1107 works. *Genome Research*. 2003 Jan;13(11):2498–2504. Company: Cold Spring
1108 Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institu-
1109 tion: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory
1110 Press Publisher: Cold Spring Harbor Lab. <https://doi.org/10.1101/gr.1239303>.
- 1111 Carraro G, Langerman J, Sabri S, Lorenzana Z, Purkayastha A, Zhang G, et al. Tran-
1112 scriptional analysis of cystic fibrosis airways at single-cell resolution reveals altered
1113 epithelial cell states and composition. *Nature Medicine*. 2021 May;27(5):806–814.
1114 Bandiera_abtest: a Cg_type: Nature Research Journals Number: 5 Primary_atype:
1115 Research Publisher: Nature Publishing Group Subject_term: Diseases;Respiratory
1116 tract diseases Subject_term_id: diseases;respiratory-tract-diseases. [https://doi.org/](https://doi.org/10.1038/s41591-021-01332-7)
1117 [10.1038/s41591-021-01332-7](https://doi.org/10.1038/s41591-021-01332-7).

- 1118 Thurman AL, Li X, Villacreses R, Yu W, Gong H, Mather SE, et al. A Single-Cell
1119 Atlas of Large and Small Airways at Birth in a Porcine Model of Cystic Fibrosis.
1120 American Journal of Respiratory Cell and Molecular Biology. 2022 Jun;66(6):612–
1121 622. Publisher: American Thoracic Society - AJRCMB. [https://doi.org/10.1165/
1122 rcmb.2021-0499OC](https://doi.org/10.1165/rcmb.2021-0499OC).
- 1123 Duan Y, Li G, Xu M, Qi X, Deng M, Lin X, et al. CFTR is a negative regulator
1124 of $\gamma\delta$ T cell IFN- γ production and antitumor immunity. Cellular & Molecular
1125 Immunology. 2021 Aug;18(8):1934–1944. Number: 8 Publisher: Nature Publishing
1126 Group. <https://doi.org/10.1038/s41423-020-0499-3>.
- 1127 Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S,
1128 et al. The GeneCards Suite: From Gene Data Mining to Disease Genome
1129 Sequence Analyses. Current Protocols in Bioinformatics. 2016;54(1):1.30.1–
1130 1.30.33. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpbi.5>. [https://
1131 doi.org/10.1002/cpbi.5](https://doi.org/10.1002/cpbi.5).
- 1132 Santos L, Nascimento R, Duarte A, Railean V, Amaral MD, Harrison PT, et al.
1133 Mutation-class dependent signatures outweigh disease-associated processes in cys-
1134 tic fibrosis cells. Cell & Bioscience. 2023 Feb;13(1):26. [https://doi.org/10.1186/
1135 s13578-023-00975-y](https://doi.org/10.1186/s13578-023-00975-y).
- 1136 Saint-Criq V, Gray MA. Role of CFTR in epithelial physiology. Cellular and Molecular
1137 Life Sciences. 2017;74(1):93–115. <https://doi.org/10.1007/s00018-016-2391-y>.
- 1138 Simões FB, Kmit A, Amaral MD. Cross-talk of inflammatory mediators and airway
1139 epithelium reveals the cystic fibrosis transmembrane conductance regulator as a
1140 major target. ERJ Open Research. 2021 Dec;7(4):00247–2021. [https://doi.org/10.
1141 1183/23120541.00247-2021](https://doi.org/10.1183/23120541.00247-2021).

- 1142 Valdivieso AG, Dugour AV, Sotomayor V, Clauzure M, Figueroa JM, Santa-Coloma
1143 TA. N-acetyl cysteine reverts the proinflammatory state induced by cigarette smoke
1144 extract in lung Calu-3 cells. Redox Biology. 2018 Jun;16:294–302. [https://doi.org/](https://doi.org/10.1016/j.redox.2018.03.006)
1145 [10.1016/j.redox.2018.03.006](https://doi.org/10.1016/j.redox.2018.03.006).
- 1146 Wang Y, Tang L, Yang L, Lv P, Mai S, Xu L, et al. DNA Methylation-
1147 Mediated Low Expression of CFTR Stimulates the Progression of Lung Adenocar-
1148 cinoma. Biochemical Genetics. 2022 Apr;60(2):807–821. [https://doi.org/10.1007/](https://doi.org/10.1007/s10528-021-10128-w)
1149 [s10528-021-10128-w](https://doi.org/10.1007/s10528-021-10128-w).

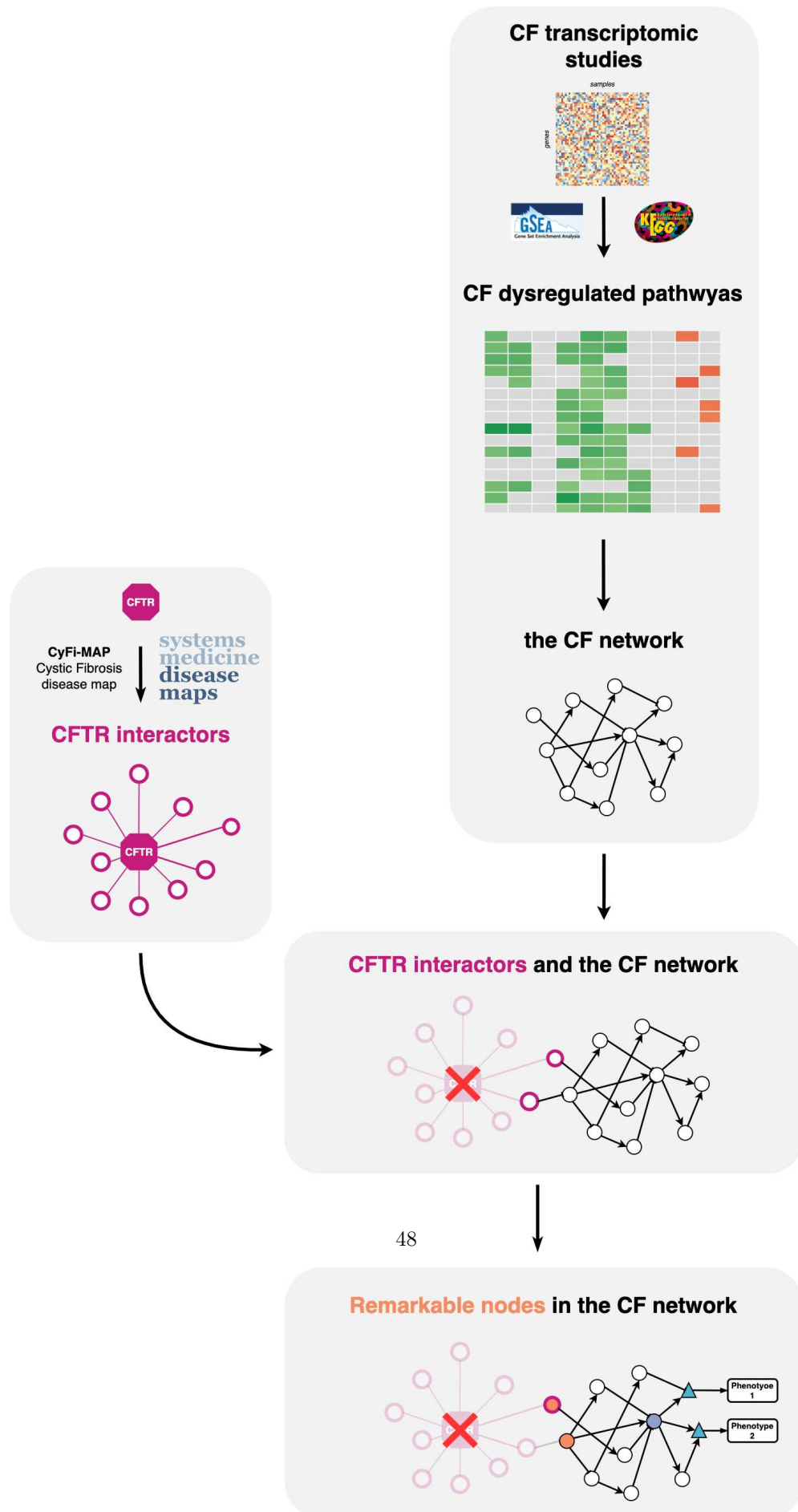


Fig. 1 Global approach followed to build the CF network: A meta-analysis of CF transcriptomic data allowed the identification of dysregulated pathways and the construction of the corresponding CF network. This network comprises known CFTR interactors that can be viewed as source nodes initiating the propagation of dysregulations.

Table 1 List of the 10 datasets considered in the meta-analysis, indicating the number of CF and NCF samples in each study.

Study	Year	Tissue sample	Cell culture	Technology	nb CF, NCF	Dataset	References
Verhaeghe	2007	Tracheal	Cell line	Microarray	3, 3	E-MEXP-980	(Verhaeghe et al., 2007)
Ogilvie (Nasal)	2011	Nasal	Primary culture	Microarray	27, 18	E-MEXP-436	(Ogilvie et al., 2011)
Ogilvie (Bronchial)	2011	Bronchial	Primary culture	Microarray	8, 17	E-MEXP-436	(Ogilvie et al., 2011)
Voisin	2013	Nasal	Primary culture	Microarray	5, 5	GSE40445	(Clarke et al., 2013)
Clarke	2014	Bronchial	Cell line	Microarray	3, 3	GSE39843	(Voisin et al., 2014)
Balloy	2015	Bronchial	Primary culture	RNA-Seq	4, 3	ERP010372	(Balloy et al., 2015)
Zoso	2019	Bronchial	Primary culture	RNA-Seq	7, 6	GSE127696	(Zoso et al., 2019)
Ling	2020	Bronchial	Primary culture	RNA-Seq	7, 5	GSE138167	(Ling et al., 2020)
Saint-Criq (UNC)	2020	Bronchial	Primary culture	RNA-Seq	3, 2	GSE154905	(Saint-Criq et al., 2020)
Saint-Criq (SC)	2020	Bronchial	Primary culture	RNA-Seq	3, 3	GSE154905	(Saint-Criq et al., 2020)

Table 2 Number of detected genes, CF and NCF samples, tested pathways and dysregulated pathways per study with a corrected p-value < 0.25 and a $|\log_2FC| > 1$ thresholds at the gene scale.

Study	nb detected genes	nb CF, NCF	nb up-regulated pathways	nb down-regulated pathways
Verhaeghe	22880	3, 3	11	2
Ogilvie (Nasal)	19880	27, 18	0	0
Ogilvie (Bronchial)	19880	8, 17	33	0
Voisin	13144	3, 3	10	0
Clarke	14118	5, 5	40	2
Balloy	39430	4, 3	24	2
Zoso	18846	7, 6	2	0
Ling	28138	7, 5	1	6
Saint-Criq (UNC)	39434	3, 2	11	0
Saint-Criq (SC)	39432	3, 3	1	9

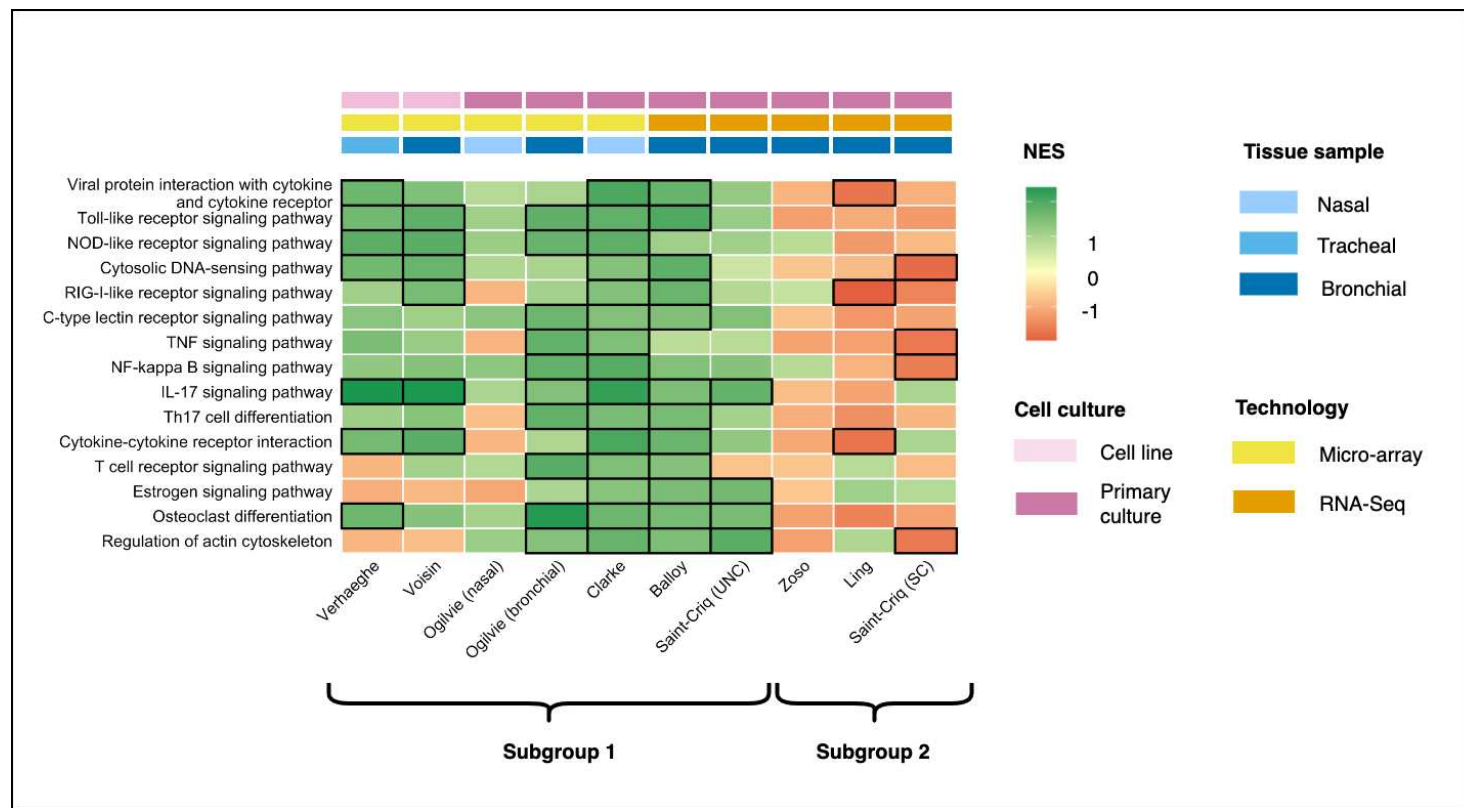


Fig. 2 Heatmap of the GSEA Normalized Enrichment Scores (NES) of the biological pathways differentially expressed in at least 3 studies. The datasets can be clustered in two subgroups based on their NES: Subgroup 1 and Subgroup 2, respectively in agreement and in contradiction with CF physio-pathology. Black boxes around the tiles represent the pathways significantly differentially expressed in the corresponding dataset.

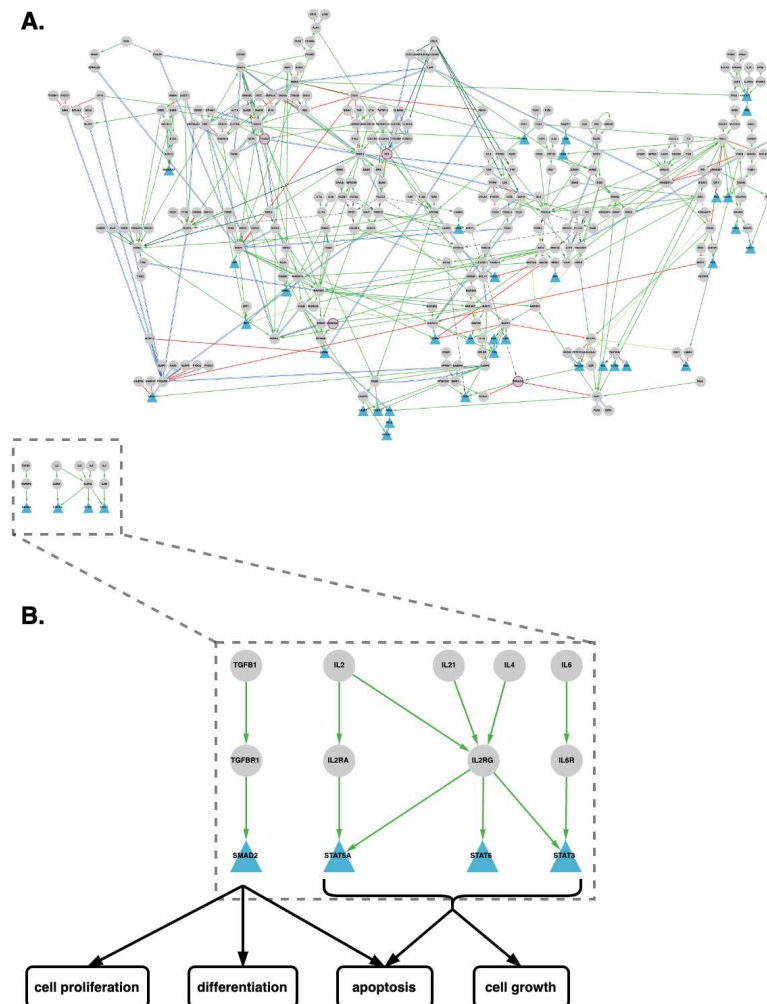


Fig. 3 The CF network(A): The main component comprises 317 nodes connected by 517 interactions and two small unconnected components shown in (B): the two unconnected components correspond to the TGF β and the JAK-STAT signalling pathways. The cellular phenotypes triggered by the sink nodes of the two components are surrounded by black contours.

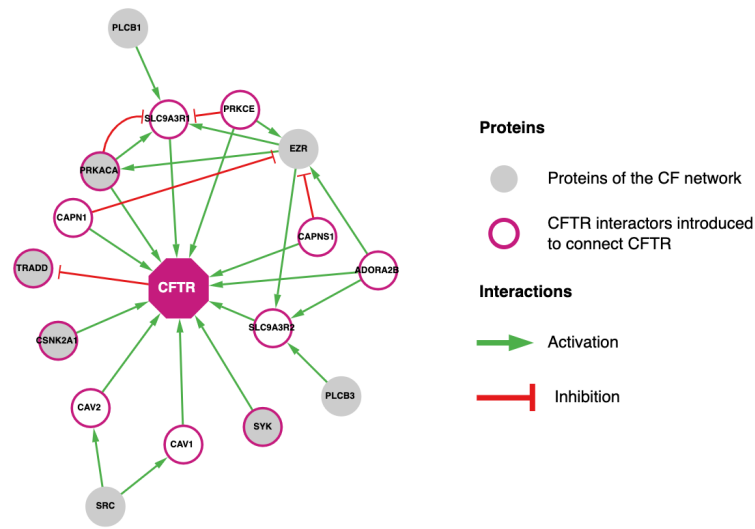


Fig. 4 CFTR interactors in the CF network: Known protein-protein interactions involving CFTR interactors in CFTR PPI.

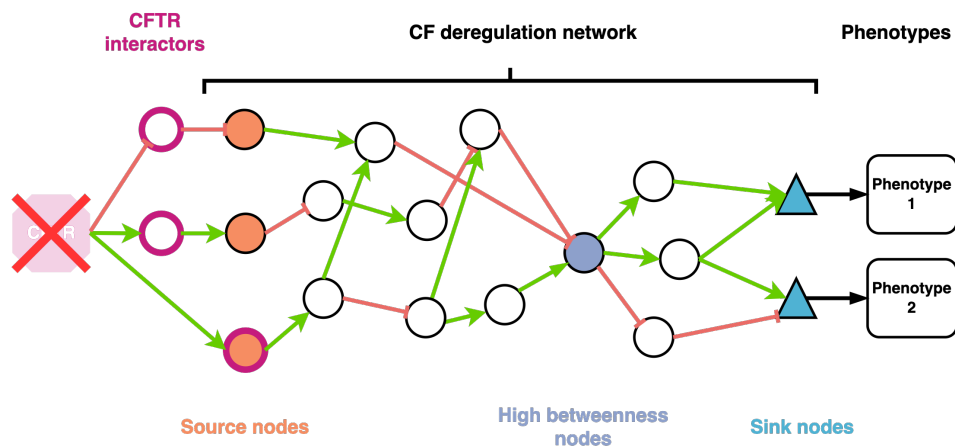


Fig. 5 Illustration of propagation of dysregulation and remarkable nodes in the CF network: the source nodes (orange disks) are CFTR interactors or connected to CFTR interactors via a single intermediate protein (magenta circles). Nodes with high betweenness centrality (purple disks) are proteins through which much information flows within the network. Sink nodes (blue triangles) modulate their corresponding phenotypes.

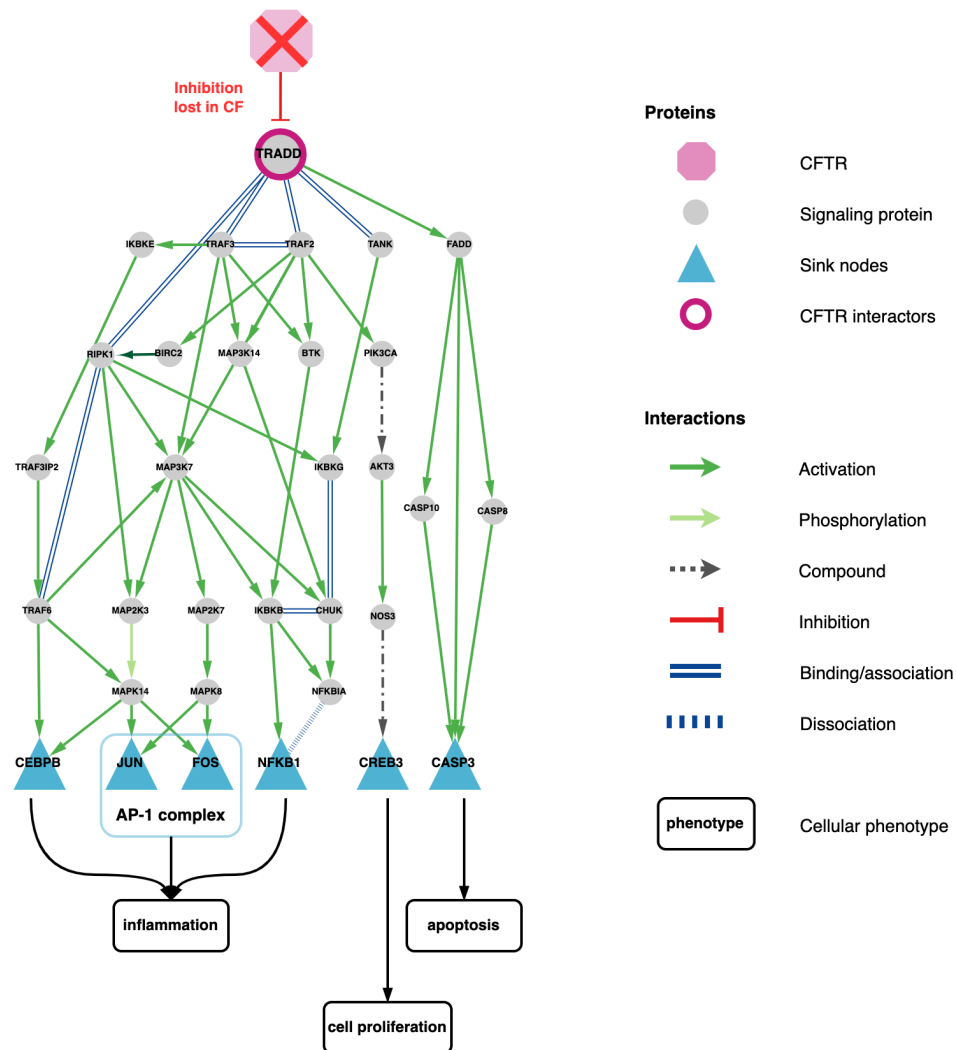


Fig. 6 Extract from the CF network showing the TRADD protein connected to the TNF- α signalling pathway, and to 5 other sink nodes, including FOS and JUN which form the AP-1 transcription factor, downstream of the MAPK cascade. The cellular phenotypes triggered by the sink nodes are surrounded by black contours. Note that TRADD is connected to the 35 sink nodes, but only part of the nodes downstream of TRADD in the network are represented.

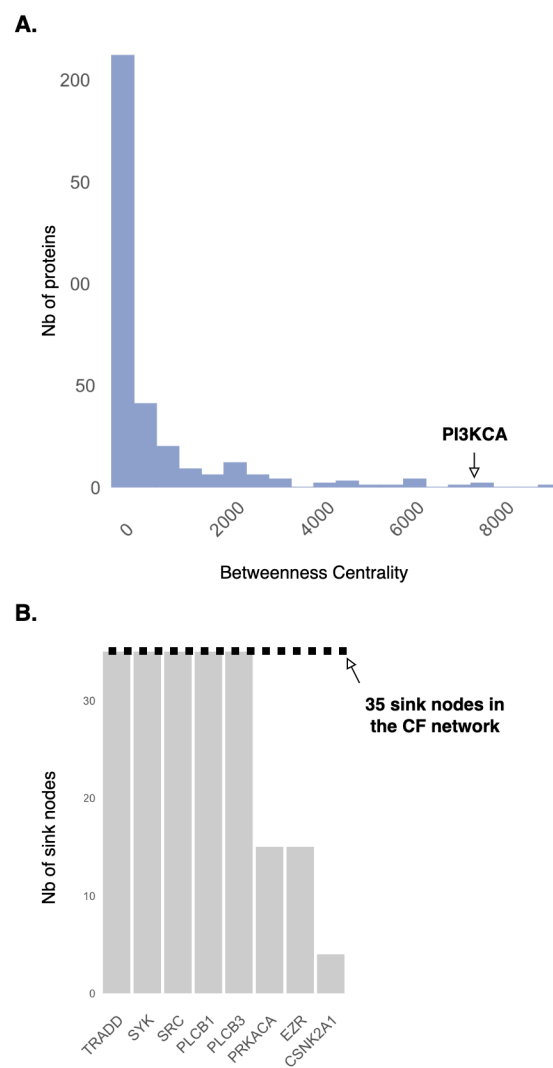


Fig. 7 (A) Histogram of the betweenness centrality measures for all nodes in the CF signalling network; (B) Number of sink nodes to which each of the 8 source nodes are connected.

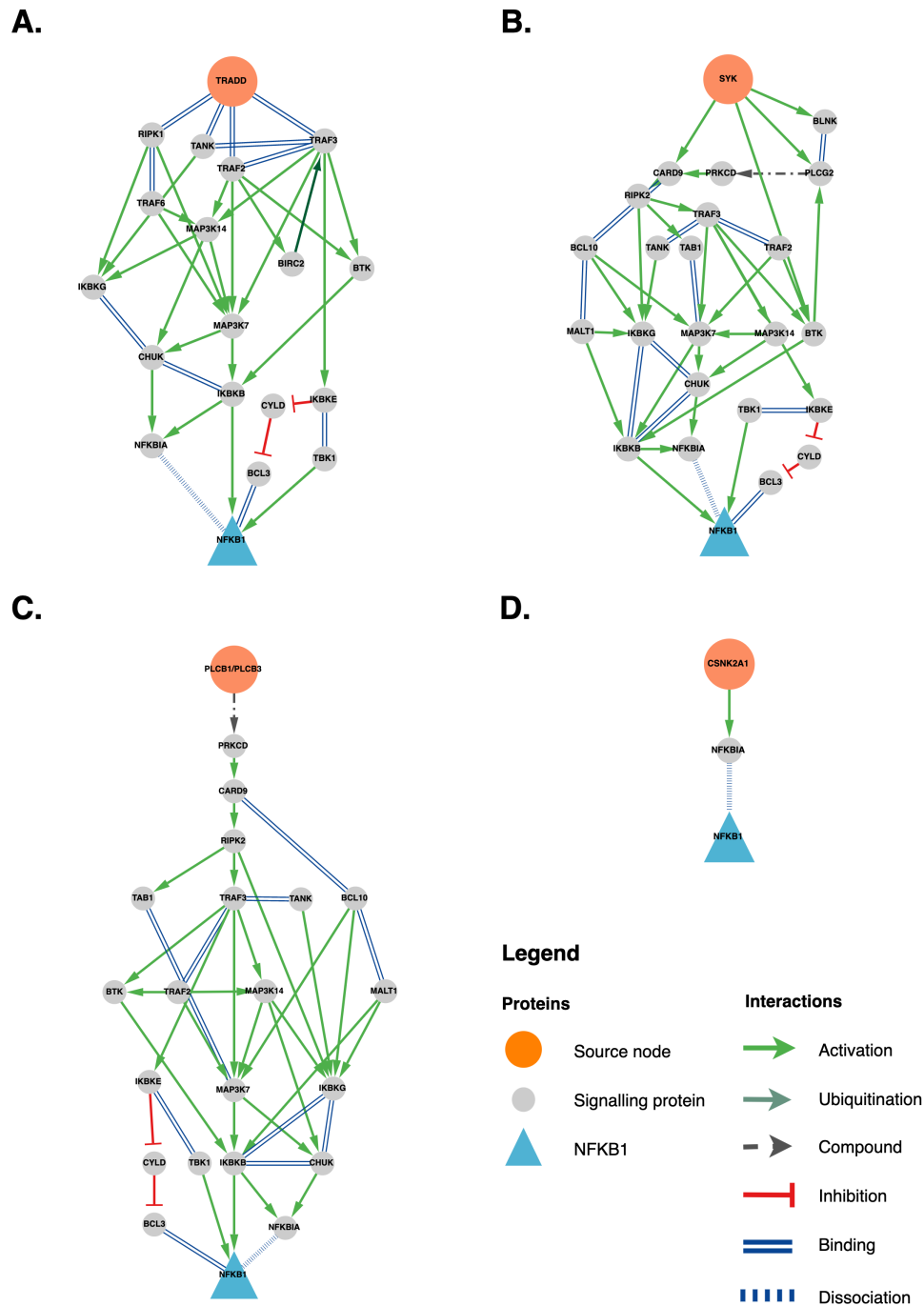


Fig. 8 Subnetworks of the CF network illustrating the connections between the source nodes TRADD, SYK, PLCB1/3, and CSNK2A1 and the sink node NFKB1.