

Interactions between an inflammatory response to infection and protein trafficking pathways favor correction of defective protein trafficking in Cystic Fibrosis

Jerry Mobley Wright^{1,*}, Elisabeth Joseloff², Yuri Nikolsky³, Tatiana Serebriyskaya³, Diana Wetmore⁴

¹Department of Physiology, Johns Hopkins Medical Institutions, 725 N. Wolfe St. Baltimore, MD 21205; ²Cystic Fibrosis Foundation Therapeutics, Inc. Bethesda, MD; ³GeneGo, Encinitas, CA; ⁴Emerald Biostructure, Bainbridge Island, WA; Jerry Mobley Wright- Email: jwright@jhmi.edu; Phone: 443-803-5416; *Corresponding author.

Received September 01, 2010; Accepted October 10, 2010; Published November 27, 2010

Abstract:

One unresolved issue in Cystic Fibrosis research is how functional loss of CFTR, a protein involved in chloride transport, results in chronic lung inflammation. Large scale experiments investigating protein or gene expression changes due to altered trafficking of the most common disease causing CFTR mutation ($\Delta F508$) have produced long lists of changes with no apparent connection to inflammation. Likewise, experiments documenting the effects of inflammation in bronchial epithelial cell lines have yielded no insights into CFTR trafficking. We used MetaMiner CF to combine and analyze results of several CFTR trafficking and epithelial response to infection studies which were on different platforms using different methodologies and had different objectives. The program searches a manually curated database for published experiments linking proteins or genes and displays the interactions in a more easily understood graphic format. Numerous connections were established between genes documented to correct $\Delta F508$ trafficking and a list of genes differentially expressed in bronchial epithelial cells after exposure to bacteria or virus. Of 34 genes documented to correct $\Delta F508$ trafficking, 9 were directly linked by positive expression activation mechanisms to the immune inflammatory response. Looking at interactions among the results as a whole and in detail, it is apparent that an inflammatory response produces numerous changes which favor correct trafficking of $\Delta F508$. One can take a view of the inflammatory process as potentially a corrective mechanism for dysfunctional $\Delta F508$ trafficking. This opens up a new research direction and provides new targets in the search for disease treatments.

Background:

Cystic Fibrosis (CF) is a genetic disorder resulting from mutations in the cystic fibrosis transmembrane conductance regulator (CFTR), a chloride conducting channel that regulates fluid in the gastrointestinal and respiratory tracts [1]. The most common CFTR mutation causing disease is the deletion of a phenylalanine at position 508 ($\Delta F508$) which causes the protein to be retained in the ER and be degraded instead of being trafficked to the apical membrane [2]. One of the major resulting clinical problems in CF is excessive chronic airway inflammation in response to bacterial infection and may occur even before bacterial airway infection is documented and continue after an infection is cleared [3] [4]. How a change in airway epithelial chloride conductance results in activation of an inflammatory response is not well understood.

Since its identification as a specific disease in 1938, CF patient lifetime has increased from a few months to over 30 years [5]. With an increasing number of long term patients, issues with tissue damage due to chronic inflammation and bacterial infection assume a much greater importance. Current research efforts are multi-pronged to address immediate problems of damage due to chronic inflammation and longer term projects to address the basic defect by gene therapy to replace the defective gene or find a means to correctly traffic $\Delta F508$. Additionally, there are biomarker studies

aimed at discovering relationships in disease pathways that can be used to monitor disease status, identify exacerbations before patients become symptomatic and evaluate effects of treatment.

Data integration and information retrieval from large scale studies with multiple independent data sources are significant issues in studying biological systems. Many experiments are independently developed for specific projects with different objectives on a variety of platforms using multiple methodologies. Even output file information may be in incompatible formats. For example protein results may have Swiss-Prot identification and gene microarrays results may be listed with Affymetrix identifications. Historically, the process of connecting information scattered among various experiments was for the researcher to use personal knowledge and experience, exhaustive literature surveys and professional judgment to create a comprehensive picture of the underlying processes. The process is tedious, time consuming, requires expertise in multiple fields and it is difficult to grasp an overall view when mired in the details. With the advent of highly automated testing, gene microarrays, protein arrays and an explosion in the number of scientific papers published annually, the ability of an individual to stay current and integrate multiple sources has been severely challenged.

Automated literature searching suffers from several problems when integrating information sources. One of the larger issues is the naming of genes and proteins. The same gene or protein may be reported under a variety of names or the same name may be used for different genes or proteins at different times [6]. While there has been an effort to standardize naming, each discipline still has a favored name for some of the commonly researched genes or proteins within the field. As a consequence, only manual inspection of the literature will correctly link the results in these publications [6]. Additionally a standardized vocabulary is required so that a computer can find interactions across the database and link effects to pairs of objects that can be used by a computer to construct complex networks of interactions. MetaMiner CF incorporates cystic fibrosis research oriented extensions to the MetaMiner program developed by GeneGo (St Joseph, MI). This is an automated search engine but it is based upon a manually curated database which provides a high degree of reliability in connecting literature to specific genes and proteins plus provides a graphic output which presents a more easily understood overall view of network connectivity.

Methodology:

PubMed and the NCBI GEO databases were searched for publications, reviews and experimental data related to CFTR trafficking or inflammatory responses due to bacterial infection in lungs. Where available, source microarray images were examined for quality.

Data were uploaded into MetaMiner CF where protein and gene identifiers were translated to a standard set of common objects for automated literature searching and association analysis. Identification is done from a manually curated database of standard names and aliases. The networks (directed graphs) were built using the direct interactions algorithm with identified genes or proteins as the seed nodes and protein interactions in the MetaCore database as the edges. The resulting networks are unique for a dataset and are a qualitative analysis tool as there are no weights on different interaction types and kinetic data on signaling interactions is not available. The space and time dimensions are defined by the original experiment on file at NCBI GEO and in this case are not provided. Each data set was first analyzed independently by checking for publications indicating direct interactions between objects within each set; a second analysis then identified direct interactions between sets.

Discussion:

We identified 6 publications with adequate numbers of identified genes or proteins for analysis and the associated lists were loaded into MetaMiner CF [7,8,9,10,11,12]. The interactions among 3 of these sets provided new insights into the disease mechanism.

In the Mayer experiment [12], human bronchial epithelial cell lines were exposed to UV inactivated pathogens. Bacteria, *P. aeruginosa* or *S. aureus*, or the respiratory syncytial virus (RSV) were used. Changes in gene expression profiles 4 hr after exposure were created using the Affymetrix U133 2.0 genome chip. A subset of genes changing in expression in response to both RSV and *P. aeruginosa* challenges were used to create a response model for further analysis. A search for documented direct interactions between members of this gene list resulted in an extensive network of elements directly linked to each other by experimental data (Supplementary Figure 1). Many of the network elements, such as interleukins and chemokines, are well known and well documented in the inflammatory response to infection. This figure was used as a base model of a general inflammatory immune response to infection in bronchial epithelial cells.

The interaction graph of the CFTR ER-interactome [7] was also used as a base connectivity figure because it covered much of the known CFTR ER-traffic pathway interactions and produced a well connected map with 11 of 25 elements having demonstrated direct interactions confirmed by experimental data (Supplementary Figure 2).

A set of 34 genes has been demonstrated to significantly improve $\Delta F508$ trafficking to the cell surface when expressed in a cell line stably expressing $\Delta F508$ [8]; however, the mechanisms of correction are currently unknown. These genes did not produce a well connected map within the group which suggested these are functionally independent, possibly operating through a variety of pathways or operating on a pathway which has not yet been defined. However, six of the genes, HSP70, HSBP1, CHIP, Erdj5, DJ-1 and HSP22, integrated with direct connections into the CFTR ER- trafficking map and further extended the network by providing intermediate links through the immune system response transcription factor STAT1 (Figure 1, Supplementary Figure 3).

Proteins which are secreted into the extracellular space, CCK, CCL2, Calgranulin A, and Endothelin-1, have no direct connections from them to protein trafficking in the ER. However, further analysis indicates a likely functional pathway through known interactions. CCL2 can activate STAT1 via Janus kinases linked to CCR1, CCR2 and CCR5 receptors (Supplementary Figure 4); similarly, CCK can activate both STAT1 and PPAR-gamma. The JAK/STAT connection is a well established cytokine signaling pathway [13] and STAT1 directly influences expression of multiple chaperones involved in CFTR trafficking (Figure 1).

Analysis of trafficking corrector gene set [8] in conjunction with the inflammatory response gene set [12] produced numerous unexpected connections which tied the inflammatory and trafficking responses together (Figure 1). We found that 9 of the 34 genes known to correct $\Delta F508$ trafficking (CCL2, Endothelin-1, Calgranulin A, CCK, AQP3, Galectin-3, Caveolin-2, STAT1, PPAR-gamma) were directly linked by positive expression activation mechanisms to the inflammatory response (Figure 1). In addition, 3 genes which can correct $\Delta F508$ trafficking when over expressed (CCL2, P23 co-chaperone and APQ3) are also upregulated at the mRNA level by the cellular response to infection. The P23 co-chaperone is in a position to influence CFTR maturation due to its interaction with HSP90, a critical step in trafficking of mature CFTR [14]. The principal path of cross talk between the inflammatory response and the corrector/trafficking compartments was through transcription factors, notably STAT1 and NF- κ B (Figure 1).

Mapping connections between elements rather than just comparing similarity of function between lists was a more productive approach in this case. A parallel analysis using Gene Ontology descriptions found no functional overlap between the studies (Supplementary Figure 5).

While all these connections are found in existing literature, Figure 1 is a summary of 818 publications with experimental data, it is highly unlikely that an individual would pull all the disparate information together to arrive at a model linking inflammation and trafficking in this manner. Prior models of inflammation in CF have focused on the origin and maintenance of an inflammatory response as a reaction to misfolded proteins not being degraded properly, an overwhelmed degradation system, presence of unfolded proteins, a reaction to oxidative stress, altered homeostasis of calcium or a variety of other reasons [15]. Our results suggest that the inflammatory response may also favor correct trafficking of misfolded proteins. Rather than turning off the inflammatory response entirely, it may be of therapeutic benefit to selectively activate discrete portions of the response to promote correct trafficking of $\Delta F508$ thereby reducing the overall inflammatory response.

Several diseases with an inflammatory component are proposed to be caused or influenced by protein trafficking including type 2 diabetes, insulin resistance and Alzheimer's disease [16,17]. Protein trafficking problems can arise at multiple points in the process. Insulin resistance and type 2 diabetes are proposed to occur when misfolded proteins overload an early step in the trafficking process thus setting off an inflammatory response [16]. In CF, the vast majority of severe cases involve a single mutation which fails to traffic the CFTR protein from the ER to the Golgi

and has an associated exaggerated inflammatory response; this is the focus of this analysis. Central to controlling an inflammatory response is NF- κ B, a complex structure with multiple subunits [18]. The NF- κ B complex is an integrating sensor through which the input and output signaling can change from cell to cell depending on the subunit composition of the complex and the subunit composition of its upstream regulators. Proposed targets for intervention would be downstream of NF- κ B by activating specific genes known to enhance correct trafficking of mutant CFTR but do not provide positive feedback for inflammatory signaling.

Conclusion:

A systems biology knowledge-based bioinformatics tool used experimental data to develop signaling networks for a bronchial epithelial inflammatory model and CFTR trafficking. Source data were generated on diverse platforms by independent laboratories with different experimental goals. The tool was able to integrate data and identify key pathways and targets in a protein trafficking pathway. NF- κ B and STAT1 were identified as transcription factor hubs linking inflammation and trafficking pathways. The interconnected pathways suggested an unexpected role for the inflammatory response in CF in that inflammatory reactions produce changes which favor correction of Δ F508 trafficking.

The close connection between elements of an inflammatory response and trafficking of CFTR was entirely unexpected. Much of the previous work on the connection between chronic inflammation and CFTR trafficking has focused on understanding the inflammatory triggering mechanism and developing therapies to reduce the inflammatory state [15]. The interactions between elements of the separate data sets provided important information needed to generate new insights at a systems level of organization.

These results assist in understanding the mechanism and effects of chronic inflammation in more detail thus providing new research directions and new hypothesis to test. One question immediately raised is: can specific portions of the inflammatory response be enhanced to further promote correct Δ F508 trafficking without causing additional tissue destruction? Possible future studies include use of RNAi, over expression, cell line models either primary culture or established cell line, and ion flux to test drugs which affect protein targets in inflammatory pathways and determine if they can correct Δ F508 trafficking. These results highlight the complexity of Cystic Fibrosis at the organism level and the value of integrating multiple large data sets generated by independent laboratories.

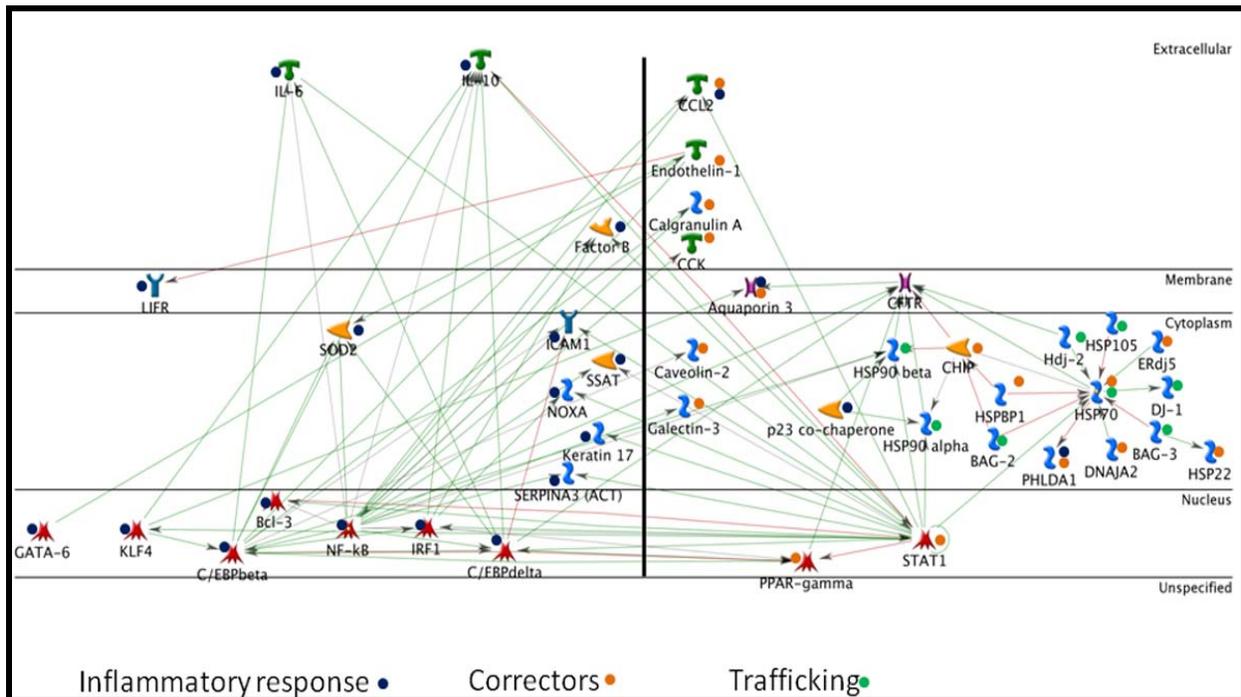


Figure 1: Connectivity map of combined multiple, independent experiments showing connections between the inflammatory response and CFTR trafficking primarily through NF- κ B and STAT1 transcription factors. Genes differentially regulated in bronchial epithelial cells in response to both *Pseudomonas* and respiratory syncytial virus are indicated by black dots. Proteins associated with CFTR trafficking are indicated by green dots. Genes demonstrated to correctly traffic Δ F508 when expressed in a cell line are indicated by orange dots. Separation of the inflammatory response and correction/trafficking compartments is denoted by the vertical black bar. Lines between objects indicate there are published experiments showing direct interactions between the two objects; arrows indicate direction of action. Red lines indicate inhibitory effects; green lines indicate stimulatory effects while grey lines indicate an association, such as binding, but no described effect.

References:

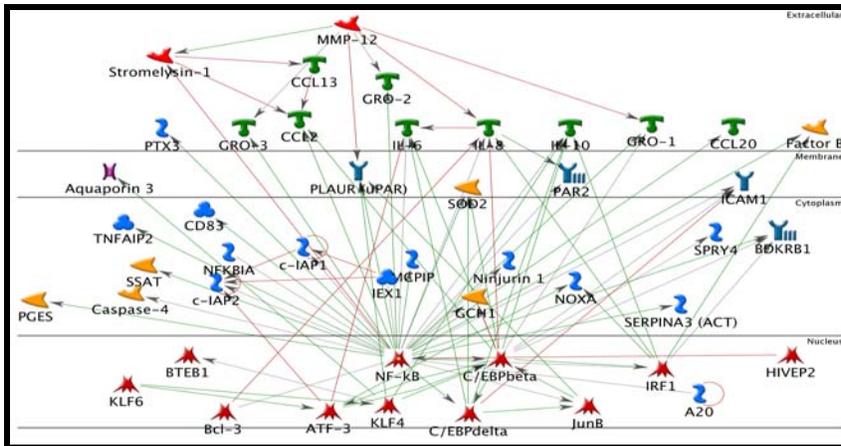
- [1] MP Anderson *et al. Science* **251**: 679 (1991) [PMID:1704151]
[2] MS Gelman *et al. J Biol Chem* **277**: 11709 (2002) [PMID:11812794]
[3] TZ Khan *et al. Am J Respir Crit Care Med* **151**: 1075 (1995) [PMID:7697234]
[4] S Rao & J Grigg *Archives of Disease in Childhood* **91**: 786 (2006) [PMID:16923862]
[5] PB Davis. *Am.J Respir Crit Care Med* **173**: 475 (2006) [PMID:16126935]
[6] D Howe *et al. Nature* **455**: 47 (2008) [PMID:18769432]
[7] X Wang *et al. Cell* **127**: 803 (2006) [PMID:17110338]
[8] AM Trzcinska-Daneluti *et al. Mol Cell Proteomics* **8**: 780 (2009) [PMID:19088066]
[9] OV Singh *et al. J Proteome Res* **5**: 562 (2006) [PMID:16512671]
[10] M Ollero *et al. Proteomics* **6**: 4084 (2006) [PMID:16791827]
[11] M Roxo-Rosa *et al. Proteomics* **6**: 2314 (2006) [PMID:16518875]
[12] AK Mayer *et al. The Journal of Immunology* **178**: 3134 (2007) [PMID:17312161]
[13] DS Aaronson & CM Horvath. *Science* **296**: 1653 (2002) [PMID:12040185]
[14] MA Loo *et al. EMBO J* **17**: 6879 (1998) [PMID:9843494]
[15] M Rottner *et al. Respir Res* **10**: 23 (2009) [PMID:19284656]
[16] GS Hotamisligil. *Cell* **140**: 900 (2010) [PMID:20303879]
[17] HW Querfurth & FM LaFerla. *N.Engl. J Med* **362**: 329 (2010) [PMID:20107219]
[18] PA Baeuerle & T Henkel. *Annu. Rev. Immunol* **12**: 141 (1994) [PMID:8011280]
[19] AK Mayer *et al. The Journal of Immunology* **178**: 3134 (2007) [PM:17312161]
[20] X Wang *et al. Cell* **127**: 803 (2006) [PM:17110338]
[21] AM Trzcinska-Daneluti *et al. Mol Cell Proteomics* **8**: 780 (2009) [PM:19088066]

Edited by VS Mathura

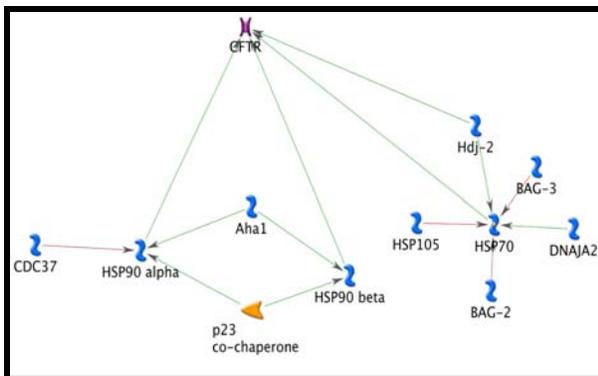
Citation: Wright *et al. Bioinformatics* 5(6): 228-233 (2010)
distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

License statement: This is an open-access article, which permits unrestricted use,

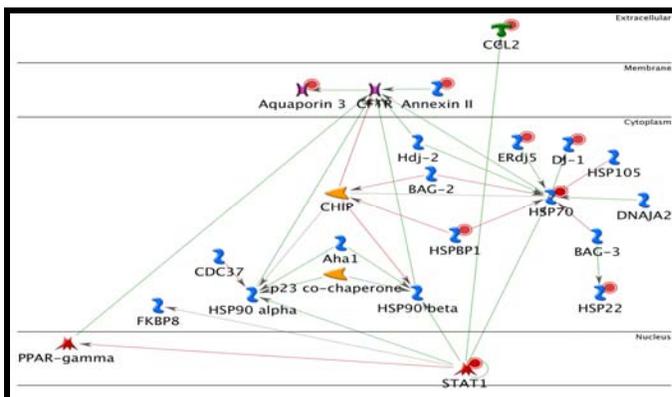
Supplementary material:



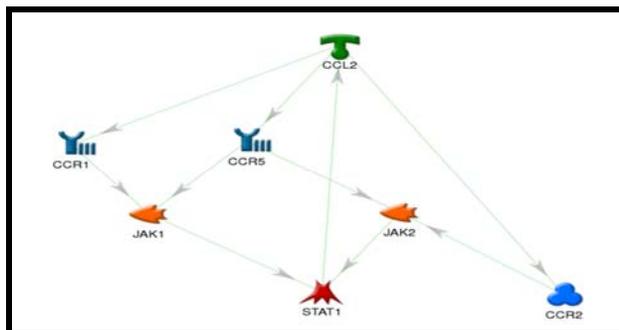
Supplementary Figure 1: An interconnected network based on curated literature-based associations of the 93 genes identified as differentially expressed in response to infection of bronchial epithelial cells [19]. 46 of the differentially expressed genes have established connectivity in existing literature. The network represents the multi-step pathways from membrane receptors via signal transduction cascades to transcriptional factors.



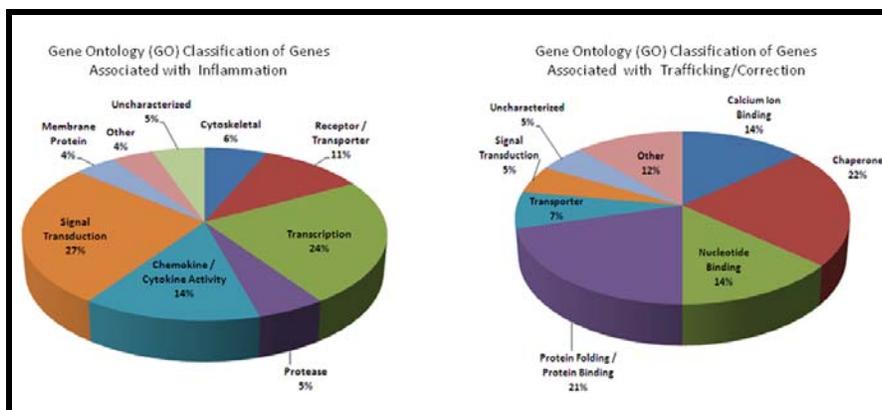
Supplementary Figure 2: Literature based connectivity map of proteins associated with CFTR trafficking [20]. 11 of the 25 proteins of the ER-CFTR interactome have documented direct interactions with each other.



Supplementary Figure 3: Interactions of known DF508 trafficking correctors [21] combined with proteins in the CFTR-ER maturation and trafficking pathway [20]. Red dots indicate genes demonstrated to correct DF508 trafficking when expressed in cell lines. CCL2 is a secreted protein downstream of STAT1 and has no direct connection into the trafficking pathway.



Supplementary Figure 4: Pathway linking CCL2, a secreted extracellular protein known to correct DF508 trafficking, to STAT1 which is a transcription factor controlling expression of genes in the ER-CFTR trafficking interactome. CCL2 activates CCR1 and CCR5 transmembrane receptors which then activate the JAK/STAT pathway.



Supplementary Figure 5: Targets identified from the publications were divided into two subsets: inflammation [19] and trafficking/correction [20,21]. Ninety-three targets were included in the infection/inflammation model and 59 targets were included in the trafficking/corrector subsets. There was little overlap between the sets based on GO molecular function classifications. [Fig 1-4]GeneGo Network Legend: Lines between objects indicate there are published experiments showing a direct connection between the two objects; arrows indicate direction of action. Red lines indicate inhibitory effects; green lines indicate stimulatory effects while grey lines indicate an association, such as binding, but no described effect.