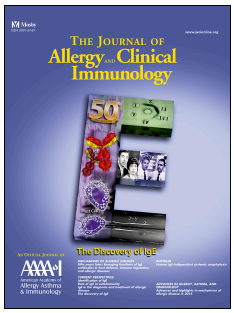
Accepted Manuscript



Differential Connectivity of Gene Regulatory Networks Distinguishes Corticosteroid Response in Asthma

Weiliang Qiu, Ph.D, Feng Guo, Ph.D, Kimberly Glass, Ph.D, Guo Cheng Yuan, Ph.D., John Quackenbush, Ph.D., Xiaobo Zhou, Ph.D., Kelan G. Tantisira, M.D.

PII: S0091-6749(17)31109-0

DOI: 10.1016/j.jaci.2017.05.052

Reference: YMAI 12920

To appear in: *Journal of Allergy and Clinical Immunology*

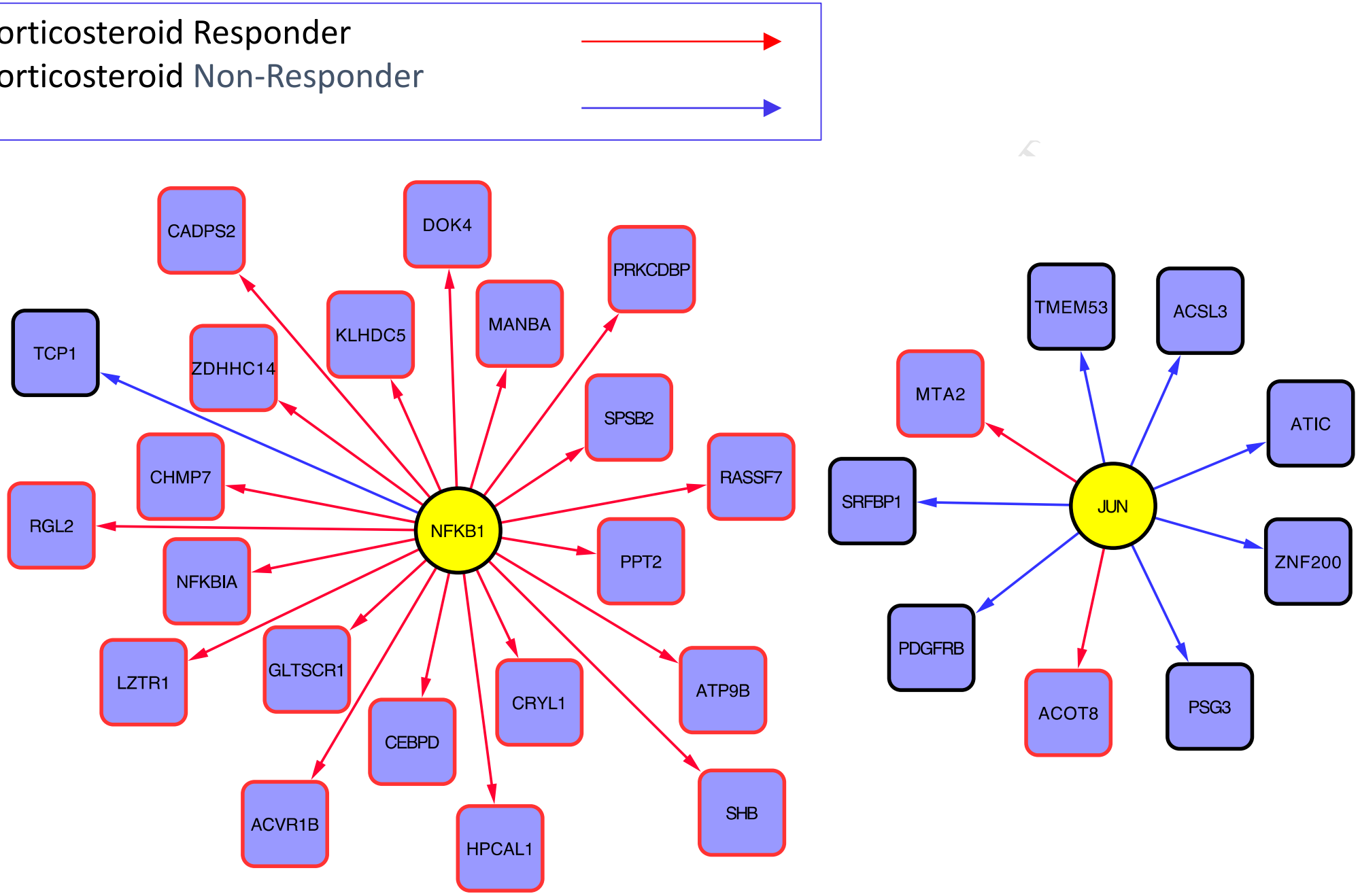
Received Date: 15 March 2016

Revised Date: 2 April 2017

Accepted Date: 3 May 2017

Please cite this article as: Qiu W, Guo F, Glass K, Guo Cheng Yuan Quackenbush J, Zhou X, Tantisira KG, Differential Connectivity of Gene Regulatory Networks Distinguishes Corticosteroid Response in Asthma, *Journal of Allergy and Clinical Immunology* (2017), doi: 10.1016/j.jaci.2017.05.052.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.


Corticosteroid Responder
Corticosteroid Non-Responder
CADPS2
DOK4
PRKCDBP
TCP1
ZDHHC14
KLHDC5
MANBA
SPSB2
CHMP7
RASSF7
RGL2
NFKB1
PPT2
NFKBIA
LZTR1
GLTSCR1
CRYL1
ATP9B
CEBPD
ACVR1B
HPCAL1
SHB
TMEM53
ACSL3
MTA2
ATIC
SRFBP1
JUN
ZNF200
PDGFRB
ACOT8
PSG3

ACCEPTED MANUSCRIPT

Differential Connectivity of Gene Regulatory Networks Distinguishes Corticosteroid Response in Asthma

Weiliang Qiu[1] Ph.D*., Feng Guo[1] Ph.D*., Kimberly Glass[1] Ph.D. Guo Cheng Yuan[2,3] Ph.D., John Quackenbush[2,3] Ph.D., Xiaobo Zhou[1] Ph.D., Kelan G. Tantisira[1,4] M.D.

[1] Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA.
[2] Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA;
[3] Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA;
[4] Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA.
*, These authors contributed equally to the work.

Conflicts of interest: none

The authors are thankful for the support from NIH R01 HL092197, K23 HG003983, U01 HL065899, R01 HL111759, P01 HL105339, R01 HL127200, and R33 HL120794

ACCEPTED MANUSCRIPT

Abstract

Background: Variations in drug response between individuals have prevented us from achieving high drug efficacy in treating many complex diseases, including asthma. Genetics plays an important role in accounting for such inter-individual variations in drug response. However, systematic approaches for addressing how genetic factors and their regulators determine variations in drug response in asthma treatment are lacking.

Methods: We used PANDA (Passing Attributes between Networks for Data Assimilations) to construct the gene regulatory networks associated with good responders and poor responders to inhaled corticosteroids based on a subset of 145 Caucasian asthmatic children who participated in the Childhood Asthma Management Cohort (CAMP). PANDA utilizes gene expression profiles and published relationships among genes, transcription factors (TFs), and proteins to construct the directed networks of TFs and genes. We assessed the differential connectivity between the gene regulatory network of good responders vs. that of poor responders.

Results: When compared to poor responders, the network of good responders has differential connectivity and distinct ontologies (e.g., pro-apoptosis enriched in network of good responders and anti-apoptosis enriched in network of poor responders). Many of the key hubs identified in conjunction with clinical response are also cellular response hubs. Functional validation demonstrated abrogation of differences in corticosteroid treated cell viability following siRNA knockdown of two TFs and differential downstream expression between good-responders and poor-responders.

Conclusions: We have identified and validated multiple transcription factors influencing asthma treatment response. Our results show that differential connectivity analysis can provide new insights into the heterogeneity of drug treatment effects.

Keywords: pharmacogenomics; gene expression; inhaled corticosteroids; apoptosis; system biology

Key Messages: Transcription factors showing differential connectivity between network of TFs and their targeting genes for ICS-good-responders and that for poor-responders have potential to characterize the response to corticosteroid treatment.

ACCEPTED MANUSCRIPT

Capsule Summary: Almost half of asthmatic patients do not respond well to standard treatment. We proposed a network approach to identify key transcription factors and their target genes that may determine differential drug response in asthmatic patients.

ACCEPTED MANUSCRIPT

Introduction

Asthma is the most common chronic airway disease among children and young adults(1) characterized by airflow obstruction in the small airways of affected individuals. The prevalence of asthma and its mortality have almost doubled in the last 20 years, imposing an increasing financial burden to medical care system. Despite availability of many standard treatments, including β2-agonists, corticosteroids, and leukotriene antagonists(2-4), to control asthmatic symptoms, almost half of asthmatic patients do not see improvement in symptoms. Such variation in drug response are attributed to many factors (5, 6), including genetics (7-10). Pharmacogenomics studies assess how genetic and genomic variation affects an individual's response to drug treatment(11-15).

Corticosteroids are the backbone of asthma therapies. They reduce inflammation through both gene activation and suppression. The molecular mechanisms of action of corticosteroids have been clearly described(16). Failure of any of the steps along the route where corticosteroids pass from the cell membrane to their target genes may result in non-responsiveness. For example, abnormalities in glucocorticoid receptor number, glucocorticoid receptor binding, or abnormalities in glucocorticoid-glucocorticoid receptor complex binding to DNA may result in poor response to corticosteroid therapy(17-20). Despite these insights, the molecular mechanisms underlying a given asthma patient's poor-responsiveness to corticosteroid medications remain unclear.

Gene differential expression, single nucleotide polymorphism, and expression quantitative trait loci (eQTL) analyses have identified multiple genes associated with asthma drug responsiveness(21-24). However, these analyses typically focus on one gene at a time. Gene expression itself is regulated by several mechanisms, such as transcription factors, microRNAs, and DNA methylation. For complex diseases such as asthma, genes and their regulators are believed to work together; network approaches investigating asthmatic drug response should include consideration of both the genes and their upstream regulators.

It is well known that (1) Transcription factors (TFs) play key roles in regulating gene expression; (2) TFs usually work together to co-regulate gene expression; and (3) genes with similar functions tend to co-express. Using this information, we have previously developed a message-passing model, called PANDA (Passing Attributes between Networks for Data Assimilations)(25), which assimilates information from multiple, complementary data-types in order to reverse engineer a regulatory network. In this study, we apply PANDA to a set of immortalized B-cells, which were derived from inhaled corticosteroid treated asthmatics who were part of a large clinical trial. We hypothesize that assessing the regulatory networks of poor- and good clinical responders using PANDA could uncover the molecular mechanisms by which drug response to inhaled corticosteroids (ICS) of each patient is determined.

Methods

Asthma Cohort. CAMP (Childhood Asthma Management Program) was a multicenter, randomized, double-masked clinical trial designed to determine the long-term effects of three inhaled treatments for mild to moderate childhood asthma: budesonide (a

glucocorticoid used daily) and albuterol (a short-acting beta-agonist bronchodilator used as needed); nedocromil (a non-steroid anti-inflammatory agent used daily) and albuterol; and placebo and albuterol(26). The primary outcome measure was post-bronchodilator forced expiratory volume in one-second percent of predicted (FEV_1%), observed over 4- to 6-year period. For this project, we selected 47 good-responders and 48 poor-responders from the 145 Caucasian CAMP subjects with available immortalized B-cells (LCLs) gene expression(15) . The definitions of good-responders and poor-responders were based on the change in FEV_1% between baseline and 2-month follow-up. We used the first tertile (Q1=1.10%) and the third tertile (Q3=9.78%) of FEV_1 change calculated based on the 145 subjects to partition the 145 subjects to 3 groups. The 48 subjects with FEV_1 change <Q1 were defined as poor-responders; the 47 subjects with FEV_1 change > Q3 were defined as good-responders.

LCL Microarray Experiment.

As previously described(15) immortalized B-cell lines (LCL) derived from 145 asthmatic subjects from the CAMP clinical trial(26, 27) were cultured in RPMI 1640 medium and treated with dexamethasone (10^{-6} M) or with sham (ethanol) for 6 hours(22, 28, 29). After treatment, total RNAs were extracted and applied for microarray profiling. Gene expression levels of 22,184 gene probes for each sample were measured by using Illumina HumanRef-8 V2 chip (Illumina, San Diego, CA).

Gene expression QC.

The gene expression data contained 201 arrays for dexamethasone-treated LCLs and 193 arrays for sham-treated LCLs. Approximately 10% of subjects had replicate arrays. We first did data quality check for the 2 treatment types of arrays separately. We then pooled the paired samples together and did log2 transformation and quantile normalization.

After data quality checking, 20,917 gene probes in 17,193 genes for 145 pairs of arrays were kept. The log2 difference of expression levels between dexamethasone-treated cell lines and sham-treated cell lines was used to measure the effect of drug treatment on the gene expression. Furtherdetails about data quality control can be found in the Supplementary Documents.

PANDA algorithm.

PANDA (Passing Attributes between Networks for Data Assimilation) is a message-passing model to construct directed networks between TFs and genes using multiple sources of genomic information to predict regulatory relationships(25). The nodes in a PANDA network are TFs or genes. The directed edges extend from TFs to genes. Each edge has a weight value indicating the probability that a TF regulates a gene.

To seed the PANDA algorithm we used a mapping between TF motifs and target genes. Descriptions of the creation of this mapping can be found in Glass et al. (2014 and 2015)(30, 31). This mapping includes 255,051 pairs of (TF, gene) and 13,979 unique genes.

There are 13,191 genes shown in both our gene expression data and the mapping file. These 13,191 genes correspond to 240,939 pairs of (TF, gene) and correspond to 16,368 gene probes in our gene expression data.

Statistical Analysis.

We first evaluated the effect of glucocorticoid on gene expression by comparing the PANDA network for the 145 dexamethasone-treated cell lines with the network for their corresponding controls, 145 sham-treated cell lines. Specifically, we obtained a PANDA network for dexamethasone-treated cell lines and a PANDA network for sham-treated cell lines, respectively. In each network, edges connect TFs and their targeted genes. Due to the differences of gene expression between dexamethasone-treated cell lines and sham-treated cell lines, the edges and edge weights between the two networks are not the same. We call a TF as differentially connected if the TF connects to different sets of genes between the two PANDA networks. Differential connectivity between the 2 PANDA networks (dexamethasone vs. sham) was revealed. We then compared the PANDA networks between the 47 good-responders and 48 poor-responders, focusing on the differential connectivity between the log2-difference in expression response (log2(dexamethasone) – log2(sham)) as the expression metric for a given individual.

Denote $d_i = w_{i,resp} - w_{i,nonresp}$, where $w_{i,resp}$ is the edge weight for the i-th pair of (TF, gene) for the good responders and $w_{i,nonresp}$ is the edge weight for the corresponding pair for the poor responders. We constructed the good-responder network of TFs and their targeted genes by adding edges to (TF, gene) pairs corresponding to the largest 10000 d_i. We then constructed the poor-responder network of TFs and their targeted genes by adding edges to (TF, gene) pairs corresponding to the smallest 10000 d_i. Denote S_g as the set of TFs in the good-responder network. Denote S_p as the set of TFs in the poor-responder network. For the common TFs in both S_g and S_p, we calculated the number of edges that a TF has for good-responder network and poor-responder network, separately. Denote them as nEdge(g) and nEdge(p), respectively. We then calculated the difference of the edges nDiff=nEdge(g) – nEdge(p) and ratio of edges nRatio=nEdge(g)/nEdge(p) for each TF. The difference gives an absolute magnitude, while the ratio provides a gene specific difference in the magnitude of differential regulation for a TF. For differentially connected TFs (i.e. TFs in both good-responder network and poor-responder network), we tested if they are differentially expressed between good responders and poor responders using two sample t test.

To assess the statistical significance of the ratio of edges, we performed a permutation analysis. Specifically, we performed 1000 randomizations wherein we randomly divided LCLs into two groups and generated two corresponding networks. Denote Sg(r) and Sp(r) as the set of TFs in the top 10000 pairs of (TF, gene) in these two “random” networks. As above, we identified the TFs in Sg(r) and Sp(r) , calculated the number of edges for those TFs in Sg(r) and Sp(r), denoted as nEdge(gr) and nEdge(pr), respectively, and determined the ratio of edges, nEdge(gr)/nEdge(pr), for each TF. For each TF with a ratio greater than 1, we counted the total number of permutations where nEdge(gr)/nEdge(pr)>nEdge(g)/nEdge(p); for TFs with a ratio less than 1, we counted the

total number of permutations where nEdge(gr)/nEdge(pr)<nEdge(g)/nEdge(p). We then divided by the number of permutations in which that TF appears in the top (TF,gene) pairs to estimate how many times "by chance" one would expect to find a ratio more extreme that what was observed.

Pathway Enrichment Analysis.

We obtained for each network the set of genes that were regulated by the TFs with differential connectivity (i.e., TFs with large absolute difference of edge weights between the network of good-responders and that of poor-responders). We then compared the enriched pathways of the 2 sets of genes by using the functional annotation tool: the Database for Annotation, Visualization and Integrated Discovery (DAVID)(32, 33). We expected that the 2 sets of enriched pathways would be different.

Gene silencing of TFs and dexamethasone treatment in LCLs

The TFs having most differential regulation may determine if a patient is a good-responder of ICS treatment or not. We chose to validate two key TFs (*NFKB1* and *JUN*) with a different ratio of edges in good vs. poor responders using a knock down experiment. These 2 TFs were at polar ends of the ratio of the numbers of edges that a TF has in good-responder network to those in poor-responder network.

We hypothesized that knocking down the *NFKB1* paralog (*RELA*) or *JUN* would result in modulation of the difference between poor-responders and good-responders. To validate this hypothesis, we cultured LCLs from 7 good responders and 7 poor responders in RPMI 1640 medium (Life Technologies) supplemented with 15% fetal bovine serum (FBS). One siRNA with best knock down efficiency in three individual siRNAs (Life Technologies) targeting *JUN* or *RELA* was chosen for the gene silencing experiment. Gene targeting siRNA and negative contol siRNA were transfected into LCL cell lines using 4D-Nucleofector X kit (Lonza) with Nucleofector Program (EC-117). Optimal amount of siRNA (20-100 pmol siRNA / 10^6 cells) were trasnsfection in each reaction. Twenty-four hours after transfection, cells were seeded into serum free RPMI 1640 medium to starve for 24 hours followed by dexamethasone or vehicle control treatment (1 μM) for 2 hours. Total RNA was extracted from 7 good responder cell lines and 7 poor responder cell lines with or without dexamethasone treatment. By RT-PCR, we measured expression levels of 4 down-stream genes (*CEBPD, MANBA, PPT2,* and *TCP1*) for *NFKB1* silencing, 5 down-stream genes (*ACOT8, ACSL3, MTA2, PDGFR8,* and *TMEM53*) of *JUN*, and one house-keeping gene (*GAPDH*). *RELA* and *JUN* were measured for detecting the knock down efficiency. We also measured expression of *NR3C1* gene that encodes the glucocorticoid receptor to determine whether there is endogenous expression difference of NR3C1 among good and poor responders. We generated histograms of the knock-down efficiencies of *RELA* and *JUN* across subjects. We also performed western blotting to confirm TFs knockdown efficiency in LCL lines. The gene expression level of a gene in PCR analysis is inversely proportional to CT level, where CT stands for cycles to reach to threshold. For a given gene, let delta CT = CT.gene – CT.*GAPDH*, where CT.gene is the CT level of the gene and CT.*GAPDH* is the CT level of *GAPDH*. For each down-stream gene, we performed general linear model analysis to test if delta CT levels for good-responders are different from those for poor-

responders after dexamethasone treatment, adjusted for knock-down status (with TFs siRNA knock-down versus control siRNA knock-down)..

Noting the key difference in apoptosis in the ontology analysis (**Figure 3** and **Supplemental Table 2**), that is, pro-apoptosis enriched in network of good responders and anti-apoptosis enriched in network of poor responders, we also measured cell viability in LCL lines after dexamethasone treatment. The cells were transfected by two different *JUN* siRNA (*JUN* siRNA-1 and *JUN* siRNA-2), RELA siRNA and control siRNA. After starving for 24 hours, each siRNA transfected cells were split into 8 wells and each 4 wells were treated with 0 μM (ethanol, solution vehicle) or 50 μM dexamethasone for 24 hours. The cell viability was detected by alamarBlue assay kit (ThermoFisher DAL1100). The results were normalized to relative ethanol treated cells.

Results

Table 1 shows the population characteristics for the 145 subjects, the 48 poor-responders, and the 47 good-responders, separately. There are no significant differences between good-responders and poor-responders for age, gender, or baseline FEV_1 as a percent of predicted (FEV_1%). The two groups have significant difference between ΔFEV1 (the change in FEV_1% between baseline and 2-month follow-up) based on the definition of the two groups.

The networks of (TF, gene) for dexamethasone-treated LCLs and for sham-treated LCLs are shown in **Figures 1A and 1B**, respectively. The complete set of TFs and their targeted genes are shown in **Supplemental Table 1**. The differential connectivity between the two networks indicates the effect of dexamethasone treatment on the transcription factor regulation of genes.

We next evaluate the differential connectivity between the network of good-responders and that of poor responders. It would be difficult to visualize differential connectivity if we used all top 10000 (TF, gene) pairs. Hence, we first illustrate the differential connectivity by using only the top 50 pairs of (TF, gene) that had the largest absolute differences of edge weights between the 2 networks (**Figure 2**). In **Figure 2**, the red edges are from the network of good-responders and the blue edges are from the network of poor-responders. **Figure 2** indicates that the 2 networks demonstrate extensive difference in connectivity. Interestingly, seven out of the nine hot spots in **Supplemental Figure 1** (networks of dexamethasone treatments) are also in **Figure 2** (networks of responsiveness), suggesting that key regulators of overall corticosteroid response are also important in regulating the clinical extremes of response. The target genes in good responders are listed in **Supplemental Table 3** and the targets in poor responders are listed in **Supplemental Table 4**.

The set S_g (for good responders) contains 32 TFs and the set S_p (for poor responders) contains 35 TFs. There are 31 TFs appearing in both S_g and S_p, including *JUN* and *NFKB1* (**Table 2**). *NFKB1* has the highest ratio (nEdge(g)/nEdge(p)=20, p=0.039 by permutation analysis) and *JUN* has the lowest ratio (nEdge(g)/nEdge(p)=0.29, p=0.019

by permutation analysis). Among the 31 unique TFs, 6 TFs have permutation p-value < 0.05. For each of the 105 TFs in this study, we performed two-sample t-test to test if a TF is differentially expressed between dexamethasone-treated LCLs and sham-treated LCLs. Fifty-one of the 105 TFs have two-sample t-test p-value < 0.05 (**Supplemental Table 5**). Sixteen of the 31 common TFs are differentially expressed between dexamethasone-treated LCLs and sham-treated LCLs (**Table 2**). Of note, neither *JUN* nor *NFKB1* were differentially expressed between dexamethasone-treated LCLs and sham-treated LCLs. We obtained the QQplot of –log10 (t-test p-value) between non-differentially connected TFs and differentially connected TFs (**Supplemental Figure 3**), which showed no significant difference between the 2 sets of TFs in terms two-sample t-test p-values. We also performed Fisher's exact test to assess if the differential expression of TFs (TFs with two-sample t-test p-value<0.05 versus TFs with p-value≥0.05) is associated with differential connection (TFs that are in both good-responder network and poor-responder network versus TFs that are not in both networks). Sixteen TFs are both differentially expressed (DE) and differentially connected (DC); 39 are neither DE nor DC; 35 TFs are DE, but not DC; and 15 TFs are DC, but not DE. The p-value for the Fisher's exact test is 0.8308. Both parallel boxplots and Fisher's exact test showed no association between differential expression and differential connectivity of TFs in two networks of good responders and poor responders.

DAVID functional annotation analysis showed that (1) the 164 genes that are only in S_g (for good responders) were enriched in 67 biological processes, including immune response and pro-apoptosis; and (2) the 225 genes that are only in S_p (for poor responders) were enriched in 33 biological processes, including anti-apoptosis. There are 2 overlapping biological processes (DNA metabolic process and DNA repair). **Figure 3** is the heatmap of –log10(p-value) for enriched GO biological processes, where the p-values are for testing if a biological process is enriched or not based on the list of genes in the top 500 pairs of (TF, gene).

We next experimentally assessed the validity of regulatory networks built using PANDA. We hypothesized that these key TFs that differentiate the good responder network from poor responder network contribute to differential effects of dexamethasone on cells. We assessed functional response along two lines: differential targeting of downstream gene expression between good and poor responders and differential cell viability (since apoptosis was a key phenotype differentiating response status in our pathway annotations). Among 31 TFs revealed by PANDA method (**Table** 2), we choose *NFKB1* and *JUN* for validation. Knock down efficiencies of *RELA* and *JUN* across subjects are shown in **Supplemental Figure 2**, which indicates that *RELA* knock-down efficiencies are good (median=80%, range=71% - 85%) and *JUN* knock-down is less efficient (median=32%, range=8% - 48%). Western blotting image about the knock-down efficiency is shown in **Supplemental Figure 5**.

The results of functional validation analysis of differential targeting showed that one of the four down-stream genes (*CEBPD*) of *NFKB1* has statistically lower delta CT level (i.e. higher expression level) in good responders than in poor responders and that one of the five tested down-stream genes (*TMEM53*) of *JUN* has statistically higher delta CT

level (i.e., lower expression level) in good responders than in poor responders, after dexamethasone treatment and adjusting for siRNA knockdown (**Table 3 and Figure 4**). The parallel boxplots of delta CT levels for all genes are shown in **Supplemental Figure 6** and **Supplemental Figure 7**. For the cell viability assays, we compared the mean of the baseline-adjusted cell viability of good-responder with that of poor-responder for each of the 4 cell types (negative control, *RELA* siRNA, *JUN* siRNA -1 and *JUN* siRNA-2) with 50 μM dexamethasone treatment. Supplemental **Figure 4a** shows the histogram of the cell viability adjusted for baseline for the 4 replicates. For negative control, there exists significant difference between good-responders and poor-responders in terms of cell viability. However, after knocking down *RELA* or *JUN*, the difference markedly decreased (Supplemental **Figure 4b**).

Discussion

In this article, we showed that (1) PANDA can be applied to gene expression data generated from immortalized B-cell lines; (2) many TFs hubs obtained from (TF, gene) networks characterizing the corticosteroid treatment effects on gene expression globally also appeared in (TF, gene) networks characterizing the ICS-responsiveness specifically; (3) the network of good responders and that of poor responders have different connectivity and distinct ontologies; and (4) one down-stream gene (*CEBPD*) of *NFKB1* and one down-stream gene (*TMEM53*) of *JUN* are differentially expressed between good responders and poor-responders adjusting for siRNA knock-down, after dexamethasone treatment.

In the analysis of corticosteroid response alone, we identified nine key “hub” TFs, each with at least nine differentially connected edges between treated and untreated (sham) cells. This suggests that these TFs may be transcriptional regulatory “hot spots”. A literature search indicated that there is biologic plausibility for a role in glucocorticoid signaling for at least seven of the nine TFs (**Supplemental Figure 1**)(34-41). For example, *ETS1* is a regulator of human glucocorticoid receptor 1A promoter(42).

The differential connectivity between the network of (TF, gene) we obtained for ICS-good responders and that for ICS-poor responders revealed multiple TFs that may help to explain why some asthmatic children do not respond well to ICS treatment. From a mechanistic perspective, it is assuring that many of these TFs also play a role in the differential connectivity between the network of dexamethasone-treated cells and that of sham-treated cells. Indeed, seven of the nine key “hub” TFs in differentially connected in dexamethasone vs. sham were also present in the analysis of clinical responders (c.f. **Supplemental Figure 1**). These include *GATA2*, *ETS1*, *YY1*, and *NFIC1*, which all have well documented roles in corticosteroid biology(35, 36, 38-40). These findings support the notion that factors innately involved in the global response to corticosteroids may also modulate treatment response differences between subjects. Further studies of these TFs might help find a way to improve the efficiency of current ICS treatment.

A key feature of PANDA is the emphasis on differential connectivity as compared to differential expression of the transcription factors. Therefore, it is not surprising that the

same sets of TFs regulated different sets of downstream genes between the good responders and poor responders. Hence, the two networks had different ontologies. For the good responders, these included regulation of the immune response and metabolic processes (**Figure 3**). However, the most striking of the ontologic differences was that the network within the good clinical responders supported "pro-apoptosis" pathways, whereas the network derived from the poor responders indicated the presence of "anti-apoptosis" pathway regulation. Corticosteroid induced apoptosis is a known key mechanism related to resolution of asthmatic inflammation and helps to differentiate severe vs. non-severe asthma(43-46). Our findings add to this literature by eliciting the differential interactions between key transcription factors and their downstream targets as they modulate corticosteroid induced apoptosis in good and poor responders.

Given this background, we chose to validate key TFs with a different ratio of edges in good vs. poor responders using both a downstream targeting approaches and a cellular apoptosis assay. In this functional validation analysis, we knocked-out two key TFs (*NFKB1* and *JUN*) from our network analysis using siRNA. *NFKB1* (nuclear factor of kappa light polypeptide gene enhancer in B-cells 1) encodes a 105 kD protein which can undergo cotranslational processing by the 26S proteasome to produce a 50 kD protein. Inappropriate activation of *NFKB1* has been associated with a number of inflammatory diseases while persistent inhibition of *NFKB1* leads to inappropriate immune cell development or delayed cell growth. *JUN* (jun proto-oncogene) encodes a protein that is highly similar to the viral protein, and which interacts directly with specific target DNA sequences to regulate gene expression. Molecular network analysis of endometriosis reveals a role for c-Jun-regulated macrophage activation(47). In the functional validation analysis, we identified one downstream target (*CEBPD*) of *NFKB1* and one downstream targets (*TMEM53*) of *JUN* that are statistically differentially expressed following dexamethasone treatment in good vs. poor responders, supporting differential connectivity of the transcription factors leading to different downstream gene targeting, as exemplified by differential expression, between good and poor ICS responders. We also observed that without knocking-out the *NFKB1* paralog (RELA) or *JUN*, the cell viability of good responders is statistically greater from that of poor responders. Importantly, such differences were largely abrogated after depletion of RelA or JUN by siRNA. Overall, this functional data suggest that RelA and JUN are playing important roles that determine differential cellular response to dexamethasone in these two groups of LCLs and, through generalization, to the differences in clinical response underlying the two groups. Thus, a malfunction of *NFKB1* or *JUN* may explain the poor-responsiveness of ICS treatment and that modulation of the connectivity related to these genes may be of therapeutic benefit. Further investigation is warranted.

Our study has a couple of limitations. One limitation of the present study is that the permutation p-values in **Table 2** are modest. We obtained Benjamini-Hochberg (BH) corrected p-values for the 31 TFs that we identified as targeted in the top/bottom edges to control for multiple testing (**Supplemental Table 6**). Although no TFs are significant at a BH-pvalue<0.05, the same six TFs we noted as nominally significant previously

(including *FEV, GATA3, JUN, NFATC2, NFKB1, and SPI1*) were identified at a BH-pvalue<0.15.

Additionally, when we investigated the siRNA mediated changes in expression, we identified only one gene for each of the two TFs (*NFKB1* and *JUN*). While only one downstream gene for each TF was validated, there are plausible biologic and experimental reasons for this. From an experimental perspective, we note that the network was built upon >200 LCL cell lines and we only choose 7 cell lines for validation due to feasibility reasons. Therefore, it is likely the sum total of target genes as inferred by the network analysis as derived from many samples may not be co-regulated by the specific TF in exactly the same fashion as a single cell line under the same condition. Biologically, we also note that while there is usually a one-to-many relationship between a given TF and its downstream targets, there is often also a co-regulation of gene expression by multiple TFs. Therefore it is entirely possible that siRNA knockdown of a TF may not significantly alter the expression of such a gene. Both of these points suggest that the regulation of gene expression remains complex and that functional validation of networks remains imperfect.

To assess if the results of the network analysis is robust to the choice of the number of top edges used to define the networks, we calculated log2 ratio of the number of edges that a TF has in good-/poor- responder network for different numbers of top edges used (**Supplemental Table 7**). We note that the ratios and especially the direction of enrichment are largely stable across different numbers of edges, especially as the number of selected edges increases (leading to more robust estimates). In addition, the higher thresholds (more edges) are likely more reliable in a sense since there is more overall information with greater numbers of edges (even if these edges are not the highest weight edges).

In permutation study, some TFs might be more easily found in the top and bottom edges compared to others. However, detailed examination shows that of the 31 TFs in **Table 2**, 25 appear in both the positive and negative edges in all 1000 permutations and 29 appear in at least 800 permutations. The last column of **Supplemental Table 6** includes the number of permuted pairs of networks for which each TF appears in both the positive and negative set of edges.

While we used immortalized B-cells for this analysis, these cells were directly derived from clinical subjects participating in CAMP. Moreover, we recently described that the glucocorticoid receptor (the primary receptor for ICS), upon stimulation with dexamethasone, functions differently in these cells in good clinical inhaled corticosteroid responders vs. poor responders(48). Using immortalized cell lines allows an analysis to be repeated many times on genetically identical cells, which is desirable and cost-effective for repeatable scientific experiments. One potential limitation of using immortalized cell lines is that immortalization might alter the biology of the cell. However, B-lymphocytes are crucial inflammatory mediators in asthma. Moreover, Ding et al.(49) recently reported that 70% of cis-eQTL in LCLs is shared with skin. Furthermore, our analysis showed the differential connectivity between good-responders

and poor-responders. Combined, this evidence suggests that the results from a LCL analysis may also have large overlap with those from an analysis based on primary cells.

Conclusion

In conclusion, we have used PANDA to elucidate differences between good vs. poor clinical corticosteroid responders in asthma. Our functional results from two key transcription factors suggest that differential drug response networks built by PANDA method are valid; further validation of other novel transcription regulators may yield additional biologic and translational insights into corticosteroid response. Our results indicate that biology between responders and poor-responders does not necessarily emanate from differential expression, but may instead be from differential connectivity.

ACCEPTED MANUSCRIPT

References:

1. Masoli M, Fabian D, Holt S, Beasley R, Global Initiative for Asthma P. The global burden of asthma: executive summary of the GINA Dissemination Committee report. Allergy. 2004;59(5):469-78.
2. Drazen JM, Silverman EK, Lee TH. Heterogeneity of therapeutic responses in asthma. Br Med Bull. 2000;56(4):1054-70.
3. Liggett SB. Pharmacogenetic applications of the Human Genome project. Nat Med. 2001;7(3):281-3.
4. Silverman EK, Hjoberg J, Palmer LJ. Application of Pharmacogenetics to the Therapeutics of Asthma. In: Eissa NT, Huston D, editors. Therapeutic Targets of Airway Inflammation. 177. New York: Marcel Dekker; 2002.
5. Byrd RP, Jr., Krishnaswamy G, Roy TM. Difficult-to-manage asthma. How to pinpoint the exacerbating factors. Postgrad Med. 2000;108(6):37-40, 5-6, 9-50 passim.
6. Araujo AC, Ferraz E, Borges Mde C, Filho JT, Vianna EO. Investigation of factors associated with difficult-to-control asthma. J Bras Pneumol. 2007;33(5):495-501.
7. Ortega VE. Pharmacogenetics of beta2 adrenergic receptor agonists in asthma management. Clin Genet. 2014;86(1):12-20.
8. Ortega VE, Wechsler ME. Asthma pharmacogenetics: responding to the call for a personalized approach. Curr Opin Allergy Clin Immunol. 2013;13(4):399-409.
9. Miller SM, Ortega VE. Pharmacogenetics and the development of personalized approaches for combination therapy in asthma. Curr Allergy Asthma Rep. 2013;13(5):443-52.
10. Esposito S, Tenconi R, Lelii M, Preti V, Nazzari E, Consolo S, et al. Possible molecular mechanisms linking air pollution and asthma in children. BMC Pulm Med. 2014;14:31.
11. Lima JJ, Blake KV, Tantisira KG, Weiss ST. Pharmacogenetics of asthma. Curr Opin Pulm Med. 2009;15(1):57-62.
12. Dahlin A, Litonjua A, Irvin CG, Peters SP, Lima JJ, Kubo M, et al. Genome-wide association study of leukotriene modifier response in asthma. Pharmacogenomics J. 2015.
13. Tantisira KG, Damask A, Szefler SJ, Schuemann B, Markezich A, Su J, et al. Genome-wide association identifies the T gene as a novel asthma pharmacogenetic locus. Am J Respir Crit Care Med. 2012;185(12):1286-91.
14. Dahlin A, Litonjua A, Lima JJ, Tamari M, Kubo M, Irvin CG, et al. Genome-Wide Association Study Identifies Novel Pharmacogenomic Loci For Therapeutic Response to Montelukast in Asthma. PLoS One. 2015;10(6):e0129385.
15. Qiu W, Rogers AJ, Damask A, Raby BA, Klanderman BJ, Duan QL, et al. Pharmacogenomics: novel loci identification via integrating gene differential analysis and eQTL analysis. Hum Mol Genet. 2014;23(18):5017-24.
16. Trevor JL, Deshane JS. Refractory asthma: mechanisms, targets, and therapy. Allergy. 2014;69(7):817-27.
17. Al-Hajjaj MS. Difficult-to-treat asthma, is it really difficult? Ann Thorac Med. 2011;6(1):1-2.

18. Adcock IM, Lane SJ, Brown CR, Peters MJ, Lee TH, Barnes PJ. Differences in binding of glucocorticoid receptor to DNA in steroid-resistant asthma. J Immunol. 1995;154(7):3500-5.
19. Sher ER, Leung DY, Surs W, Kam JC, Zieg G, Kamada AK, et al. Steroid-resistant asthma. Cellular mechanisms contributing to inadequate response to glucocorticoid therapy. J Clin Invest. 1994;93(1):33-9.
20. Chang PJ, Michaeloudes C, Zhu J, Shaikh N, Baker J, Chung KF, et al. Impaired nuclear translocation of the glucocorticoid receptor in corticosteroid-insensitive airway smooth muscle in severe asthma. Am J Respir Crit Care Med. 2015;191(1):54-62.
21. Smirnov DA, Morley M, Shin E, Spielman RS, Cheung VG. Genetic analysis of radiation-induced changes in human gene expression. Nature. 2009;459(7246):587-91.
22. Maranville JC, Luca F, Richards AL, Wen X, Witonsky DB, Baxter S, et al. Interactions between glucocorticoid treatment and cis-regulatory polymorphisms contribute to cellular response phenotypes. PLoS Genet. 2011;7(7):e1002162.
23. Grundberg E, Adoue V, Kwan T, Ge B, Duan QL, Lam KC, et al. Global analysis of the impact of environmental perturbation on cis-regulation of gene expression. PLoS genetics. 2011;7(1):e1001279.
24. Qiu W, Rogers AJ, Damask A, Raby BA, Klanderman BJ, Duan QL, et al. Pharmacogenomics: novel loci identification via integrating gene differential analysis and eQTL analysis. Human molecular genetics. 2014;23(18):5017-24.
25. Glass K, Huttenhower C, Quackenbush J, Yuan GC. Passing messages between biological networks to refine predicted interactions. PLoS One. 2013;8(5):e64832.
26. Childhood Asthma Management Program Research Group. The Childhood Asthma Management Program (CAMP): design, rationale, and methods. Control Clin Trials. 1999;20(1):91-120.
27. The Childhood Asthma Management Program Research Group. Long-term effects of budesonide or nedocromil in children with asthma. The Childhood Asthma Management Program Research Group. N Engl J Med. 2000;343(15):1054-63.
28. Yoshida NL, Miyashita T, U M, Yamada M, Reed JC, Sugita Y, et al. Analysis of gene expression patterns during glucocorticoid-induced apoptosis using oligonucleotide arrays. Biochem Biophys Res Commun. 2002;293(4):1254-61.
29. Hakonarson H, Bjornsdottir US, Halapi E, Bradfield J, Zink F, Mouy M, et al. Profiling of genes expressed in peripheral blood mononuclear cells predicts glucocorticoid sensitivity in asthma patients. Proc Natl Acad Sci U S A. 2005;102(41):14789-94.
30. Glass K, Quackenbush J, Silverman EK, Celli B, Rennard SI, Yuan GC, et al. Sexually-dimorphic targeting of functionally-related genes in COPD. BMC Syst Biol. 2014;8:118.
31. Glass K, Quackenbush J, Spentzos D, Haibe-Kains B, Yuan GC. A network model for angiogenesis in ovarian cancer. BMC Bioinformatics. 2015;16:115.
32. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57.

576 33. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools:
577 paths toward the comprehensive functional analysis of large gene lists. Nucleic
578 Acids Res. 2009;37(1):1-13.
579 34. Li CS, Zhang Q, Lim MK, Sheen DH, Shim SC, Kim JY, et al. Association of FOXJ1
580 polymorphisms with systemic lupus erythematosus and rheumatoid arthritis in
581 Korean population. Exp Mol Med. 2007;39(6):805-11.
582 35. Suzuki M, Egashira N, Kajiya H, Minematsu T, Takekoshi S, Tahara S, et al.
583 ACTH and alpha-subunit are co-expressed in rare human pituitary corticotroph cell
584 adenomas proposed to originate from ACTH-committed early pituitary progenitor
585 cells. Endocr Pathol. 2008;19(1):17-26.
586 36. Geng CD, Vedeckis WV. c-Myb and members of the c-Ets family of
587 transcription factors act as molecular switches to mediate opposite steroid
588 regulation of the human glucocorticoid receptor 1A promoter. The Journal of
589 biological chemistry. 2005;280(52):43264-71.
590 37. Ma XL, Liu ZP, Ma JX, Han C, Zang JC. Dynamic expression of Runx2, Osterix
591 and AJ18 in the femoral head of steroid-induced osteonecrosis in rats. Orthop Surg.
592 2010;2(4):278-84.
593 38. Nunez BS, Vedeckis WV. Characterization of promoter 1B in the human
594 glucocorticoid receptor gene. Mol Cell Endocrinol. 2002;189(1-2):191-9.
595 39. Breslin MB, Vedeckis WV. The human glucocorticoid receptor promoter
596 upstream sequences contain binding sites for the ubiquitous transcription factor,
597 Yin Yang 1. J Steroid Biochem Mol Biol. 1998;67(5-6):369-81.
598 40. Riffel AK, Schuenemann E, Vyhlidal CA. Regulation of the CYP3A4 and
599 CYP3A7 promoters by members of the nuclear factor I transcription factor family.
600 Mol Pharmacol. 2009;76(5):1104-14.
601 41. Reiprich S, Stolt CC, Schreiner S, Parlato R, Wegner M. SoxE proteins are
602 differentially required in mouse adrenal gland development. Mol Biol Cell.
603 2008;19(4):1575-86.
604 42. Geng CD, Vedeckis WV. c-Myb and members of the c-Ets family of
605 transcription factors act as molecular switches to mediate opposite steroid
606 regulation of the human glucocorticoid receptor 1A promoter. J Biol Chem.
607 2005;280(52):43264-71.
608 43. Uddin M, Nong G, Ward J, Seumois G, Prince LR, Wilson SJ, et al. Prosurvival
609 activity for airway neutrophils in severe asthma. Thorax. 2010;65(8):684-9.
610 44. Vazquez-Tello A, Halwani R, Hamid Q, Al-Muhsen S. Glucocorticoid receptor-
611 beta up-regulation and steroid resistance induction by IL-17 and IL-23 cytokine
612 stimulation in peripheral mononuclear cells. J Clin Immunol. 2013;33(2):466-78.
613 45. White SR, Dorscheid DR. Corticosteroid-induced apoptosis of airway
614 epithelium: a potential mechanism for chronic airway epithelial damage in asthma.
615 Chest. 2002;122(6 Suppl):278S-84S.
616 46. Melis M, Siena L, Pace E, Gjomarkaj M, Profita M, Pirazzoli A, et al. Fluticasone
617 induces apoptosis in peripheral T-lymphocytes: a comparison between asthmatic
618 and normal subjects. Eur Respir J. 2002;19(2):257-66.
619 47. Beste MT, Pfaffle-Doyle N, Prentice EA, Morris SN, Lauffenburger DA,
620 Isaacson KB, et al. Molecular network analysis of endometriosis reveals a role for c-
621 Jun-regulated macrophage activation. Sci Transl Med. 2014;6(222):222ra16.

48. Cornejo S, Tantisira K, Raby BA, Weiss ST, Kaplan F. Nuclear bioavailability of the glucocorticoid receptor in a pediatric asthma cohort with variable corticosteroid responsiveness. Pediatr Res. 2015;78(5):505-12.
49. Ding J, Gudjonsson JE, Liang L, Stuart PE, Li Y, Chen W, et al. Gene expression in skin and lymphoblastoid cells: Refined statistical method reveals extensive overlap in cis-eQTL signals. Am J Hum Genet. 2010;87(6):779-89.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Figure legends:

Figure 1. The network of TFs and their targeting genes output by PANDA algorithm based on dexamethasone-treated LCSs (Figure 1a) and based on sham-treated LCLs (Figure 1b). The edges are directed from TFs (circles) to their targeting genes (rectangles). The 6 key TFs that are highlighted in Suppl. Figure 1 are labeled here.

Figure 2. We illustrate the differential connectivity by using the networks of top 50 (TF, gene) pairs in terms of the absolute difference of the edge weights between the 2 networks. The red edges are for the network of responders; the blue edges are for the network of poor-responders.

Figure 3. Heatmap of –log10(p-value) for enriched biological processes, where the p-values are for testing if a biological process is enriched or not based on the list of genes in the top 500 pairs of (TF, gene). Yellow color means larger -log10(p-value) compared to red color.

Figure 4. Parallel boxplots of normalized gene expression across different combination of Response status (good vs poor) and knock-down status (control knock-down vs siRNA knock-down) for RELA knock-down experiment and JUN knock-down experiment, Respectively.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 1. Population characteristics for the 145 Caucasian CAMP children

	All	Poor responders	Good-responders	p-value (nonresp vs resp)
Variable	N=145	N=48	N=47	
Age (year)	8.81±2.13	8.49±2.09	8.99±2.20	0.29
Female (n, %)	64, 44%	23, 48%	25, 53%	0.76**
FEV_1%	1.62±0.45	1.620.43	1.51±0.47	0.18
ΔFEV1*	7.13±14.66	-5.22±6.77	22.12±14.86	<0.01

*: the change in FEV_1% between baseline and 2-month follow-up.

**: For comparing female proportion between poor-responders and good-responders, we applied chi-squared test; for other comparisons, we applied Wilcoxon's rank sum test.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 2. The characteristics of the 31 unique TFs in both Sg and Sp.

TF	nEdge(g)	nEdge(p)	nDiff	nOverlap	nRatio	pval.Perm	stat.DiffExprs	pval.DiffExprs
TFAP2A	565	287	278	0	1.97	0.309	-2.42	**1.67E-02**
GATA3	548	817	-269	0	0.67	**0.042**	3.65	**3.60E-04**
SP1	893	630	263	2	1.42	0.249	-4.09	**6.84E-05**
SOX5	749	995	-246	6	0.75	0.160	-1.21	2.29E-01
ARID3A	760	998	-238	3	0.76	0.164	-15.15	**2.17E-32**
NFATC2	423	633	-210	0	0.67	**0.022**	-1.52	1.30E-01
PAX5	346	151	195	0	2.29	0.327	-0.96	3.37E-01
NKX2-5	273	373	-100	0	0.73	0.207	-1.20	2.33E-01
PRRX2	546	643	-97	1	0.85	0.163	-0.44	6.64E-01
AHR	292	386	-94	0	0.76	0.153	1.78	7.67E-02
BRCA1	230	138	92	0	1.67	0.075	5.48	**1.68E-07**
YY1	237	146	91	0	1.62	0.110	0.42	6.77E-01
SPI1	251	163	88	0	1.54	**0.033**	-8.13	**1.29E-13**
ARNT	223	309	-86	0	0.72	0.081	-2.52	**1.27E-02**
NFIC	339	263	76	0	1.29	0.147	-4.05	**8.03E-05**
NKX3-1	228	303	-75	0	0.75	0.257	-2.85	**4.90E-03**
SOX10	352	280	72	1	1.26	0.210	-0.60	5.50E-01
EGR1	328	371	-43	0	0.88	0.374	-5.31	**3.67E-07**
MAFG	60	18	42	0	3.33	0.054	-7.61	**2.48E-12**
NFE2L1	60	18	42	0	3.33	0.054	-6.26	**3.68E-09**
GATA2	701	660	41	0	1.06	0.288	-0.25	8.07E-01
ETS1	650	609	41	0	1.07	0.293	8.42	**2.33E-14**
SPIB	91	60	31	0	1.52	0.132	-18.52	**4.10E-41**
KLF4	46	18	28	0	2.56	0.381	-0.92	3.60E-01
FEV	34	11	23	0	3.09	**0.046**	-0.45	6.55E-01
ZNF354C	697	677	20	0	1.03	0.360	1.29	2.00E-01
NFKB1	20	1	19	0	20.00	**0.039**	-1.38	1.69E-01

ACCEPTED MANUSCRIPT

GABPA	34	17	17	0	2.00	0.225	-4.20	**4.56E-05**
CTCF	13	4	9	0	3.25	0.409	-0.71	4.76E-01
JUN	2	7	-5	0	0.29	**0.019**	-1.22	2.24E-01
NFYA	7	8	-1	0	0.88	0.451	-4.87	**2.77E-06**

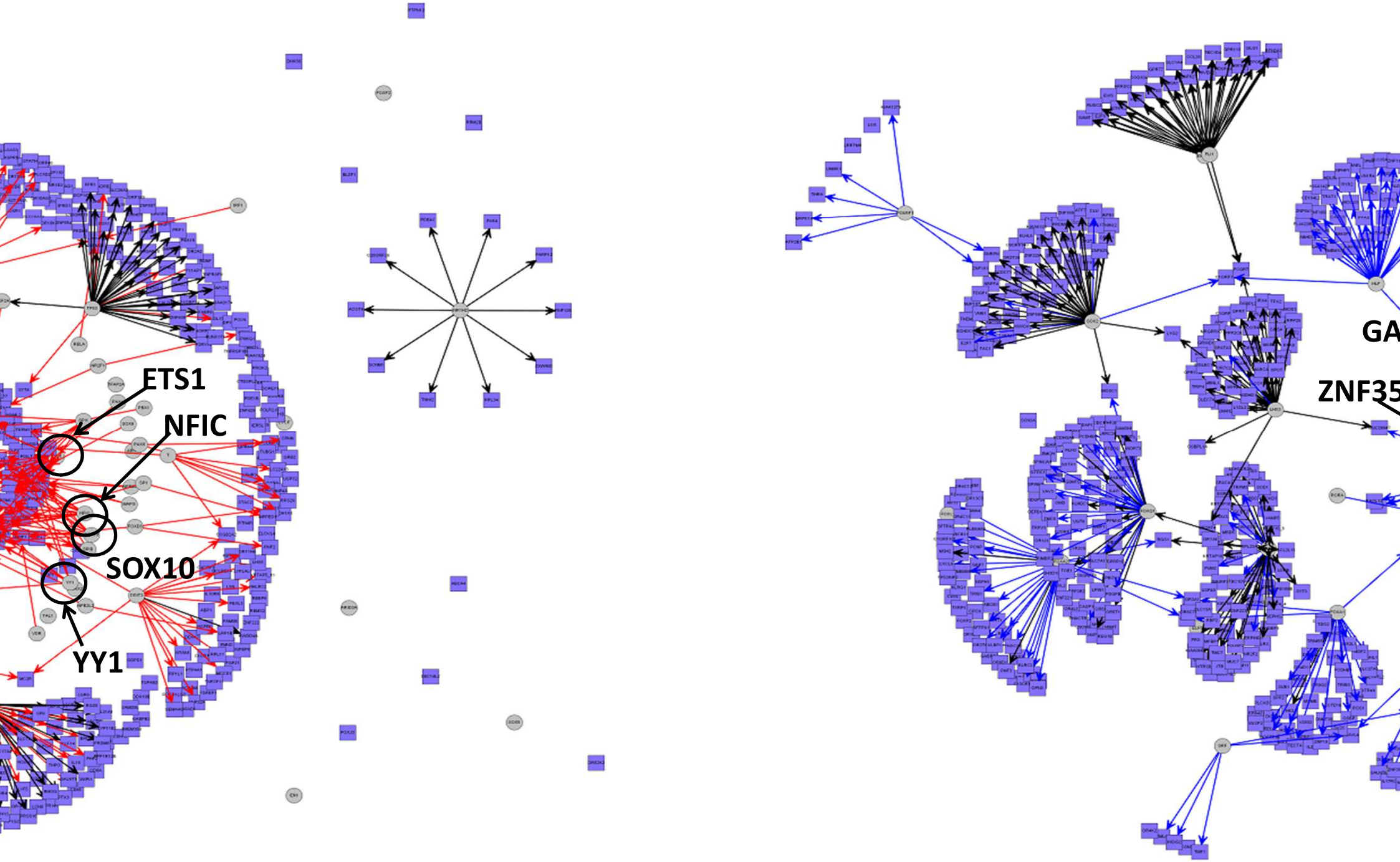
nDiff=nResp-nNonResp. nRatio=nResp/nNonResp. pval.Perm=p-value for the significance of nRatio by permutation. Six TFs (GATA3, NFATC2, SPI1, FEV, NFKB1, and JUN) have pval.Perm<0.05. stat.DiffExprs and pval.DiffExprs are test statistic and p-value for testing if a TF is differentially expressed between dexamethasone-treated cell lines and sham-treated cell lines by using two sample t test. Positive stat.DiffExprs indicates mean expression levels of the TF in dexamethasone-treated cell lines is higher than that in sham-treated cell lines. Sixteen TFs have pval.DiffExprs < 0.05.

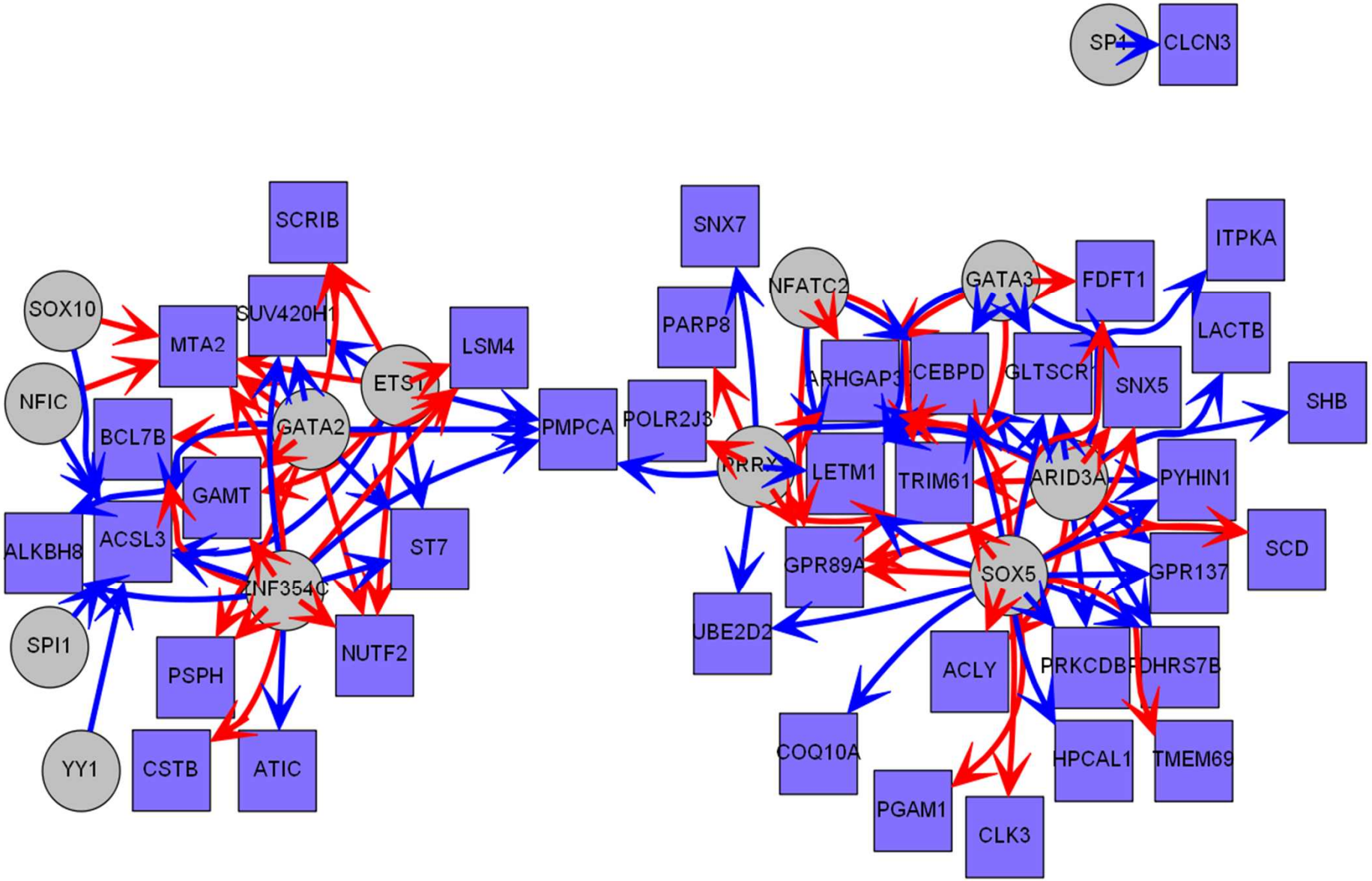
Table 3. P-values for testing if expression level s of a down-stream gene normalized by GAPDH for good-responders are different from those for poor-responders.

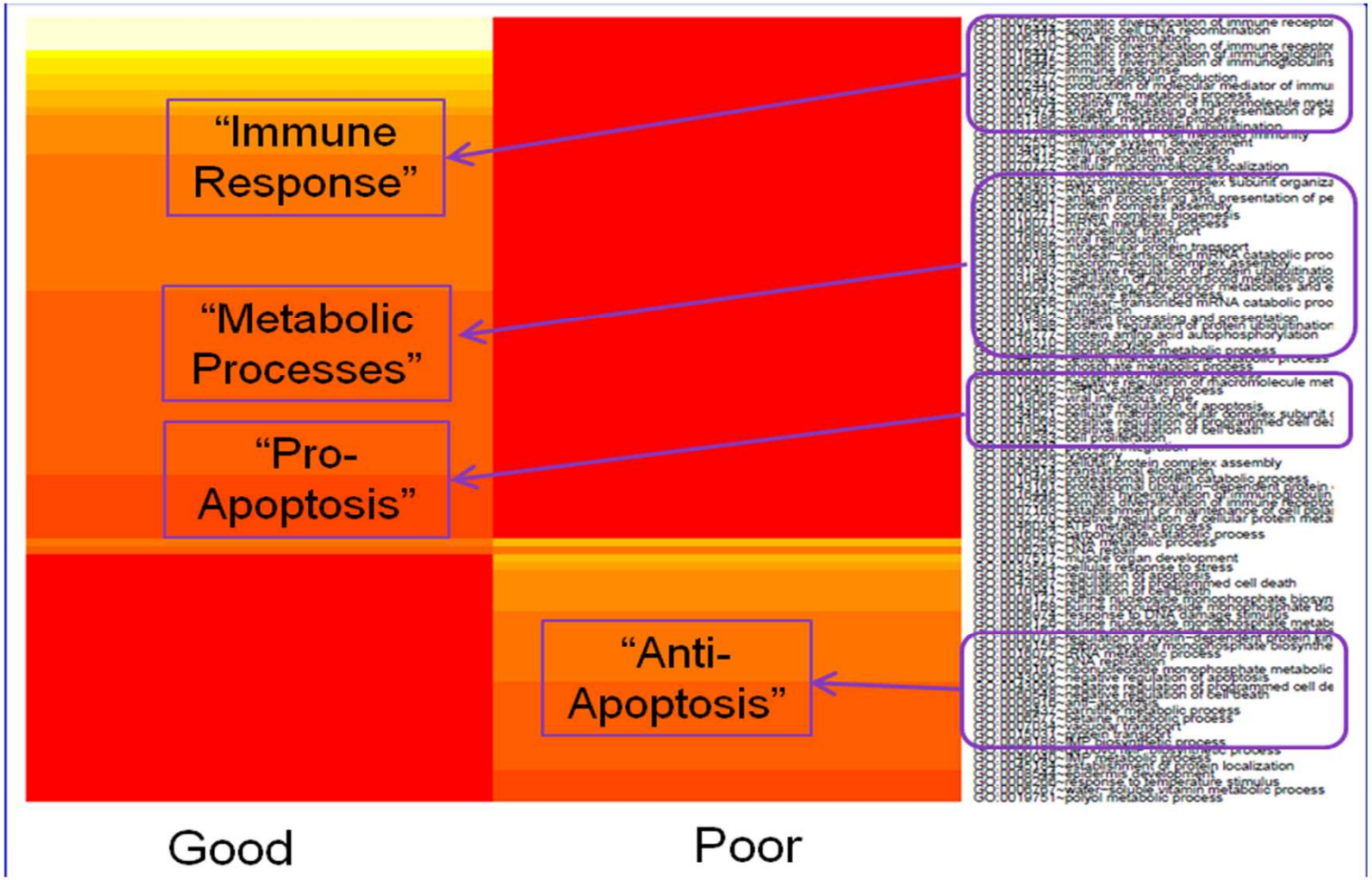
TF	down-stream gene	statistic	p-value
RELA	MANBA	0.75	4.59E-01
RELA	NR3C1	1.48	1.45E-01
RELA	PPT2	0.00	1.00E+00
RELA	TCP1	-1.95	5.60E-02
RELA	CEBPD	4.77	**1.47E-05**
JUN	NR3C1	-0.64	5.27E-01
JUN	MTA2	-2.01	5.00E-02
JUN	TMEM53	-2.72	**8.88E-03**
JUN	ACOT8	-1.38	1.73E-01
JUN	PDGFR8	-0.71	4.83E-01
JUN	ACSL3	-1.30	2.00E-01

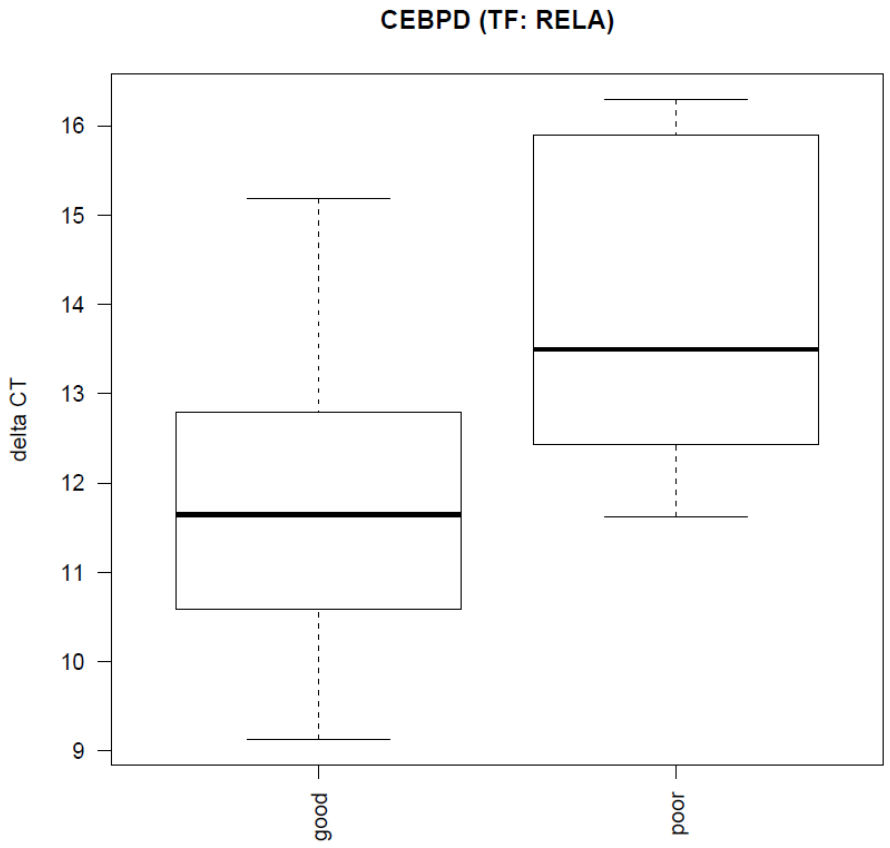
Statistic: the t-value for testing if the expression level in good responder is different from that in poor-responder by using general linear model. Negative value of ‘statistic’ means the expression level in good responder is higher than that in poor-responder.

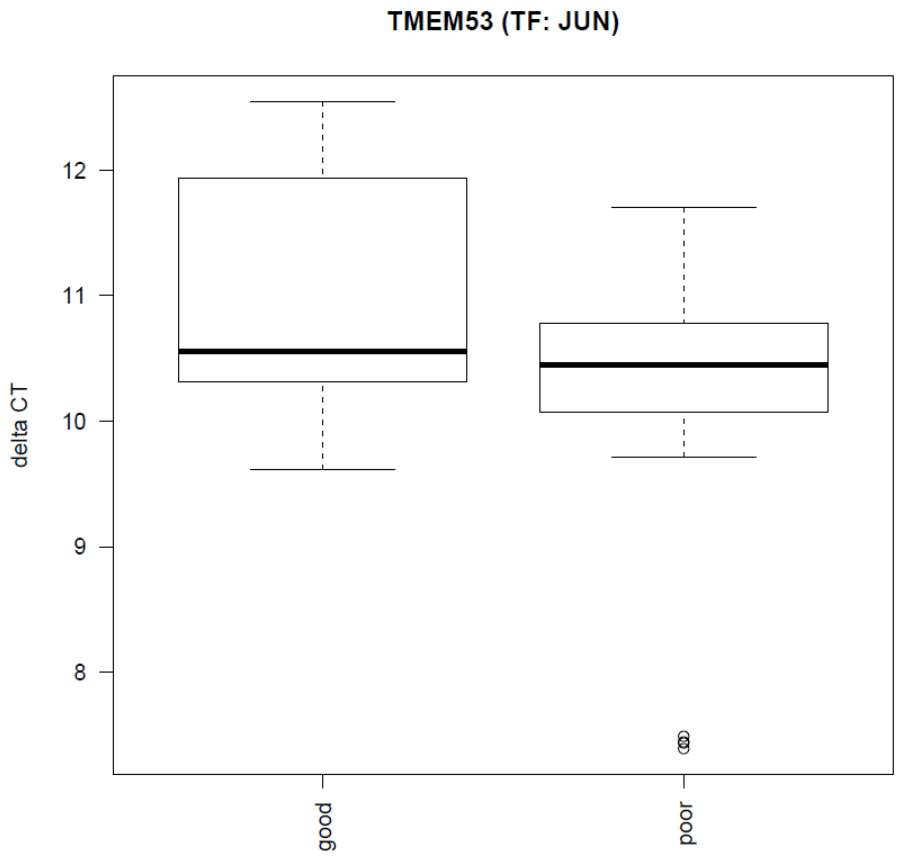
ACCEPTED MANUSCRIPT


ETS1
NFIC
SOX10
YY1
GA
ZNF35


SP1
CLCN3
SCRIB
SOX10
MTA2
SUV420H1
LSM4
NFIC
ETS1
BCL7B
GATA2
PMPCA
GAMT
ALKBH8
ACSL3
ST7
ZNF354C
SPI1
PSPH
NUTF2
YY1
CSTB
ATIC
SNX7
PARP8
POLR2J3
NFATC2
GATA3
FDFT1
ITPKA
LACTB
ARHGAP3
CEBPD
GLTSCR1
SNX5
SHB
PRRX1
LETM1
TRIM61
ARID3A
PYHIN1
GPR89A
SOX5
GPR137
SCD
UBE2D2
ACLY
PRKCDBP
DHRS7B
COQ10A
HPCAL1
TMEM69
PGAM1
CLK3


"Immune Response"
"Metabolic Processes"
"Pro-Apoptosis"
"Anti-Apoptosis"
Good
Poor


CEBPD (TF: RELA)
delta CT
16
15
14
13
12
11
10
9
good
poor


TMEM53 (TF: JUN)
delta CT
12
11
10
9
8
good
poor