DIABETES TECHNOLOGY & THERAPEUTICS
Volume 7, Number 2, 2005
© Mary Ann Liebert, Inc.

A Computational Model to Define the Molecular Causes of Type 2 Diabetes Mellitus

JACK POLLARD, JR., Ph.D.,[1] ATUL J. BUTTE, M.D., Ph.D.,[2,3] STEVE HOBERMAN, S.B.,[1] MONICA JOSHI, S.B., O.I.T.,[1] JOSH LEVY, M.S.,[1] and JACQUES PAPPO, D.D.S., D.M.Sc.[1]

ABSTRACT

Background: Metabolic abnormalities associated with type 2 diabetes mellitus (DM2) are caused in part by inadequate insulin action and resulting changes in gene expression in the skeletal muscle. Two recent, independent studies of human skeletal muscle biopsies from ethnically diverse DM2 patients have identified coordinated reductions in the expression of the oxidative phosphorylation (OXPHOS) genes. Whether these reductions are a consequence or a cause of impaired insulin sensitivity remains an open question.

Methods: To address this question and to define the underlying molecular causes consistent with the expression changes reported in the muscle studies, we created a large-scale computable model to analyze the molecular actions and effects of insulin on muscle gene expression. The model enables computer-aided reasoning using over 210,000 molecular relationships assembled from the DM2 literature.

Results: We integrated the data from these muscle biopsy studies into the model and used computer-aided causal reasoning to discover mechanisms that can link alterations in OXPHOS genes to decreases in glucose transport, insulin signaling, and risk factors associated to post-transplant diabetes mellitus.

Conclusions: The emerging hypotheses describe biologic effects in DM2 and offer important cues for molecular targeted therapy.

INTRODUCTION

Type 2 diabetes mellitus (DM2), a complex multigenic disease, requires system-level measurements and modeling to improve our mechanistic understanding of its pathophysiology and corresponding therapeutic intervention. While the human genome sequence and genome-wide profiling technologies have facilitated system-level measurements, methods to interpret and integrate these measurements into models of discrete signaling, metabolic, and gene regulatory mechanisms have lagged behind.[1]

In this paper, we describe a novel approach to define discrete mechanisms of gene regulation in skeletal muscle biopsies from DM2 patients. We integrate genome-wide profiling measurements with system-level models of molecular cause-and-effect relationships and use computer-aided causal reasoning to discover mechanisms that can be causally linked

[1]Genstruct, Inc., Cambridge, Massachusetts.
[2]Research Division, Joslin Diabetes Center, Boston, Massachusetts.
[3]Division of Informatics and Endocrinology, Children's Hospital, Boston, Massachusetts.

to altered expression profiles in DM2. The emerging hypotheses describe biologic effects in DM2 and offer important cues for molecular targeted therapy.

BACKGROUND

Metabolic abnormalities associated with DM2 are in part caused by changes in transcriptional regulatory networks of skeletal muscle. Two recent, independent transcript profile studies of human skeletal muscle biopsies from DM2 patients have shown the coordinated down-regulation in genes associated with oxidative phosphorylation and ATP biosynthesis and those expressed in the mitochondria (collectively called OXPHOS).[2,3] These studies explained altered OXPHOS transcription by regulation of the transcriptional activity of *NRF1* and *PPARGC1a* (gene names are given in the Appendix); however, discrete mechanisms that link OXPHOS transcription to both its downstream effects and its upstream causes need to be defined in order to determine if altered OXPHOS transcription is a cause or an effect of the characteristics of DM2 such as impaired insulin signaling. In this paper, we have developed a computational approach using a model of human skeletal muscle and computer-aided causal reasoning to interpret expression profiles to define discrete mechanisms of regulation in DM2.

A number of statistical and network reconstruction methods have been applied to interpret expression profiles. Statistical methods can identify associations between genes according to their expression profiles (e.g., clustering,[4,5] principal component analysis,[6] or relevance networks[7]), but frequently the associations are abstract, and even promoter analysis cannot define a discrete mechanism of regulation. Additionally, numerous network reconstruction methods from systems identification and engineering have been used to infer potential transcriptional networks solely from expression profiles (Boolean networks,[8] linear models,[9] differential equation models,[10] and Bayesian networks[11]).

While these network reconstruction methods in principle are capable of defining discrete mechanisms of regulation, as typically applied they require large numbers of expression profiles to define even small networks (e.g., 25 gene regulatory networks require >100 expression profiles for typical Bayesian methods[12]), and even under these ideal conditions, the direction and sign of the influence of regulatory relationships can be very difficult to determine solely from the expression profiles. Irrespective of the reconstruction methods employed, numerous regulatory models can be generated that are consistent with the changes in expression, and prior biological knowledge in the form of cause-and-effect relationships is crucial to evaluating and ranking potential models.[13–15]

The use of prior knowledge has received considerable attention, but its implementation on a large scale has been hampered because of the difficulty in assembling large amounts of knowledge into a unified and computationally tractable form.[1,12] We have overcome this challenge, and in the Causal modeling in biology section we describe our approach to assembling knowledge in the form of a large-scale, causal model of human skeletal muscle. In the RNA expression state changes analysis and Reverse and forward causal analysis sections we describe how we engraft profiling data into this model and use computer-aided causal analysis to define discrete mechanisms of regulation for signal, metabolic, and gene expression networks. In the Results section we apply both our causal model and analysis techniques to interpret the expression profiles from the Scandinavian[2] and Mexican-American[3] studies. In the Conclusions section we discuss our mechanistic findings relative to altered OXPHOS transcription and other characteristics of DM2.

METHODS

Causal modeling in biology

Qualitative causal models. A model that describes the influence exerted by one component in a system (e.g., protein abundance or activity) on another component is called a causal model. Causal models provide a statistical framework to infer (1) causes for changes in the system's components as it transitions between

states (e.g., what are the potential causes for changes in muscle gene expression profiles in DM2?) and (2) effects on the components resulting from system perturbations (e.g., what muscle gene expression changes are attributable only to increased insulin levels?). Causal models are also computationally attractive because this inferencing can be automated since the model can be rendered as a directed graph. In Figure 1 the transcriptional activity of *PPARG* and *NRF1* as well as the transcriptional activity of the transcriptional co-activator *PPARGC1a* are represented as nodes. The edges and their directionality assert that the transcriptional activity of *PPARGC1a* positively influences the transcriptional activity of *PPARG* and *NRF1*. Additionally, the influence of indirectly connected nodes can be identified by tracing connecting paths and multiplying the signs of the intervening edges as in the case of the *AGTR1*.

This kind of causal model is qualitative in nature in that it describes only the vector of influence that *PPARGC1a*'s transcriptional activity exerts on *PPARG*'s and *NRF1*'s transcriptional activity. The qualitative causal model does not describe the magnitude of the influence or how multiple influences should be integrated. Qualitative causal models occupy a tractable middle ground in a taxonomy of models between (1) more abstract association models that cannot be used to infer causes or effects but require no prior knowledge and (2) more detailed quantitative models that can be used to infer causes and effects but require very detailed knowledge such as diffusion or rate constants in a specific subcellular environments (for recent reviews, see Rice and Stolovitzky[16] and Ideker et al.[1]).

Creating and populating a causal model describing human skeletal muscle biology. Our modeling and analyses techniques are aimed at facilitating inference on the causes for and effects of the changes in RNA expression levels. Our analysis techniques require leveraging prior knowledge to gain efficiency and resolution over purely statistical approaches; therefore, we employ an ontology to structure and render computable knowledge about the causes for changes in activities and abundance of the components in the muscle and the effects those causes engender. At the highest level our ontology describes the measurable components in molecular biology (protein abundance, catalytic activities, biological processes, etc.) and the relationships between the components (associative or causal, including the sign of the influence). Our ontology also describes the biological context of the components and their relationships by linking them to specific organisms, tissues, cell types, and subcellular compartments.

Using this ontology as a framework, we conducted a large-scale knowledge extraction and unification effort to mine several diverse biological databases such as Locuslink, GO, OMIM, CSNDB, KEGG, and Homologene to create a model of human skeletal muscle biology and DM2. Because the potential causes for changes in the measurable components of the muscle are not recorded in a machine-readable

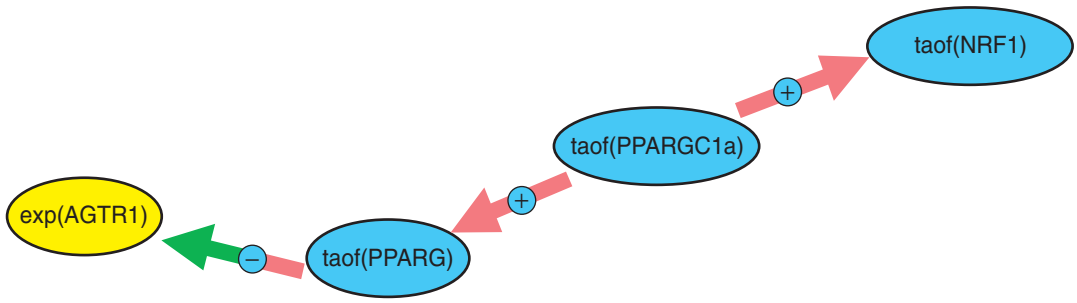

FIG. 1. Causal model showing how the transcriptional activity (taof) of *PPARG* and *NRF1* are influenced by transcriptional activity of the transcriptional co-activator *PPARGC1a*. Increases are marked in yellow, and decreases are marked in blue. Arrows trace causal paths of influence and signs indicate the direction of the influence. See Appendix for gene abbreviations.

form, we augmented our model through manual and semiautomated methods by extracting causal relationships from more than 12,000 distinct scientific articles whose abstracts were contained in Medline and deemed relevant to skeletal muscle biology. The model contains more than 157,000 molecular components (e.g., genes, protein, metabolites, pathways, etc.) and more than 210,000 relationships between those components (greater than 24,000 of which are causal). For example, the model describes the molecular effects of increased insulin, fatty acids, and glucose on muscle as well as many other causes for changes in gene expression including glucose transport, lipid metabolism, insulin signaling, glucose oxidation, and glucosamine metabolism.

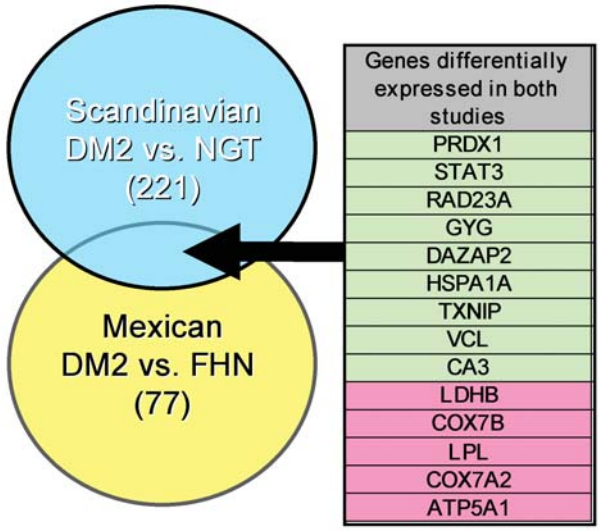

FIG. 2. Overlap of differentially expressed genes between the Scandinavian[3] and Mexican-American[2] studies. FHN, family history negative; NGT, normal glucose tolerance. See Appendix for gene abbreviations.

RNA expression state changes analysis

RNA expression profiles of human skeletal muscle biopsies from DM2 patients (five samples) compared with family history-negative control subjects (five samples) from the Mexican-American study[3] and DM2 patients (18 samples) compared with normal glucose tolerance control subjects (17 samples) from the Scandinavian study[2] were separately analyzed using R and the Bioconductor's affy and limma packages.[17,18] Briefly, raw intensities were normalized within each array with locally weighted linear regression (loess), low-intensity signals were filtered, and global scaling on each array was performed to make the intensities between the DM2 patients and control subjects comparable within each study.

Separate differential expression analyses were done on the normalized data from each study. A gene-by-gene analysis of variance was performed comparing the DM2 patients with control subjects; genes that were differentially expressed ($P < 0.05$ after false discovery rate correction) in two or more comparisons between the DM2 and the control subjects were used to classify the samples in the respective studies. Interestingly, only 14 genes were consistently modulated in both studies (Fig. 2). Using over-representation analysis the original studies concluded and our re-analysis confirms that OXPHOS genes were present in the modulated set of genes at disproportionately high frequencies. Since our causal model is qualitative in nature, we categorized the expression changes before using them in causal analysis. The expression changes were divided into three categories: genes that were increased, decreased, or unchanged in DM2 muscle compared with control.

Reverse and forward causal analysis

Our causal analysis algorithms take as inputs (1) prior knowledge in the form of a causal model and (2) a set of changes in the components of a system profiled in different states (e.g., changes in the muscle gene expression profiles comparing the DM2 state with the normal state). In the case of RNA expression data, reverse causal analysis interrogates the model to find immediate upstream transcriptional controllers (e.g., the abundance or activity of transcription factors, co-activators, co-repressors, or modulators of transcript stability) whose increase or decrease hypothetically could be a cause of the observed changes in the RNA profiles. Once immediate upstream controllers are located, forward causal analysis then scores each hypothetical cause by comparing the predicted RNA profile with the observed profile. As shown in Figure 3, reverse causal analysis finds five transcriptional mod-

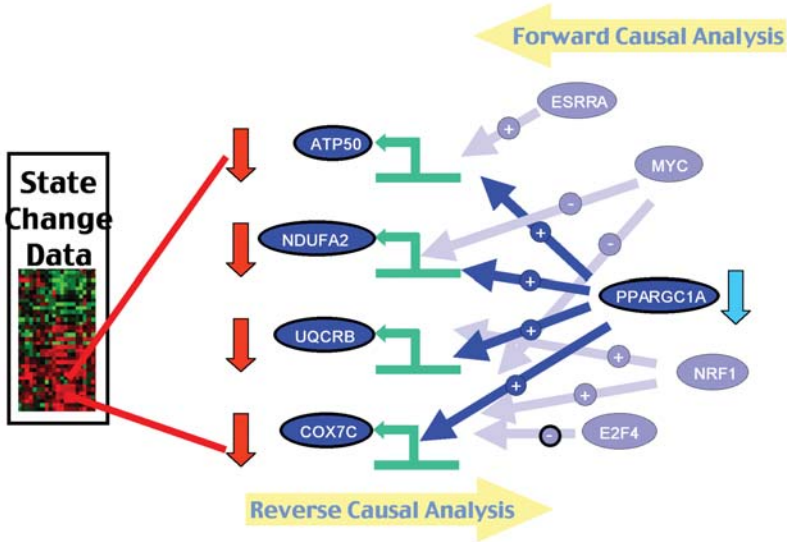

FIG. 3. Reverse and forward causal analysis. Reverse causal analysis finds five hypothetical causes for the changes in the OXPHOS genes *ATP50*, *NDUFA2*, *UQCRB*, and *COX7C*. Forward causal analysis reveals that only decreased *PPARGC1a* can explain all the observed changes. See Appendix for gene abbreviations.

ulators that could account for the observed changes in RNA profiles of *ATP50*, *NDUFA2*, *UQCRB*, and *COX7C*. Forward causal analysis predicts that only a decrease in the transcriptional activity of *PPARGC1a* can explain all of the changes with no contradictions.

Each hypothesis is scored according to two probabilistic scoring metrics—"richness" and "concordance"—which examine orthogonal aspects of the probability of a hypothetical cause explaining a given number of state changes. "Richness" is the probability that the number of observed changes in a given model could have occurred by chance alone. "Concordance" is the probability that the number of observed changes that match the directionality (e.g., increased or decreased abundance) of the changes predicted by the model could have occurred by chance alone. Only hypotheses that pass preset metric thresholds are used as inputs for continued upstream exploration in the model in progressive cycles of reverse and forward causal analyses. Collections of hypothetical causes that are highly concordant with the observed expression profiles are the output of causal analysis and form the inferred mechanism of regulation.

RESULTS

Assessing model competence

Before using the model to infer mechanisms of regulation in DM2 muscle, we probed the competency of the model to do reasoning in the muscle. Competency was assessed by introducing a set of perturbations, and then using forward causal analysis to predict effects caused by such perturbations. Numerous perturbations were used to probe the model, and Table 1 details a few examples. The predicted efficacy of the thiazolidinedione class of drugs used clinically as insulin sensitizers is shown in Figure 4. Through forward causal analysis, the model accurately predicted (1) enhanced insulin sensitivity (via *IRS1* tyrosine phosphorylation) and increased glucose import (via increased *GLUT4* expression), (2) decreased cytokine-induced insulin resistance (via decreased *TNFa* abundance and expressed *IL-6*), and (3) enhanced lipid metabolism (via increased abundance of expressed *LPL* caused by increased transcriptional activity of *PPARG*).

The competence of the model to discover the transcriptional network of the OXPHOS genes is shown in Figure 5A. Since much modulation of the OXPHOS genes is predicated on decreases in the transcriptional activity of the coactivator *PPARGC1a*, forward causal analysis was performed using *PPARGC1a* as a starting point. The causal analysis confirmed that a decrease in *PPARGC1a* transcriptional activity and consequently a decrease in *NRF1* transcriptional activity could explain the observed changes in expression profiles of OXPHOS genes from both of the studies. Furthermore, the causal analysis outlined numerous plausi-

Table 1. Assessment of Causal Model Competence for Muscle Biology and DM2

Perturbation	*Mechanism in model*	*Prediction via causal analysis*
Effect of PTP1B inhibitor on insulin signaling	PTP1B dephosphorylates insulin receptor. PTP1B inhibition increases insulin receptor signaling	Increased insulin receptor activation and increased insulin signaling
Effect of thiazolidinediones mediated through PI3K kinase activity	Increase tyrosine phosphorylation of Cbl, which activates Cbl-dependent PI3K kinase activity and its downstream signaling	Increased insulin signaling
Effect of inactivation of AMP-kinase on exercise-stimulated glucose import	AMP-kinase stimulates glucose import	Decreased exercise-stimulated glucose import
Effect of MEK inhibitor LY492002 or PTEN overexpression on ETS2 transcriptional activity	LY492002 inhibits ETS2 activation, and PTEN phosphatase activity decreases ETS2 activity through inhibition of ETS2 phosphorylation	Decreased ETS2 transcriptional activity
Effect of p85a knockout on insulin signaling as measured by AKT kinase activity	Increased of AKT abundance via inhibition of PI3K kinase activity	Increased AKT phosphorylation and kinase activity
Effect of GSK3B inhibitor on insulin-stimulated glucose transport	GSK3B inhibits insulin stimulated glucose transport	Increased insulin-stimulated glucose transport
Effect of increase in TNF protein abundance on insulin signaling	TNF inhibits IRS1 binding by promoting its serine phosphorylation	Inhibition of insulin signaling
Effect of ENPP1 phosphatase activity on insulin signaling	ENPP1 dephosphorylates insulin receptor	Inhibition of insulin signaling

See Appendix for gene abbreviations.

ble mechanisms by which decreased transcriptional activity of *PPARGC1a* could cause decreases in insulin sensitivity and glucose import in muscle. An example mechanism mediated directly through decreased *GLUT4* expression or indirectly through decreased *UCP2* expression is shown in Figure 5B. Reverse causal analyses using the expression profiles from the studies allowed us examine all such mechanisms known by the model and rank them by concordance to the observed profiles and linkage to insulin signaling as well as glucose import.

Causal analysis of expression changes in DM2 muscle

The causal model contained transcriptional control information for approximately 56% of genes that were modulated combined across the studies[2,3] (101 of 221 Scandinavian and 30 of 77 Mexican-American modulated genes). Reverse causal analyses was performed on each study separately, and regulatory mechanisms were scored and ranked based on concordance between predicted and observed expression changes. Intersection of the causal analyses across the two studies revealed a regulatory mechanism that is highly significant according to the richness and concordance metrics and can explain 49 of the 131 gene expression changes known to the model (Table 2 and Fig. 6). The mechanism details that decreased calcineurin signaling can modulate *PPARGC1a* protein abundance by decreasing the transcriptional activity of the *MEF2* family,[19] which in turn can decrease the transcriptional activity of *PPARGC1a* and correspondingly explain the decreases in *NRF1*-regulated genes. A decrease in *PPARGC1a* protein abundance is further supported by the observed changes in expression of *STAT3* and *LPL* in both studies.

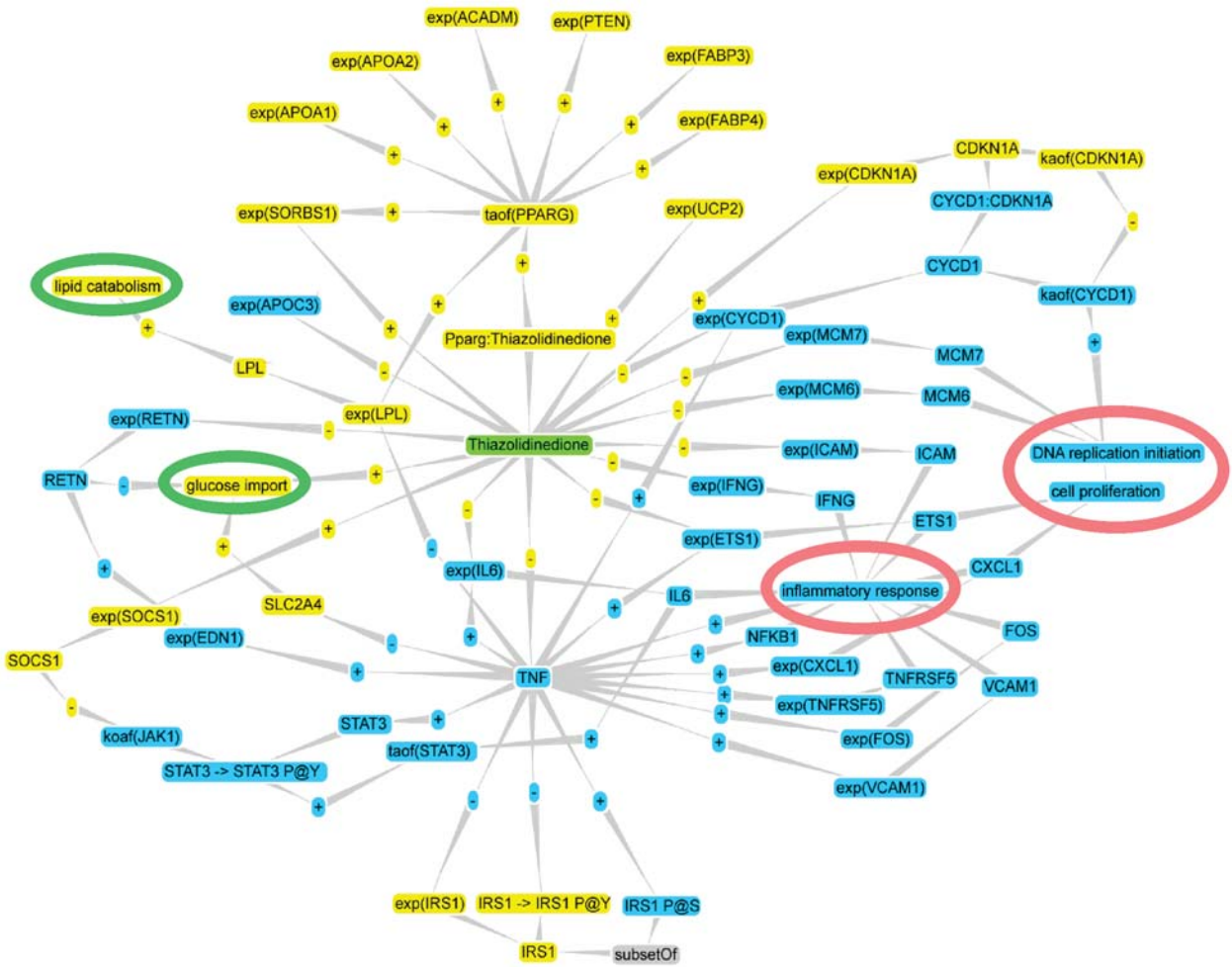

FIG. 4. In silico prediction of mechanism of insulin-sensitizing thiazolidinedione action and downstream biological effects. Via forward causal analysis the muscle model indicates increased lipid catabolism and glucose import (green circles) as well as decreased cell proliferation and inflammation response (red circles). Predicted increases are marked in yellow, and predicted decreases are marked in blue. See Appendix for gene abbreviations.

Forward causal analysis from calcineurin inhibition also causally predicted decreased glucose import and insulin signaling as evidenced by decreased *GLUT4* expression mediated by *MEF2A* and decreased *IRS1* tyrosine phosphorylation mediated by *DUSP1*.[20] Additionally, this mechanism can be linked to post-transplant diabetes mellitus (PTDM).

The calcineurin inhibitors tacrolimus (FK506) and cyclosporine A are essential immunosuppressive drugs for the clinical management of rejection in organ transplantation. While calcineurin is widely distributed [brain, heart, liver, kidney, pancreas, skeletal muscle (reviewed by Rusnak and Mertz[21])], the rationale for use in transplant rejection protocols has been the targeting of leukocyte-associated calcineurin as a means to suppress leukocyte function[22,23] and prolong graft survival. The hypothetical mechanism presented here suggests that inhibition of skeletal muscle calcineurin by these agents could impair glucose uptake and insulin signaling at the level of muscle and explains why immunosuppressive therapy with these inhibitors represents a significant independent risk factor for the development of PTDM,[24] which itself significantly compromises graft and patient survival.[25,26]

In order to assess bias in the model and determine if the mechanisms we generated are not simply artifacts of our particular causal model, we performed a series of negative con-

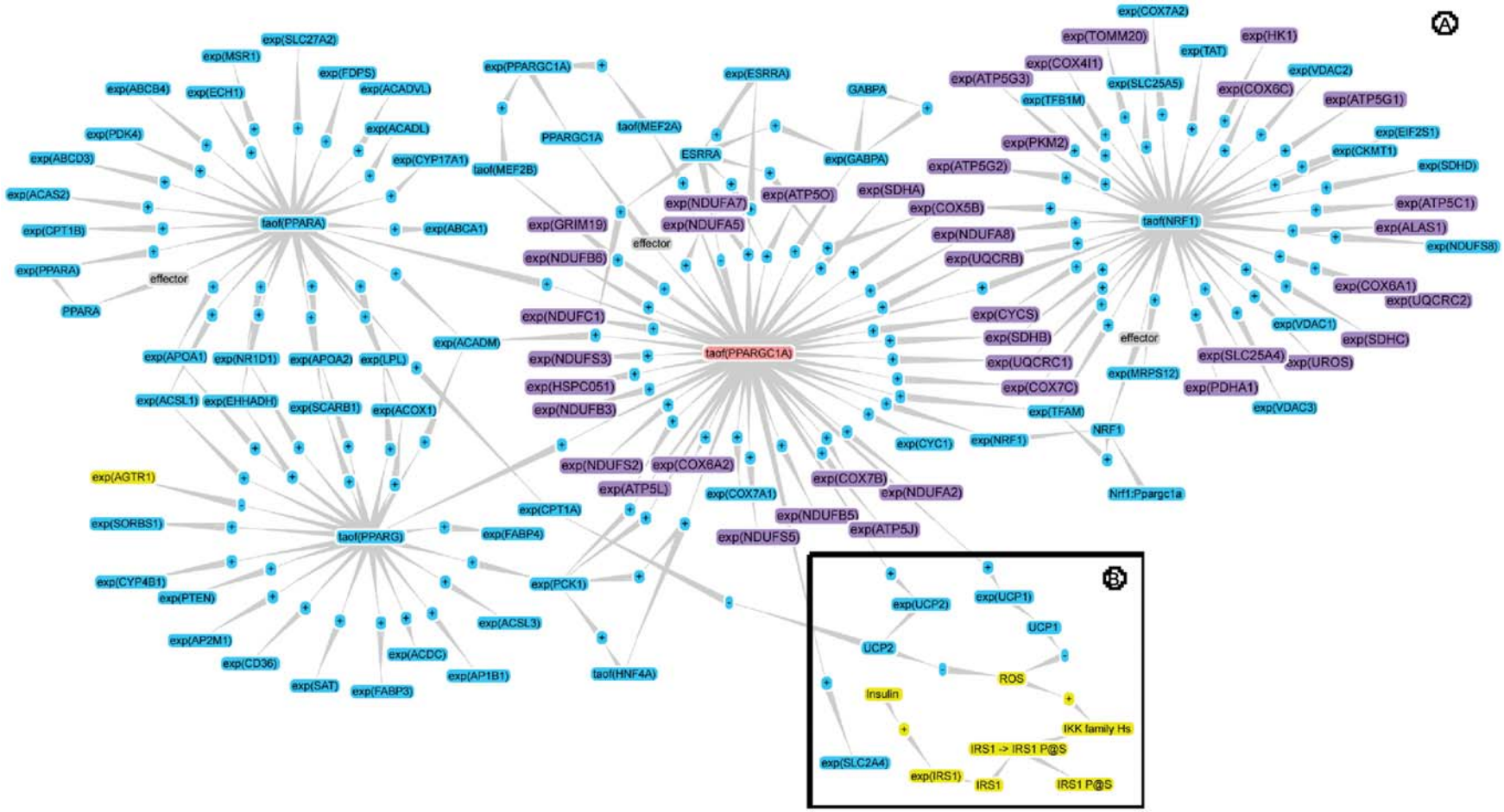

FIG. 5. **A:** Forward causal analysis on the muscle model following decreased *PPARGC1a* perturbation. Predicted increases are marked in yellow, and predicted decreases are marked in blue. The expression of genes observed to change in the studies is marked in purple. **B:** One of the many mechanisms that can link decreases in OXPHOS expression to decreases in insulin sensitivity and glucose import in muscle. See Appendix for gene abbreviations.

TABLE 2. HYPOTHETICAL REGULATORY MECHANISMS PRESENT IN BOTH STUDIES

Causal hypothesis[a]	*Change in activity or abundance*	*Predictions*[b]				P *values*[c]	
		Possible	*Correct*	*Contradiction*	*Conflict*	*Richness*	*Concordance*
taof(PPARGC1A)	Decrease	35	25	0	0	2.4E-06	9.8E-04
taof(NRF1)	Decrease	36	22	2	0	2.5E-02	1.9E-01
taof(MEF2 family Hs)	Decrease	71	47	2	0	1.6E-05	1.7E-03
paof(CalA-Calcineurin family Hs)	Decrease	62	49	3	0	3.3E-06	1.7E-03

[a]A causal hypothesis is a change in abundance or activity of a component in the model that would cause observable changes in other components in the model. toaf, transcriptional activity; paof, phosphatase activity. A set of causal hypotheses form the regulatory mechanism. See Appendix for gene abbreviations.

[b]Possible is the number of expression changes downstream of the causal hypothesis for which the model can make predictions. Correct is for a given causal hypothesis the number of predicted changes that matched the observed direction of change, whether increased, decreased, or unchanged. Contradiction is for a given causal hypothesis the number of predicted changes that do not match the observed direction of change. Conflict is for a given causal hypothesis the number of expression changes ambiguously determined from the model.

[c]Richness is for a given causal hypothesis the probability that the number of observed changes in a given model could have occurred by chance alone. Concordance is for a given causal hypothesis the probability that the number of observed changes that match the directionality of the changes predicted by the model could have occurred by chance alone.

trol causal analyses. We generated randomized expression profiles with similar model connectivity as our experimentally derived profiles and performed reverse causal analysis on these randomized profiles. We observed negligible overlap between the hypotheses generated from random and experimental profiles. We also noted that the scores that we observed

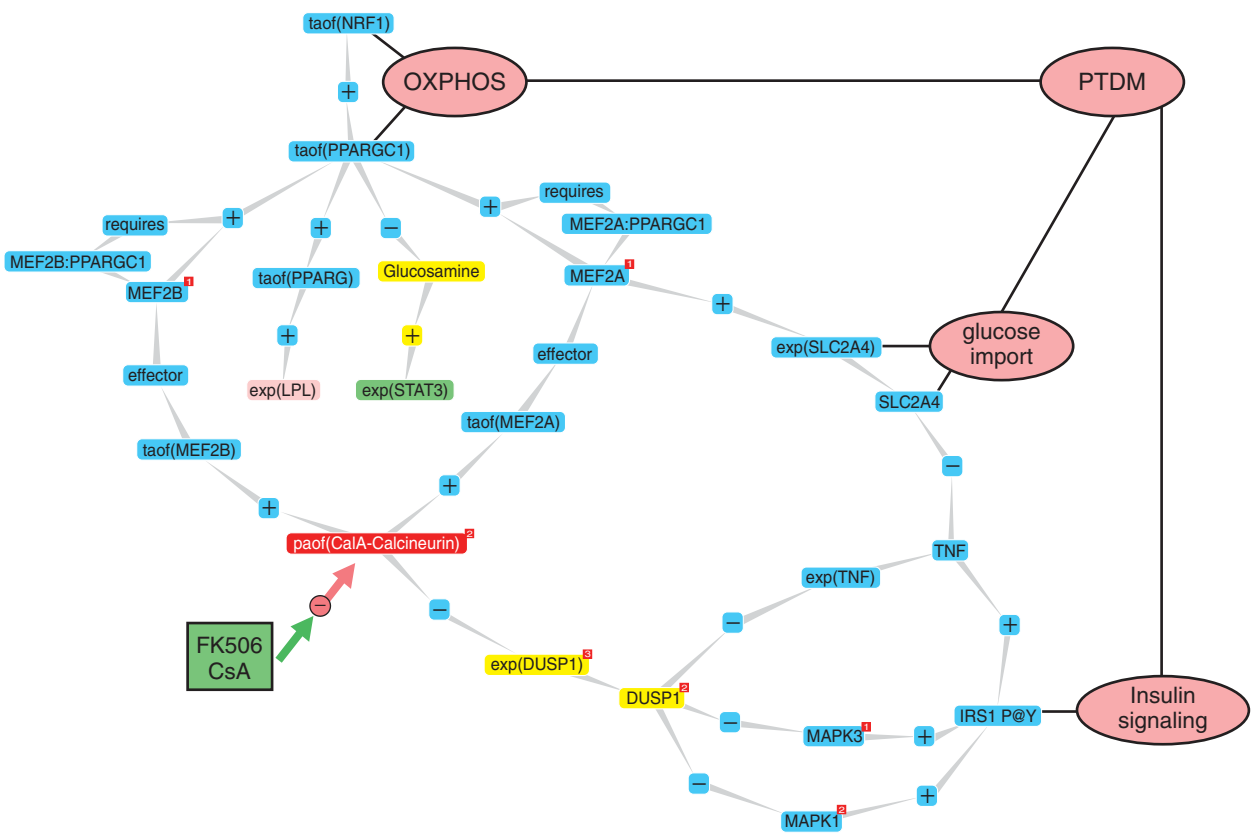

FIG. 6. Intersection of the causal analyses across the two studies revealed a common regulatory mechanism that can be linked to PTDM after treatment with the calcineurin inhibitor tacrolimus (FK506) or cyclosporine A (CsA). toaf, transcriptional activity; paof, phosphatase activity. See Appendix for gene abbreviations.

with our randomized profiles were much less significant than the scores from experimental profiles (data not shown).

CONCLUSIONS

In this paper, we describe a novel approach to define discrete, altered mechanisms of regulation in DM2. Our approach integrates genome-wide profiling measurements with system-level models of molecular cause-and-effect relationships and uses computer-aided causal analysis to define discrete mechanisms of regulation. Through causal analysis we showed evidence of discovery of those genes mechanistically linked to alteration in RNA profiles in diabetes, confirmed the coordinate down-regulation of genes involved in OXPHOS transcription, and further predicted glucose transport and insulin signaling outcomes. We also showed the utility of this model for hypothesis generation and discovery of a previously unrecognized mechanisms of regulation in DM2, as evidenced by the finding of skeletal muscle calcineurin expression linked to insulin sensitivity and glucose tolerance in PTDM. The emerging hypotheses describe biologic effects in DM2 and offer important cues for molecular targeted therapy. These findings show the utility of causal reasoning models for hypotheses generation and suggest a powerful new approach for interpretation of transcript profiling in drug discovery.

APPENDIX

Gene abbreviations

ABCB4	ATP-binding cassette, subfamily B (MDR/TAP), member 4
ABCD3	ATP-binding cassette, subfamily D (ALD), member 3
ACADM	acyl-CoA dehydrogenase, C-4 to C-12 straight chain
ACAS2	acetoacetyl-CoA synthetase
ACOX1	acyl-CoA oxidase 1, palmitoyl
ACSL1	acyl-CoA synthetase long-chain family member 1
ACSL3	acyl-CoA synthetase long-chain family member 3
AGTR1	angiotensin II receptor, type 1
AKT1	v-akt murine thymoma viral oncogene homolog 1
ALAS1	aminolevulinate, delta-, synthase 1
AP2M1	adaptor-related protein complex 2, mu 1 subunit
APOA1	apolipoprotein A-I
APOA2	apolipoprotein A-II
APOC3	apolipoprotein C-III
ATP5A1	ATP synthase, H^+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, cardiac muscle
ATP5C1	ATP synthase, H^+ transporting, mitochondrial F1 complex, gamma polypeptide 1
ATP5G1	ATP synthase, H^+ transporting, mitochondrial F0 complex, subunit c (subunit 9), isoform 1
ATP5G2	ATP synthase, H^+ transporting, mitochondrial F0 complex, subunit c (subunit 9), isoform 2
ATP5G3	ATP synthase, H^+ transporting, mitochondrial F0 complex, subunit c (subunit 9) isoform 3
ATP5L	ATP synthase, H^+ transporting, mitochondrial F0 complex, subunit g
ATP5O	ATP synthase, H^+ transporting, mitochondrial F1 complex, O subunit (oligomycin sensitivity-conferring protein)
CA3	carbonic anhydrase III, muscle specific
CBL	Cas-Br-M (murine) ecotropic retroviral transforming sequence

CCND1	cyclin D1 (PRAD1: parathyroid adenomatosis 1)
CD36	CD36 antigen (collagen type I receptor, thrombospondin receptor)
CDKN1A	cyclin-dependent kinase inhibitor 1A (p21, Cip1)
CKMT1	creatine kinase, mitochondrial 1 (ubiquitous)
COX4I1	cytochrome c oxidase subunit IV isoform 1
COX5B	cytochrome c oxidase subunit Vb
COX6A1	cytochrome c oxidase subunit VIa polypeptide 1
COX6A2	cytochrome c oxidase subunit VIa polypeptide 2
COX6C	cytochrome c oxidase subunit VIc
COX7A2	cytochrome c oxidase subunit VIIa polypeptide 2 (liver)
COX7B	cytochrome c oxidase subunit VIIb
COX7C	cytochrome c oxidase subunit VIIc
CPT1A	carnitine palmitoyltransferase 1A (liver)
CPT1B	carnitine palmitoyltransferase 1B (muscle)
CXCL1	chemokine (C-X-C motif) ligand 1
CYCS	cytochrome *c*, somatic
CYP4B1	cytochrome P450, family 4, subfamily B, polypeptide 1
DAZAP2	DAZ-associated protein 2
DUSP1	dual specificity phosphatase 1
E2F4	E2F transcription factor 4, p107/p130-binding
EDN1	endothelin 1
EHHADH	enoyl-CoA, hydratase/3–hydroxyacyl CoA dehydrogenase
EIF2S1	eukaryotic translation initiation factor 2, subunit 1 alpha, 35 kDa
ENPP1	ectonucleotide pyrophosphatase/phosphodiesterase 1
ESRRA	estrogen-related receptor alpha
ETS1	v-ets erythroblastosis virus E26 oncogene homolog 1 (avian)
ETS2	v-ets erythroblastosis virus E26 oncogene homolog 2 (avian)
FABP3	fatty acid binding protein 3, muscle and heart (mammary-derived growth inhibitor)
FABP4	fatty acid binding protein 4, adipocyte
FOS	v-fos FBJ murine osteosarcoma viral oncogene homolog
GABPA	GA binding protein transcription factor, alpha subunit 60 kDa
GRIM19	cell death-regulatory protein GRIM19
GSK3B	glycogen synthase kinase 3 beta
GYG	glycogenin
HK1	hexokinase 1
HNF4A	hepatocyte nuclear factor 4, alpha
HSPA1A	heat shock 70-kDa protein 1A
HSPC051	ubiquinol-cytochrome c reductase complex (7.2 kDa)
ICAM1	intercellular adhesion molecule 1 (CD54), human rhinovirus receptor
IFNG	interferon, gamma
IKK	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase family
IL6	interleukin 6 (interferon, beta 2)
IRS1	insulin receptor substrate 1
JAK1	Janus kinase 1 (a protein tyrosine kinase)
LDHB	lactate dehydrogenase B
LPL	lipoprotein lipase
MAPK1	mitogen-activated protein kinase 1
MCM6	MCM6 minichromosome maintenance deficient 6 (MIS5 homolog, *Schizosaccharomyces pombe*) (*Saccharomyces cerevisiae*)
MCM7	MCM7 minichromosome maintenance deficient 7 (*S. cerevisiae*)

MEF2C	MADS box transcription enhancer factor 2, polypeptide C (myocyte enhancer factor 2C)
MEK	mitogen-activated protein kinase kinase 1
MRPS12	mitochondrial ribosomal protein S12
MYC	v-myc myelocytomatosis viral oncogene homolog (avian)
NDUFA2	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 2, 8 kDa
NDUFA5	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 5, 13 kDa
NDUFA7	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 7, 14.5 kDa
NDUFA8	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 8, 19 kDa
NDUFB3	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 3, 12 kDa
NDUFB6	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 6, 17 kDa
NDUFC1	NADH dehydrogenase (ubiquinone) 1, subcomplex unknown, 1, 6 kDa
NDUFS2	NADH dehydrogenase (ubiquinone) Fe-S protein 2, 49 kDa (NADH-coenzyme Q reductase)
NDUFS3	NADH dehydrogenase (ubiquinone) Fe-S protein 3, 30 kDa (NADH-coenzyme Q reductase)
NDUFS8	NADH dehydrogenase (ubiquinone) Fe-S protein 8, 23 kDa (NADH-coenzyme Q reductase)
NFKB1	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (p105)
NR1D1	nuclear receptor subfamily 1, group D, member 1
NRF1	nuclear respiratory factor 1
PCK1	phosphoenolpyruvate carboxykinase 1 (soluble)
PDHA1	pyruvate dehydrogenase (lipoamide) alpha 1
PDK4	pyruvate dehydrogenase kinase, isoenzyme 4
PI3K	phosphatidylinositol 3-kinase
PKM2	pyruvate kinase, muscle
PPARA	peroxisome proliferative activated receptor, alpha
PPARG	peroxisome proliferative activated receptor, gamma
PPARGC1A	peroxisome proliferative activated receptor, gamma, coactivator 1, alpha
PRDX1	peroxiredoxin 1
PTEN	phosphatase and tensin homolog (mutated in multiple advanced cancers 1)
PTP1B	protein tyrosine phosphatase, non-receptor type 1b
RAD23A	RAD23 homolog A (*S. cerevisiae*)
RETN	resistin
ROS	reactive oxygenated species
SAT	spermidine/spermine N^1-acetyltransferase
SCARB1	scavenger receptor class B, member 1
SDHA	succinate dehydrogenase complex, subunit A, flavoprotein (Fp)
SDHB	succinate dehydrogenase complex, subunit B, iron sulfur (Ip)
SDHC	succinate dehydrogenase complex, subunit C, integral membrane protein, 15 kDa
SDHD	succinate dehydrogenase complex, subunit D, integral membrane protein
SLC25A4	solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 4
SLC25A5	solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 5
SLC2A4	solute carrier family 2 (facilitated glucose transporter), member 4
SOCS1	suppressor of cytokine signaling 1
SORBS1	sorbin and SH3 domain containing 1
STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)
TAT	tyrosine aminotransferase
TFAM	transcription factor A, mitochondrial

TFB1M	transcription factor B1, mitochondrial
TNF	tumor necrosis factor (TNF superfamily, member 2)
TNFRSF5	tumor necrosis factor receptor superfamily, member 5
TOMM20	translocase of outer mitochondrial membrane 20 homolog (yeast)
TXNIP	thioredoxin-interacting protein
UCP1	uncoupling protein 1 (mitochondrial, proton carrier)
UCP2	uncoupling protein 2 (mitochondrial, proton carrier)
UQCRB	ubiquinol-cytochrome c reductase binding protein
UQCRC1	ubiquinol-cytochrome c reductase core protein I
UQCRC2	ubiquinol-cytochrome c reductase core protein II
UROS	uroporphyrinogen III synthase (congenital erythropoietic porphyria)
VCAM1	vascular cell adhesion molecule 1
VCL	vinculin
VDAC1	voltage-dependent anion channel 1
VDAC2	voltage-dependent anion channel 2
VDAC3	voltage-dependent anion channel 3

ACKNOWLEDGMENTS

We would like to thank Dexter Pratt, Toby Segaran, Justin Sun, David Kightley, Ted Slater, and William Ladd for their contributions to this work.

REFERENCES

1. Ideker T, Ozier O, Schwikowski B, Siegel AF: Discovering regulatory and signalling circuits in molecular interaction networks. Bioinformatics 2002;18 (Suppl 1):S233–S240.
2. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Leher J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC: PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003;34:267–273.
3. Patti ME, Butte AJ, Crunkhorn S, Cusi K, Berria R, Kashyap S, Miyazaki Y, Kohane I, Costello M, Saccone R, Landaker EJ, Goldfine AB, Mun E, DeFronzo R, Finlayson J, Kahn CR, Mandarino LJ: Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes: Potential role of PGC1 and NRF1. Proc Natl Acad Sci USA 2003;100:8466–8471.
4. Eisen MB, Spellman PT, Brown PO, Botstein D: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 1998;95:14863–14868.
5. Tamayo P, Slonim D, Mesirov J, Zhu Q, Kitareewan S, Dmitrovsky E, Lander ES, Golub TR: Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. Proc Natl Acad Sci USA 1999;96: 2907–2912.
6. Raychaudhuri S, Stuart JM, Altman RB: Principal components analysis to summarize microarray experiments: application to sporulation time series. In: Altman RB, Dunker KA, Hunter L, Lauderdale K, Klein TE, eds. Biocomputing 2000: Proceedings of the Pacific Symposium. Singapore: World Scientific Publishing Co., 2000:455–466.
7. Butte AJ, Kohane IS: Mutual information relevance networks: Functional genomic clustering using pairwise entropy measurements. In: Altman RB, Dunker KA, Hunter L, Lauderdale K, Klein TE, eds. Biocomputing 2000: Proceedings of the Pacific Symposium. Singapore: World Scientific Publishing Co., 2000: 418–429.
8. Liang S, Fuhrman S, Somogyi R: Reveal, a general reverse engineering algorithm for inference of genetic network architectures. In: Altman RB, Dunker AK, Hunter L, Klein TE, eds. Biocomputing '98: Proceedings of the Pacific Symposium. Singapore: World Scientific Publishing Co., 1998:18–29.
9. D'Haeseleer P, Liang S, Somogyi R: Genetic network inference: From co-expression clustering to reverse engineering. Bioinformatics 2000;16:707–726.
10. Zak DE, Gonye GE, Schwaber JS, Doyle FJ 3rd: Importance of input perturbations and stochastic gene expression in the reverse engineering of genetic regulatory networks: Insights from an identifiability analysis of an in silico network. Genome Res 2003;13:2396–2405.
11. Pe'er D, Regev A, Elidan G, Friedman N: Inferring subnetworks from perturbed expression profiles. Bioinformatics 2001;17(Suppl 1):S215–S224.
12. Le PP, Bahl A, Ungar LH: Using prior knowledge to improve genetic network reconstruction from microarray data. In Silico Biol 2004;4(2):0027.

13. Hartemink AJ, Gifford DK, Jaakkola TS, Young RA: Using graphical models and genomic expression data to statistically validate models of genetic regulatory networks. In: Altman RB, Dunker KA, Hunter L, Lauderdale K, Klein TE, eds. Biocomputing 2001: Proceedings of the Pacific Symposium. Singapore: World Scientific Publishing Co., 2001:422–433.
14. Masys DR: Linking microarray data to the literature. Nat Genet 2001;28:9–10.
15. Jenssen TK, Laegreid A, Komorowski J, Hovig E: A literature network of human genes for high-throughput analysis of gene expression. Nat Genet 2001; 28:21–28.
16. Rice J, Stolovitzky G: Making the most of it: Pathways reconstruction and integrative simulation using the data at hand. BIOSILICO 2004;2:70–77.
17. Gentleman RR: A language for data analysis and graphics. J Computat Graph Stat 5:299–314.
18. Dudoit S, Gentleman RC, Quackenbush J: Open source software for the analysis of microarray data. Biotechniques 2003;Suppl:45–51.
19. Handschin C, Rhee J, Lin J, Tarr PT, Spiegelman BM: An autoregulatory loop controls peroxisome proliferator-activated receptor gamma coactivator 1alpha expression in muscle. Proc Natl Acad Sci USA 2003; 100:7111–7116.
20. Lim HW, New L, Han J, Molkentin JD: Calcineurin enhances MAPK phosphatase-1 expression and p38 MAPK inactivation in cardiac myocytes. J Biol Chem 2001;276:15913–15919.
21. Rusnak F, Mertz P. Calcineurin: Form and function. Physiol Rev 2000;80:1483–1521.
22. Koprak S, Matheravidathu S, Springer M, Gould S, Dumont FJ: Down-regulation of cell surface CXCR6 expression during T cell activation is predominantly mediated by calcineurin. Cell Immunol 2003;223:1–12.
23. Hartel C, Schumacher N, Fricke L, Ebel B, Kirchner H, Muller-Steinhardt M: Sensitivity of whole-blood T lymphocytes in individual patients to tacrolimus (FK 506): Impact of interleukin-2 mRNA expression as surrogate measure of immunosuppressive effect. Clin Chem 2004;50:141–151.
24. Heisel O, Heisel R, Balshaw R, Keown P: New onset diabetes mellitus in patients receiving calcineurin inhibitors: A systematic review and meta-analysis. Am J Transplant 2004;4:583–595.
25. Cosio FG, Pesavento TE, Kim S, Osei K, Henry M, Ferguson RM: Patient survival after renal transplantation: IV. Impact of post-transplant diabetes. Kidney Int 2002;62:1440–1446.
26. Kasiske BL, Snyder JJ, Gilbertson D, Matas AJ: Diabetes mellitus after kidney transplantation in the United States. Am J Transplant 2003;3:178–185.

Address reprint requests to:
Dr. Jack Pollard, Jr.
Project Lead, Metabolic Disorders
Genstruct
One Alewife Center, 1st Floor
Cambridge, MA 02140

E-mail: jpollard@genstruct.com