



Gene Targeting in Disease Networks

Deborah Weighill¹, Marouen Ben Guebila¹, Kimberly Glass^{1,2,3}, John Platig^{2,3}, Jen Jen Yeh⁴ and John Quackenbush^{1,2*}

¹ Department of Biostatistics, Harvard T. H. Chan School of Public Health, Harvard University, Boston, MA, United States,

² Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, United States, ³ Harvard Medical School, Harvard University, Boston, MA, United States, ⁴ Departments of Surgery and Pharmacology, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

OPEN ACCESS

Edited by:

Pierre De Meyts,
Université catholique de Louvain,
Belgium

Reviewed by:

James Lim,
University of Arizona, United States
Ankush Sharma,
University of Oslo, Norway

*Correspondence:

John Quackenbush
johnq@hsph.harvard.edu

Specialty section:

This article was submitted to
Systems Biology,
a section of the journal
Frontiers in Genetics

Received: 05 January 2021

Accepted: 15 March 2021

Published: 23 April 2021

Citation:

Weighill D, Ben Guebila M,
Glass K, Platig J, Yeh JJ and
Quackenbush J (2021) Gene
Targeting in Disease Networks.
Front. Genet. 12:649942.
doi: 10.3389/fgene.2021.649942

Profiling of whole transcriptomes has become a cornerstone of molecular biology and an invaluable tool for the characterization of clinical phenotypes and the identification of disease subtypes. Analyses of these data are becoming ever more sophisticated as we move beyond simple comparisons to consider networks of higher-order interactions and associations. Gene regulatory networks (GRNs) model the regulatory relationships of transcription factors and genes and have allowed the identification of differentially regulated processes in disease systems. In this perspective, we discuss gene targeting scores, which measure changes in inferred regulatory network interactions, and their use in identifying disease-relevant processes. In addition, we present an example analysis for pancreatic ductal adenocarcinoma (PDAC), demonstrating the power of gene targeting scores to identify differential processes between complex phenotypes, processes that would have been missed by only performing differential expression analysis. This example demonstrates that gene targeting scores are an invaluable addition to gene expression analysis in the characterization of diseases and other complex phenotypes.

Keywords: cancer genomics, network medicine, gene targeting, differential targeting, gene regulatory networks

INTRODUCTION

A core tenet of molecular biology is that phenotypic differences are reflected through patterns of differential expression of key genes involved in relevant biological processes. Since its inception, whole-genome transcript profiling has been an invaluable tool for exploring these associations and has been used in a range of applications, including identification of clinically relevant molecular subtypes in cancer exhibiting different morbidities and implications for treatment together with characteristic genes associated with these phenotypes (Rouzier et al., 2005; Collisson et al., 2011; Moffitt et al., 2015; Bailey et al., 2016; Kwa et al., 2017; Rodriguez-Salas et al., 2017; Rudin et al., 2019; Sjö Dahl et al., 2019). Studies have also found that complex patterns of association between genes represented as networks can provide additional insight and that network metrics parameterizing these associations can be used to prioritize and identify crucial disease-related genes (Ramadan et al., 2016; Dimitrakopoulos et al., 2018; Horn et al., 2018; Gumpinger et al., 2020). However, there is growing evidence that the processes regulating the expression of phenotype-associated genes can provide a more holistic picture of drivers of disease and other phenotypes. Gene regulatory networks (GRNs) are often represented as directed bipartite graphs that are used to depict inferred relationships between transcription factors (TFs) and their target genes. GRNs can be characterized by calculating “gene targeting scores,” a network topology measure that captures

the complex relationships that a gene has with other TFs and genes and represents the extent to which a gene is targeted in a given system. In this perspective, we will present gene targeting scores, discuss their meaning, and show how this network-based measure provides information about disease systems beyond that found using only differential expression analysis in the investigation and characterization of human disease.

GENE REGULATORY NETWORKS: CHARACTERIZATION OF SYSTEMS

Networks are useful tools for representing and analyzing large, complex datasets because they capture information about the *relationships* within a system rather than simply the state of individual components. This is an important distinction, as can be illustrated with a small toy example first described in Glass et al. (2014; **Figure 1**). In this example, we consider the expression of four genes in nine healthy individuals and nine individuals with a disease (**Figures 1A,B**). Comparing expression levels between healthy and diseased individuals, we find that none are differentially expressed (**Figures 1C,D**). However, when looking at the co-expression of these genes in each group of individuals, we see that the genes are *differentially co-expressed* between groups. For example, in healthy individuals, gene G1 is co-expressed with gene G2 (**Figure 1E**), whereas in diseased individuals, gene G1 is co-expressed with gene G3 (**Figure 1F**). This illustrates that differential expression analysis alone may miss important correlations or regulatory relationships that distinguish biological states such as healthy and disease.

This does not mean that gene expression analysis alone is not useful. Differential expression analyses have contributed to many key advances in our understanding of disease. For example, much of our understanding of the complexities of human cancers is derived from large-scale expression profiling of cancer, such as that carried out by The Cancer Genome Atlas (TCGA)¹, where the expression-based subtypes that have been identified possess distinct clinical characteristics. In pancreatic ductal adenocarcinoma (PDAC), several studies have used expression profiling to identify molecular subtypes (Collisson et al., 2011; Moffitt et al., 2015; Bailey et al., 2016; Rashid et al., 2020; Puleo et al., 2018; Maurer et al., 2019). However, we suggest that a more comprehensive molecular characterization of diseases can be achieved by exploring inferred regulatory network differences and differential gene targeting.

In 2013, Glass et al. (2013) introduced the PANDA (Passing Attributes between Networks for Data Assimilation) framework for GRN construction. This method takes a data integration approach to GRN construction by using message passing to combine multiple data sources. PANDA predicts regulatory relationships between TFs and genes by considering three main sources of information: (1) a TF–gene network “adjacency matrix” representing an initial guess of which TFs regulate which genes based on the presence/absence of TF motif in the promoter regions of genes, (2) a protein–protein interaction

network “co-operativity matrix” that recognizes that many TFs exert their influence through regulatory complexes, and (3) a gene co-expression matrix representing gene–gene relationships initially based on correlation in expression patterns across a set of samples. These three different sources of information are iteratively updated using a message-passing algorithm, using the logic that if two genes are co-expressed, they are more likely to be co-regulated by a similar set of TFs (**Figure 2A**), and that if two TFs interact, they are more likely to bind promoter regions as a complex and co-regulate the expression of their target genes (**Figure 2B**). In this process, the TF–gene “edge weights” in the adjacency matrix are updated to reflect the evidence supporting a regulatory interaction; the refinement of edge weights through message passing has been found to improve the prediction accuracy of GRNs, validated through prediction of chromatin immunoprecipitation (ChIP)-seq-derived TF binding.

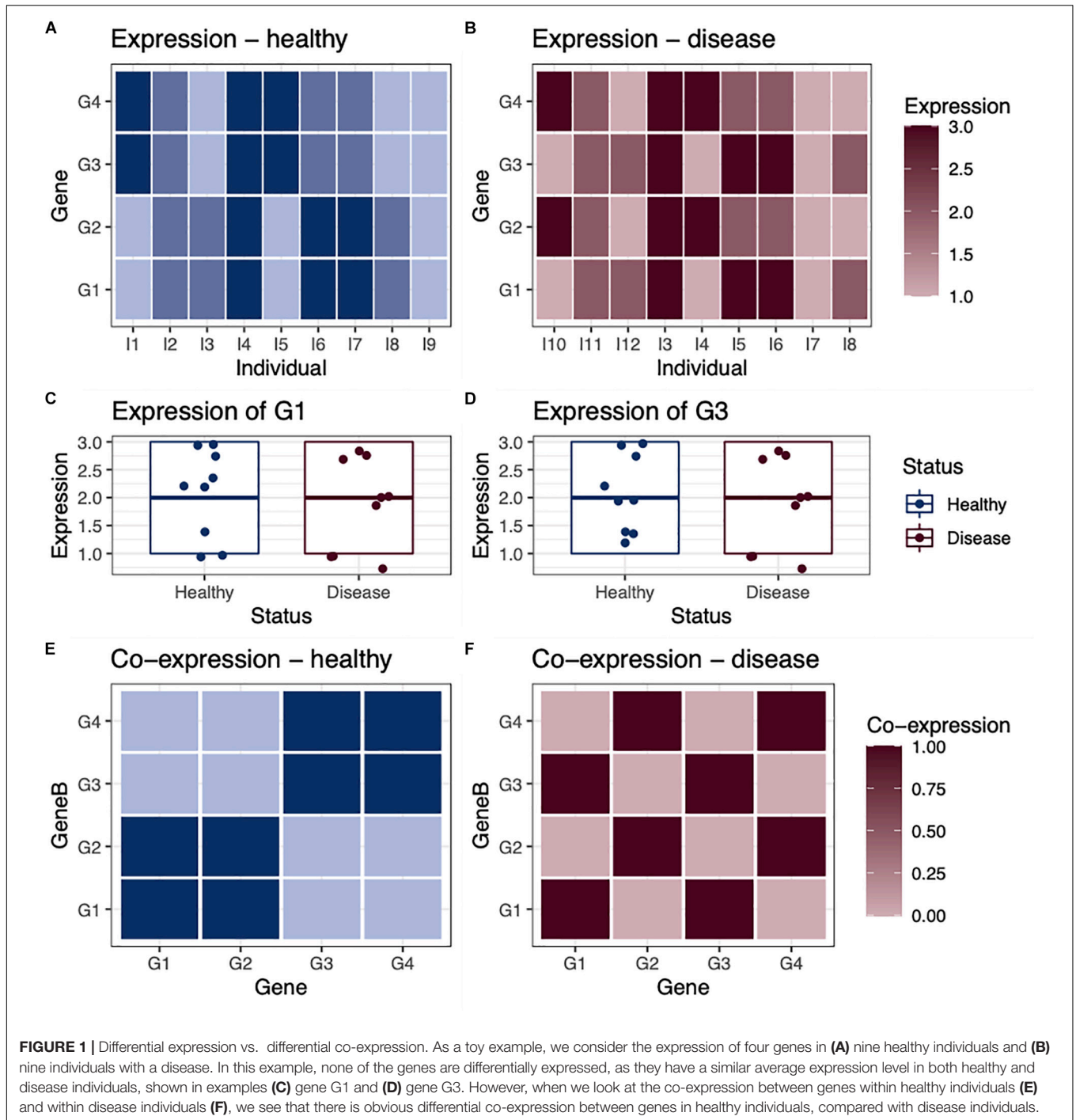
PANDA has been used to investigate gene regulatory relationships in several disease contexts, including chronic obstructive pulmonary disease (COPD) (Glass et al., 2014), asthma (Qiu et al., 2018), ovarian cancer (Glass et al., 2015), and colorectal cancer (Vargas et al., 2016; Lopes-Ramos et al., 2018). In addition, single-sample versions of PANDA GRNs, derived using a method called LIONESS (Linear Interpolation to Obtain Network Estimates for Single Samples) (Kuijjer et al., 2019b), have been used to study sex-linked differences in colon cancer (Lopes-Ramos et al., 2018) as well sex-related differences in gene regulation (Lopes-Ramos et al., 2020) in tissues from the Genotype-Tissue Expression (GTEx) project (Lonsdale et al., 2013).

GENE TARGETING SCORE: IDENTIFYING INFORMATIVE REGULATORY PROCESSES

The use of GRNs in the analysis of disease relies on analysis of the “gene targeting score,” a numerical score representing the extent to which a gene is targeted by TFs in a given biological context. The gene targeting score is calculated by summing the weights of all inbound regulatory edges for a gene (**Figure 2C**). Because of the way in which PANDA estimates edge weights, a gene’s targeting score synthesizes multiple lines of evidence—TF motif data, TF–TF interactions, and gene expression correlation. Thus, gene targeting scores are not necessarily correlated with absolute gene expression levels, and consequently, differential targeting is not necessarily correlated with differential gene expression.

Sonawane et al. (2017) used PANDA to construct tissue-specific GRNs for 38 tissues in GTEx and investigated the tissue specificity of TF–gene regulatory relationships. They found many tissue-specific regulatory relationships that would have been missed by using expression information alone. For example, when comparing the tissue-specific regulatory activity of TFs based on gene expression with that deduced using network targeting, they found that TF regulation of tissue-specific function was evident when using gene targeting metrics, but it was largely independent of TF expression level (Sonawane et al., 2017). PANDA analysis also identified unique, tissue-specific

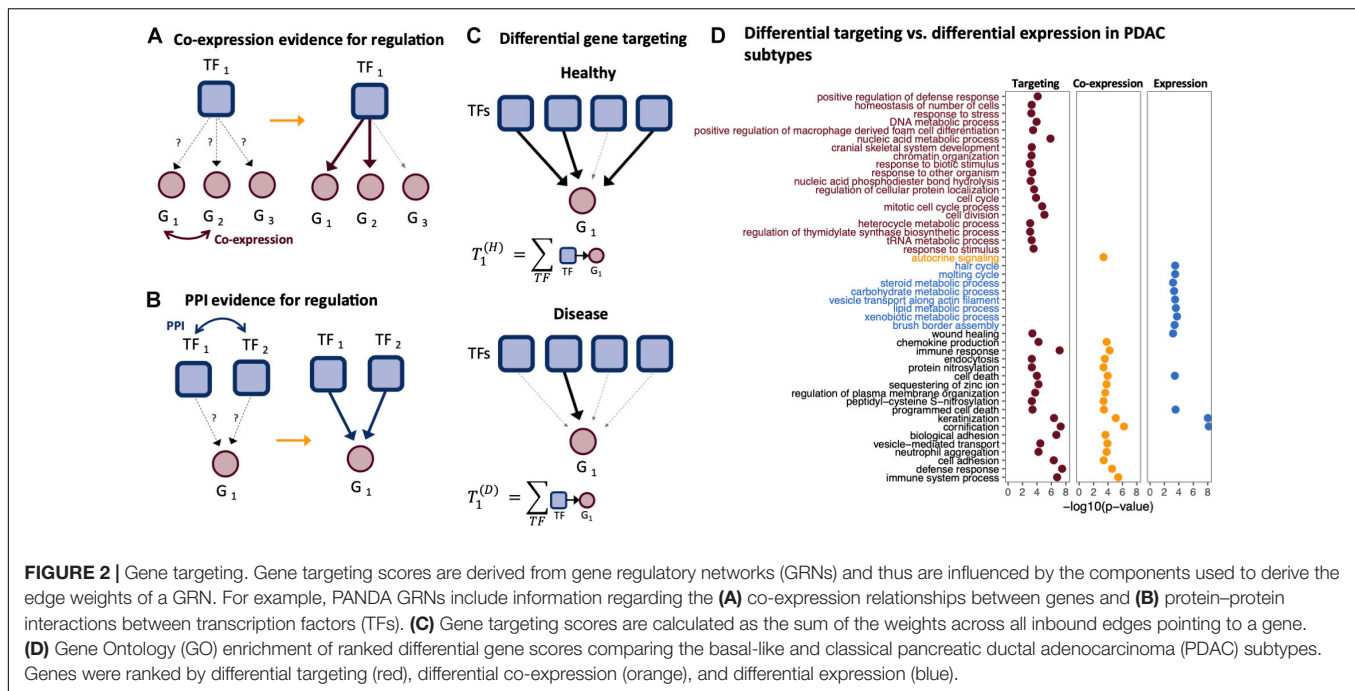
¹<https://www.cancer.gov/tcga>



targeting patterns in the TF–gene edges and found significant enrichment of tissue-specific regulatory edges targeting tissue-specific expressed genes. This example demonstrates how differences in tissue-specific regulatory relationships between TFs and genes can give rise to the distinct phenotype by altering regulation of key biological processes (Sonawane et al., 2017).

A study of COPD by Glass et al. (2014) found patterns of gene targeting that differed between men and women and

offered a possible explanation for higher disease susceptibility of women compared with men. They first compared gene expression between males and females with COPD and found little evidence of differential expression of autosomal genes. Using a resampling approach, they constructed an ensemble of 100 male and 100 female PANDA GRNs. They calculated a targeting score for each gene in each network defined as the sum of all inbound edge weights for the gene. They found several



genes that were not differentially expressed but nevertheless had significantly different targeting scores between the sexes. Pre-ranked gene set enrichment analysis based on targeting scores found many processes associated with mitochondrial function that were more highly targeted in females; these processes had previously been implicated in many aspects of COPD and lung disease (Glass et al., 2014). These results suggest that differential regulation of processes associated with disease may alter disease development and progression in meaningful ways.

Lopes-Ramos et al. (2020) investigated sex differences in gene expression and gene regulation in 29 human tissues by constructing individual-specific networks for each sample in each tissue. Differential edge weights between males and females were identified, and genes were classified as differentially targeted if at least 5% of their inbound edge weights were significantly different between males and females. This allowed genes to be classified as being male-biased if most (>60%) of the inbound differential edges were higher in males, female-biased if most (>60%) of the inbound differential edges were higher in females, and sex-divergent if the number of inbound differential edges was evenly split between being higher in males and higher in females (Lopes-Ramos et al., 2020). Consistent with previous studies, they found little differential gene expression except in breast tissue, with the median number of differentially expressed genes across tissues equal to 64. However, widespread sex-biased targeting was detected in all tissues, with a median number of differentially targeted genes across tissues equal to 169. Interestingly, the sex hormone receptors ESR1, ESR2, and AR were differentially targeted between male and female individuals in several tissues such as the breast, heart, and blood, despite the fact that those hormone receptors were not differentially expressed.

SPECIFIC EXAMPLE: PANCREATIC DUCTAL ADENOCARCINOMA SUBTYPES

PDAC is a lethal disease involving heterogeneous tumors composed of diverse cell types including tumor epithelial cells and components of the tumor microenvironment such as immune cells and fibroblasts. Molecular subtypes of PDAC have been identified through gene expression analysis (Collisson et al., 2011; Moffitt et al., 2015; Bailey et al., 2016; Puleo et al., 2018; Maurer et al., 2019), and the basal-like and classical subtypes first identified by Moffitt et al. have been associated with both prognosis and treatment response (Moffitt et al., 2015; Aung et al., 2018; Rashid et al., 2020; O’Kane et al., 2020). The basal-like subtype is associated with worse median patient survival and resistance to chemotherapy (Moffitt et al., 2015; O’Kane et al., 2020) and has characteristically high expression of keratins and laminins, both structural proteins also associated with the basal subtypes of breast and bladder cancers (Damrauer et al., 2014; McConkey et al., 2014; Weinstein et al., 2014). The classical subtype shows better response to treatment and better overall survival and is marked by increased expression of GATA binding protein 6 (GATA6), a TF involved in cell differentiation.

To identify factors driving these subtypes, we compared differential gene expression and differential GRN gene targeting scores between basal-like and classical subtypes of 150 PDAC tumors using TCGA (Cancer Genome Atlas Research Network, 2017) transcripts per kilobase million (TPM) expression data processed using Recount (Collado-Torres et al., 2017). We used PANDA and LIONESS to construct sample specific GRNs and chose to limit our analysis to those genes with a high standard

deviation of logTPMs [$\text{sd}(\log\text{TPMs}) > 0.4$] across samples. For each gene in each individual tumor, a gene targeting score was calculated as the sum of all inbound edge weights surrounding each gene. Separately, we also calculated a sample-specific co-expression network for each tumor (Kuijjer et al., 2019a) and for each gene in each sample, and we calculated a gene co-expression score equal to the sum of each gene's co-expression edges surrounding the gene. We used limma (Ritchie et al., 2015) to compare the expression data, the correlation networks, and GRNs between the basal-like and classical subtypes, allowing us to identify differentially expressed genes, differentially co-expressed genes, and differentially targeted genes, respectively.

The three genes found to be most significantly differentially targeted in GRNs, but not differentially expressed, between basal-like and classical subtypes are folate receptor beta (FOLR2), hedgehog interacting protein (HHIP), and the CD209 antigen C-type lectin domain family 4 member L (CD209). FOLR2 encodes the folate receptor 2 protein and is known to be overexpressed in tumor-associated macrophages (Tie et al., 2020). HHIP codes for the hedgehog interacting protein; the hedgehog signaling pathway regulates cell differentiation and proliferation and is activated in several cancers including PDAC (Yang et al., 2010; Honselmann et al., 2015; Gu et al., 2016). CD209 codes for a C-type lectin domain family 4 protein and is a dendritic cell marker. The roles that these play in PDAC have not yet been explored.

Ranking genes according to three different metrics, differential targeting, differential co-expression, and differential expression, we performed ranked gene set enrichment analysis (Eden et al., 2009; Supek et al., 2011) to identify significantly over-represented biological process Gene Ontology (GO) terms. Both the differential targeting analysis and differential expression analysis identified keratinization, cornification, cell death, and wound healing as differentiating basal-like and classical samples. However, several immune-related processes, epigenetic, and cell cycle processes found by differential targeting analysis were missed using differential expression alone (Figure 2). Functional enrichment using co-expression scores to rank genes identified some processes similar to those found using differential targeting but missed several important pathways related to cell cycle and other processes, such as chromatin organization.

The identification of keratinization as enriched in differentially expressed genes is consistent with previous studies that identified genes encoding keratins and laminins as biomarkers for basal-like tumors (Moffitt et al., 2015; O'Kane et al., 2020). The fact that both differential expression and differential targeting identified keratinization and cell adhesion as biological processes distinguishing PDAC subtypes serves as an internal consistency check. Differential targeting alone identified processes related to the immune system and speaks to the importance of the tumor microenvironment, which is known to influence PDAC prognosis and drug response. A high degree of tumor-associated macrophage infiltration has been linked to lower survival (Karamitopoulou, 2019), which is a known hallmark of the basal-like subtype, and it is possible that the differential targeting analysis is detecting cross-talk between the tumor and the tumor

microenvironment. The differential targeting of epigenetic functions between subtypes is consistent with reports that the basal-like and classical subtypes have distinct epigenetic landscapes (Lomber et al., 2018).

This analysis of PDAC subtypes, although abbreviated, demonstrates the power of using GRN inference and gene targeting score analysis to identify regulatory processes that characterize distinct phenotypes—including processes that are distinct from those that are associated with patterns of gene expression. The biologically relevant differences we see in targeting but not expression or co-expression suggest that regulatory control, even if not activated, is important in defining health and disease.

DISCUSSION

There is growing experimental evidence of the importance of complex regulatory processes in distinguishing phenotypes in health and disease. For example, the Wilms tumor-1 (WT1) TF is a master regulator that targets several essential genes in kidney podocyte cells. Ettou et al. (2020) investigated WT1-based gene regulation during podocyte injury and found that WT1 maintained open chromatin in the regions of its target genes but that the expression level of WT1 was not universally associated with the intensity of its binding. They also found that WT1 could cause either an increase or a decrease in the expression of its target genes. The role of complex regulatory processes is further illustrated by the work of Carnesecchi et al. (2020), who investigated how a single TF could regulate different developmental programs in various cell lineages. They showed that the Ubx TF forms different complexes with distinct binding partners in various cell lineages despite the fact that most of the interaction partners showed no differential expression across the lineages.

Taken together, the results reported by Ettou and Carnesecchi illustrate the complexity of regulatory processes and the importance that regulatory targeting plays in defining phenotype, even in instances in which a key regulator does not itself substantially change in its expression levels. Their results also point to the importance of modeling both “direct” and “indirect” regulation of genes by TFs and the complexes they form. Among the methods for GRN inference, PANDA (and by extension, PANDA+LIONESS) is singular in considering interactions between TF proteins in its model. PANDA's integrative approach using TF–TF interactions, predicted TF–gene regulatory relationships, and gene co-expression data, refines the inputs to optimize agreement between them; the resulting networks provide unique insight into regulatory processes that are linked to phenotypes.

The work summarized in this perspective demonstrates the value of the gene targeting score as a metric for assessing the drivers of phenotypic differences. Gene targeting scores not only capture structural characteristics of regulatory networks but also allow for the identification of processes that may be activated in response to appropriate stimuli and in this way help to

define phenotypes and disease subtypes. For example, Lopes-Ramos et al. (2018) performed gene targeting analysis on gene expression data from colon cancer tumor samples and discovered differences between males and females in the regulation of genes involved drug metabolism, suggesting that male and female tumor cells are programmed to respond differently. They found that genes in drug metabolism pathways, particularly those acting through cytochrome P450, had higher targeting scores in female networks than male networks. Furthermore, higher targeting of the drug metabolism pathways was found to correlate with patient survival, indicating a mechanism for sex-divergent response to chemotherapy in colon cancer.

Our application of PANDA and LIONESS in comparing PDAC subtypes demonstrates the value of the GRN-based approach and of using network-based metrics such as gene targeting to characterize properties of biological systems. We constructed sample-specific GRNs for 150 PDAC tumors and used gene targeting to compare the topologies of networks derived from basal-like and classical tumor subtypes. We found that differential targeting analysis identified compelling differences between the two subtypes in the regulation of processes related to cell cycle, immune, and epigenetic functions, none of which were seen in a standard differential expression analysis. Given that PDAC tumors are known to exhibit immune infiltration, and that the subtypes differ in both their epigenetic landscapes and patient survival, our identification of relevant processes illustrates how GRN-based methods can provide important and relevant biological insights into disease-associated processes beyond what is seen using other analytical methods.

PANDA and LIONESS software for GRN analyses and identification of differential targeting are freely available as open-source tools with extensive documentation (netzoo.github.io)

and can easily be implemented in most analytical workflows. We hope that this review motivates the broader use and appreciation of gene targeting analysis.

DATA AVAILABILITY STATEMENT

Publicly available data analyzed in this study were obtained from recount2 (<https://jhubiostatistics.shinyapps.io/recount/>). We provide an interactive Jupyter notebook hosted in netbooks (netbooks.networkmedicine.org) to reproduce the differential targeting analysis and the PDAC gene regulatory networks are available in GRAND database (<https://grand.networkmedicine.org/cancers>).

AUTHOR CONTRIBUTIONS

DW drafted the manuscript and performed the analysis. All authors discussed, planned, reviewed and edited the manuscript.

FUNDING

DW, MBG, and JQ were supported by a grant from the US National Cancer Institute (NCI), R35CA220523. MBG and JQ were further supported by NCI grant U24CA231846. KG was supported by a grant from the US National Heart, Lung, and Blood Institute (NHLBI), K25HL133599. JP was supported by a grant from the US National Heart, Lung and Blood Institute, K25HL140186. JYJ was supported by R01CA199064, CA193650, and U24CA211000.

REFERENCES

- Aung, K. L., Fischer, S. E., Denroche, R. E., Jang, G.-H., Dodd, A., Creighton, S., et al. (2018). Genomics-driven precision medicine for advanced pancreatic cancer: early results from the COMPASS trial. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 24, 1344–1354. doi: 10.1158/1078-0432.CCR-17-2994
- Bailey, P., Chang, D. K., Nones, K., Johns, A. L., Patch, A.-M., Gingras, M.-C., et al. (2016). Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* 531, 47–52. doi: 10.1038/nature16965
- Cancer Genome Atlas Research Network (2017). Integrated genomic characterization of pancreatic ductal adenocarcinoma. *Cancer Cell* 32, 185–203.e13. doi: 10.1016/j.ccell.2017.07.007
- Carneseccchi, J., Sigismondo, G., Domsch, K., Baader, C. E. P., Rafiee, M.-R., Krijgsveld, J., et al. (2020). Multi-level and lineage-specific interactomes of the Hox transcription factor Ubx contribute to its functional specificity. *Nat. Commun.* 11:1388. doi: 10.1038/s41467-020-15223-x
- Collado-Torres, L., Nellore, A., Kammers, K., Ellis, S. E., Taub, M. A., Hansen, K. D., et al. (2017). Reproducible RNA-seq analysis using recount2. *Nat. Biotechnol.* 35, 319–321. doi: 10.1038/nbt.3838
- Collisson, E. A., Sadanandam, A., Olson, P., Gibb, W. J., Truitt, M., Gu, S., et al. (2011). Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat. Med.* 17, 500–503. doi: 10.1038/nm.2344
- Damrauer, J. S., Hoadley, K. A., Chism, D. D., Fan, C., Tiganelli, C. J., Wobker, S. E., et al. (2014). Intrinsic subtypes of high-grade bladder cancer reflect the hallmarks of breast cancer biology. *Proc. Natl. Acad. Sci. U.S.A.* 111, 3110–3115. doi: 10.1073/pnas.1318376111
- Dimitrakopoulos, C., Hindupur, S. K., Häfliger, L., Behr, J., Montazeri, H., Hall, M. N., et al. (2018). Network-based integration of multi-omics data for prioritizing cancer genes. *Bioinformatics* 34, 2441–2448. doi: 10.1093/bioinformatics/bty148
- Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10:48. doi: 10.1186/1471-2105-10-48
- Ettou, S., Jung, Y. L., Miyoshi, T., Jain, D., Hiratsuka, K., Schumacher, V., et al. (2020). Epigenetic transcriptional reprogramming by WT1 mediates a repair response during podocyte injury. *Sci. Adv.* 6:eabb5460. doi: 10.1126/sciadv.abb5460
- Glass, K., Huttenhower, C., Quackenbush, J., and Yuan, G.-C. (2013). Passing Messages between biological networks to refine predicted interactions. *PLoS One* 8, e64832. doi: 10.1371/journal.pone.0064832
- Glass, K., Quackenbush, J., Silverman, E. K., Celli, B., Rennard, S. I., Yuan, G.-C., et al. (2014). Sexually-dimorphic targeting of functionally-related genes in COPD. *BMC Syst. Biol.* 8:118. doi: 10.1186/s12918-014-0118-y
- Glass, K., Quackenbush, J., Spentzos, D., Haibe-Kains, B., and Yuan, G.-C. (2015). A network model for angiogenesis in ovarian cancer. *BMC Bioinformatics* 16:115. doi: 10.1186/s12859-015-0551-y
- Gu, D., Schlotman, K. E., and Xie, J. (2016). Deciphering the role of hedgehog signaling in pancreatic cancer. *J. Biomed. Res.* 30, 353–360. doi: 10.7555/JBR.30.20150107
- Gumpinger, A. C., Lage, K., Horn, H., and Borgwardt, K. (2020). Prediction of cancer driver genes through network-based moment propagation of mutation scores. *Bioinformatics* 36, i508–i515. doi: 10.1093/bioinformatics/btaa452

- Honselmann, K. C., Pross, M., Wellner, U. F., Deichmann, S., Keck, T., Bausch, D., et al. (2015). Regulation mechanisms of the hedgehog pathway in pancreatic cancer: a review. *JOP* 16, 25–32.
- Horn, H., Lawrence, M. S., Chouinard, C. R., Shrestha, Y., Hu, J. X., Worstell, E., et al. (2018). NetSig: network-based discovery from cancer genomes. *Nat. Methods* 15, 61–66. doi: 10.1038/nmeth.4514
- Karamitopoulou, E. (2019). Tumour microenvironment of pancreatic cancer: immune landscape is dictated by molecular and histopathological features. *Br. J. Cancer* 121, 5–14. doi: 10.1038/s41416-019-0479-5
- Kuijjer, M. L., Hsieh, P.-H., Quackenbush, J., and Glass, K. (2019a). lionessR: single sample network inference in R. *BMC Cancer* 19:1003. doi: 10.1186/s12885-019-6235-7
- Kuijjer, M. L., Tung, M. G., Yuan, G., Quackenbush, J., and Glass, K. (2019b). Estimating Sample-Specific Regulatory Networks. *iScience* 14, 226–240. doi: 10.1016/j.isci.2019.03.021
- Kwa, M., Makris, A., and Esteva, F. J. (2017). Clinical utility of gene-expression signatures in early stage breast cancer. *Nat. Rev. Clin. Oncol.* 14, 595–610. doi: 10.1038/nrclinonc.2017.74
- Lomber, G., Blum, Y., Nicolle, R., Nair, A., Gaonkar, K. S., Marisa, L., et al. (2018). Distinct epigenetic landscapes underlie the pathobiology of pancreatic cancer subtypes. *Nat. Commun.* 9:1978. doi: 10.1038/s41467-018-04383-6
- Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., et al. (2013). The genotype-tissue expression (GTEx) project. *Nat. Genet.* 45:580.
- Lopes-Ramos, C. M., Chen, C.-Y., Kuijjer, M. L., Paulson, J. N., Sonawane, A. R., Fagny, M., et al. (2020). Sex differences in gene expression and regulatory networks across 29 human tissues. *Cell Rep.* 31:107795. doi: 10.1016/j.celrep.2020.107795
- Lopes-Ramos, C. M., Kuijjer, M. L., Ogino, S., Fuchs, C. S., DeMeo, D. L., Glass, K., et al. (2018). Gene regulatory network analysis identifies sex-linked differences in colon cancer drug metabolism. *Cancer Res.* 78, 5538–5547. doi: 10.1158/0008-5472.CAN-18-0454
- Maurer, C., Holmstrom, S. R., He, J., Laise, P., Su, T., Ahmed, A., et al. (2019). Experimental microdissection enables functional harmonisation of pancreatic cancer subtypes. *Gut* 68, 1034–1043. doi: 10.1136/gutjnl-2018-317706
- McConkey, D. J., Choi, W., and Dinney, C. P. N. (2014). New insights into subtypes of invasive bladder cancer: considerations of the clinician. *Eur. Urol.* 66, 609–610. doi: 10.1016/j.eururo.2014.05.006
- Moffitt, R. A., Marayati, R., Flate, E. L., Volmar, K. E., Loeza, S. G. H., Hoadley, K. A., et al. (2015). Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat. Genet.* 47, 1168–1178. doi: 10.1038/ng.3398
- O’Kane, G. M., Grünwald, B. T., Jang, G.-H., Masoomian, M., Picardo, S., Grant, R. C., et al. (2020). GATA6 expression distinguishes classical and basal-like subtypes in advanced pancreatic cancer. *Clin. Cancer Res.* 26, 4901–4910. doi: 10.1158/1078-0432.CCR-19-3724
- Puleo, F., Nicolle, R., Blum, Y., Cros, J., Marisa, L., Demetter, P., et al. (2018). Stratification of pancreatic ductal adenocarcinomas based on tumor and microenvironment features. *Gastroenterology* 155, 1999–2013.e3. doi: 10.1053/j.gastro.2018.08.033 * 1999-2013.e3,
- Qiu, W., Guo, F., Glass, K., Yuan, G. C., Quackenbush, J., Zhou, X., et al. (2018). Differential connectivity of gene regulatory networks distinguishes corticosteroid response in asthma. *J. Allergy Clin. Immunol.* 141, 1250–1258. doi: 10.1016/j.jaci.2017.05.052
- Ramadan, E., Alinsaif, S., and Hassan, M. R. (2016). Network topology measures for identifying disease-gene association in breast cancer. *BMC Bioinformatics* 17:274. doi: 10.1186/s12859-016-1095-5
- Rashid, N. U., Peng, X. L., Jin, C., Moffitt, R. A., Volmar, K. E., Belt, B. A., et al. (2020). Purity independent subtyping of tumors (PuriST), a clinically robust, single-sample classifier for tumor subtyping in pancreatic cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 26, 82–92. doi: 10.1158/1078-0432.CCR-19-1467
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47–e47. doi: 10.1093/nar/gkv007
- Rodriguez-Salas, N., Dominguez, G., Barderas, R., Mendiola, M., García-Albéniz, X., Maurel, J., et al. (2017). Clinical relevance of colorectal cancer molecular subtypes. *Crit. Rev. Oncol. Hematol.* 109, 9–19. doi: 10.1016/j.critrevonc.2016.11.007
- Rouzier, R., Perou, C. M., Symmans, W. F., Ibrahim, N., Cristofanilli, M., Anderson, K., et al. (2005). Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin. Cancer Res.* 11, 5678–5685. doi: 10.1158/1078-0432.CCR-04-2421
- Rudin, C. M., Poirier, J. T., Byers, L. A., Dive, C., Dowlati, A., George, J., et al. (2019). Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. *Nat. Rev. Cancer* 19, 289–297. doi: 10.1038/s41568-019-0133-9
- Sjödahl, G., Jackson, C. L., Bartlett, J. M., Siemens, D. R., and Berman, D. M. (2019). Molecular profiling in muscle-invasive bladder cancer: more than the sum of its parts. *J. Pathol.* 247, 563–573. doi: 10.1002/path.5230
- Sonawane, A. R., Platig, J., Fagny, M., Chen, C.-Y., Paulson, J. N., Lopes-Ramos, C. M., et al. (2017). Understanding tissue-specific gene regulation. *Cell Rep.* 21, 1077–1088. doi: 10.1016/j.celrep.2017.10.001
- Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6:e21800. doi: 10.1371/journal.pone.0021800
- Tie, Y., Zheng, H., He, Z., Yang, J., Shao, B., Liu, L., et al. (2020). Targeting folate receptor β positive tumor-associated macrophages in lung cancer with a folate-modified liposomal complex. *Signal Transduct. Target. Ther.* 5, 1–15. doi: 10.1038/s41392-020-0115-0
- Vargas, A. J., Quackenbush, J., and Glass, K. (2016). Diet-induced weight loss leads to a switch in gene regulatory network control in the rectal mucosa. *Genomics* 108, 126–133. doi: 10.1016/j.ygeno.2016.08.001
- Weinstein, J. N., Akbani, R., Broom, B. M., Wang, W., Verhaak, R. G. W., McConkey, D., et al. (2014). Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature* 507, 315–322. doi: 10.1038/nature12965
- Yang, L., Xie, G., Fan, Q., and Xie, J. (2010). Activation of the hedgehog-signaling pathway in human cancer and the clinical implications. *Oncogene* 29, 469–481. doi: 10.1038/onc.2009.392

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Weighill, Ben Guebila, Glass, Platig, Yeh and Quackenbush. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.