Targeted disruption of glycerol kinase gene in mice: expression analysis in liver shows alterations in network partners related to glycerol kinase activity

Nicole K. MacLennan¹, Lola Rahib⁴, Cynthia Shin¹, Zixing Fang²,³, Steve Horvath²,³, Jason Dean⁵, James C. Liao⁵, Edward R.B. McCabe¹,³ and Katrina M. Dipple¹,³,⁴,*

¹Department of Pediatrics, ²Department of Biostatistics and ³Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095, USA, ⁴Department of Biomedical Engineering and ⁵Department of Chemical and Biomolecular Engineering, Henry Samueli School of Engineering and Applied Sciences at UCLA, Los Angeles, CA 90095, USA

Received October 6, 2005; Revised and Accepted December 9, 2005

Glycerol kinase deficiency (GKD) is an X-linked inborn error of metabolism with metabolic and neurological crises. Liver shows the highest level of glycerol kinase (GK) activity in humans and mice. Absence of genotype–phenotype correlations in patients with GKD indicates the involvement of modifier genes, including other network partners. To understand the molecular pathogenesis of GKD, we performed microarray analysis on liver mRNA from neonatal glycerol kinase (Gyk) knockout (KO) and wild-type (WT) mice. Unsupervised learning revealed that the overall gene expression profile of the KO mice was different from that of WT. Real-time PCR confirmed the differences for selected genes. Functional gene enrichment analysis was used to find 56 increased and 37 decreased gene functional categories. PathwayAssist analysis identified changes in gene expression levels of genes involved in organic acid metabolism indicating that GK was part of the same metabolic network which correlates well with the patients with GKD having metabolic acidemia during their episodic crises. Network component analysis (NCA) showed that transcription factors sterol regulatory element-binding protein (SREBP)-1c, carbohydrate response element-binding protein (ChREBP), hepatocyte nuclear factor-4 alpha (HNF-4α) and peroxisome proliferative-activated receptor-alpha (PPARα) had increased activity in the Gyk KO mice compared with WT mice, whereas SREBP-2 was less active in the Gyk KO mice. These studies show that Gyk deletion causes alterations in expression of genes in several regulatory networks and is the first time NCA has been used to expand on microarray data from a mouse KO model of a human disease.

INTRODUCTION

Glycerol kinase (GK) phosphorylates glycerol to glycerol 3-phosphate (G-3P), which is a source for dihydroxyacetone phosphate (DHAP), glycerolipids, glucose, glycogen and protein (1). Thus, GK is a critical link to glycolysis, glycogenesis and gluconeogenesis. Mutations in the human GK gene on Xp21 cause glycerol kinase deficiency (GKD), an X-linked inborn error of metabolism with episodes of metabolic and neurological crises that include organic acidemia and hypoglycemia (1). GK is characterized by hyperglycerolemia and can be subdivided into three different categories: isolated GKD (iGKD) (symptomatic and asymptomatic GKD) and complex GKD (cGKD) (1). iGKD involves the GK locus only, whereas cGKD involves one or more additional loci around GK including Duchenne muscular dystrophy (DMD), dosage-sensitive sex reversal-AHC critical region on the chromosome gene 1 (DAXI), interleukin-1 receptor accessory protein-like, gene 1 (IL1RAPL1) and aristaeless-related homeobox gene (ARX) (1,2). All patients with GKD evidence hyperglycerolemia and glyceroluria. Patients with symptomatic iGKD or cGKD present with vomiting, lethargy and
progressive central nervous system deterioration in childhood associated with metabolic crises. Patients with asymptomatic iGKD present with pseudo-hypertriglyceridemia in adulthood. Treatment of GKD involves limitation of glyceraldehyde and therefore triglycerides in the diet and avoiding long periods of fasting (1).

We have previously shown that there is no genotype–phenotype correlation in iGKD in humans with GK mutation analysis incapable of predicting which individuals will have the symptomatic versus the asymptomatic form of iGKD (3). This lack of genotype–phenotype correlation makes GKD a complex disease (3). Metabolic flux through associated pathways with intermediates, such as G-3P and DHAP, and modifier genes, such as glycerol phosphate dehydrogenase (GPD) that influence flux in this metabolic network, may play a role in the complexity of the GKD phenotype (3–6).

Symptomatic iGKD is modeled in the glycerol kinase (Gyk) knockout (KO) mouse (7,8). Gyk KO males mice exhibit hyperglycerolemia, glyceroluria and abnormal fat metabolism (8). In addition, they become growth retarded and die on day of life (dol) 3–4 (8). Although the murine KO phenotype is more severe than human GKD, this KO mouse provides an opportunity to study the pathogenesis of this disease.

We have shown that the Gyk KO mouse is an excellent model for the iGKD metabolic phenotype (7). We confirmed the work of others (8) that Gyk KO mice have hyperglycerolemia and elevated plasma-free fatty acids and die by dol 3–4. We further demonstrated that KO male mice are acidelemia and elevated plasma-free fatty acids and die by dol 3–4 (8). Although the murine KO phenotype is more severe than human GKD, this KO mouse provides an opportunity to study the pathogenesis of this disease.

In this study, we further characterize this murine model of GKD and identify genes that have altered expression in livers of Gyk KO mice using microarray analysis. Our analysis may reveal which pathways play a role in the molecular pathogenesis of the human disease. Gyk KO perturbs expression profiles of gene pathways that are anticipated by our current knowledge of GK, such as organic acid metabolism/transport, lipid metabolism and glucose metabolism, as well as other functional groups that might not have been predicted, including apoptosis, steroid biosynthesis and cell-cycle arrest. We also investigated the applicability of network component analysis (NCA) to identify alterations in transcription factor activities (TFAs) in Gyk KO mice and demonstrate for the first time that NCA can be applied in mammalian model organisms.

RESULTS

Microarray analysis

Unsupervised learning analysis using the 4422 most varying probe sets (corresponding to 4387 genes) was used to determine whether the liver samples had globally distinct gene expression profiles. Hierarchical clustering (HC) and multi-dimensional scaling (MDS) of the most varying genes (Supplementary Material, Fig. S1) showed that the KO mice samples clustered together and separate from the WT mice, which also clustered together (Supplementary Material, Fig. S1A). MDS plots also revealed distinct clustering of the KO and WT samples indicating that the hepatic gene expression profiles for KO males were globally distinct from those of WT (Supplementary Material, Fig. S1B).

Gene filtering between control and experimental groups using the t-test and the fold-change criteria for gene list A (as described in Materials and Methods) identified 103 genes/110 probe sets that were differentially expressed in Gyk KO versus WT mouse livers. Supervised clustering revealed that 40 out of the 110 probe sets were up-regulated (red) in KO, whereas 70 out of the 110 probes were down-regulated (green) in KO (Fig. 1).

Functional grouping

Expression analysis systematic explorer (EASE) analysis was performed with genes generated from criteria for list B (as described in Materials and Methods) and revealed 477 genes/529 probesets differentially expressed between KO and WT mice. EASE analysis identified 56 increased (Fisher’s exact p < 0.05) and 37 decreased (p < 0.05) gene expression functional categories in the Gyk KO mice. The genes that were differentially expressed in the Gyk KO mouse livers included genes in the organic acid metabolism, fatty acid metabolism, lipid metabolism, carbohydrate/glucose metabolism, positive regulation of anti-apoptosis, glycerol ether metabolism and organic acid transport functional groups (Supplementary Material, Table S1). Expression levels of genes in the cell-cycle arrest, steroid dehydrogenase activity, G-3P metabolism, RAS protein signal transduction, regulation of cell cycle, lipid metabolism, glycerol ether metabolism, death receptor activity, water channel activity and cell death functional groups were decreased in the Gyk KO mouse livers (Supplementary Material, Table S2). Functional categories with subsets of genes that were both up-regulated and down-regulated in KO mice included lipid metabolism and glycerol ether metabolism (Supplementary Material, Tables S1 and S2).

Quantitative real-time-PCR analysis

Real-time PCR analysis confirmed that the levels of Gyk (13% of WT) and glycerol phosphate 2, mitochondrial (Gpd-2) (37% of WT) were significantly decreased in KO mice (Fig. 2). Peroxisome proliferator-activated receptor (PPAR)-γ co-activating factor-1 (Pgc-I) (393% of WT), solute carrier family-7a, member-2 (Slc-7a2) (281% of WT), tyrosine aminotransferase (Tat) (374% of WT), triosephosphate isomerase-1 (Tpi-I) (231% of WT), lipoprotein lipase (Lpl) (245% of WT), hepatocyte nuclear factor-4 alpha (Hnf-4α) (263% of WT), phosphoglycerate mutase-1 (Pgam-1) (292% of WT), aldolase-1A, fructose-bisphosphate (AldolA) (262% of WT) and glucose phosphate isomerase-1 (Gpi-1) (197% of WT) were significantly increased in Gyk KO compared with WT mice (Fig. 2). Expression levels (Supplementary Material, Table S3) of these genes from GeneChip analysis correlated significantly with the real-time PCR analysis (with a Spearman’s rank correlation test, rs = 0.86, p < 0.002).

Analysis of Affymetrix Gyk probe sets

Although the three Gyk probe sets on the mouse 430 2.0 GeneChip were thought to be specific for the X-chromosome
Figure 1. HC Analysis (dChip program, 3) of significantly differentially expressed 103 genes/110 gene probesets from WT ($n = 4$) versus KO ($n = 4$) samples filtered by 2-fold change in gene expression, $p < 0.05$ and percent of present call of at least 20%. Each row represents the indicated gene. Each column corresponds to the experimental sample as listed at the top. Red indicates an up-regulation of gene expression (40 probe sets) and green indicates a down-regulation of expression (70 probe sets) in Gyk KO mice.
isoform, probe set [1445242_at] was only 3-fold decreased which was less than expected for ablation of the gene and less than the other two Gyk probe sets: [1422704_at] and [1422703_at], which had 15- and 16-fold decreases, respectively (Supplementary Material, Fig. S2). Therefore, the three Gyk probe sets were analyzed to determine the possible sequence overlap with other genes using BLAST-like alignment tool (BLAT) (9). BLAT scores of less than 30 were obtained for both the [1422703_at] probe set, designed from a 536 bp segment in the middle of Gyk and the [1422704_at] probe set, designed from a 413 bp region at the 3′ end of Gyk. Significant DNA sequence scores of 80–95 were found for probe set [1445242_at] which is based on a 538 bp region at the 5′ end of the Gyk gene. The three X-chromosome Gyk probe sets with their corresponding probe set oligonucleotides and the approximate location on the X-chromosome Gyk gene are graphically depicted in Supplementary Material, Figure S2. When the probe set sequences were submitted to BLAT, only probe set [1445242_at] had a region (62 bp) with significant sequence similar to other mouse sequences. A BLAST search of probe set [1445242_at] identified the two Gyk accession numbers from which all three probe sets were designed as well as unidentified mouse chromosomal clones, hypothetical glycaryltransferases group 1 gene and a novel protein C6orf32 (also identified as KIAA0386 and DIFF48). However, a BLAST search of probe sets [1422703_at] and [1422704_at] resulted in identification of Gyk isoform sequences only.

Pathway and network analysis

PathwayAssist analysis was used to model the network between Gyk and other metabolic pathways including organic acid and glycerol metabolism (Fig. 3). Gyk is shown as a triangle at the top of the diagram. Genes involved in organic acid metabolism that were differentially expressed in Gyk KO compared with WT mice by microarray analysis include cystathionine-beta-synthase (CBS), carnitine acetyltransferase (CRAT), cystathioninase (CTH), enoyl-coenzyme A hydratase/3-hydroxyacyl-coenzyme A dehydrogenase (EHHADH), lipoprotein lipase (LPL), N-acetylglutamate synthase (NAGS), solute carrier family 7a, member 2 (SLC7A2), tyrosine aminotransferase (TAT) and triosephosphate isomerase 1 (TPI1) (Fig. 3, ovals). Transcription factors (TF) identified by NCA that involve Gyk in regulation of the metabolic network include carbohydrate response element-binding protein (ChREBP), hepatocyte nuclear factor-4 alpha (HNF-4α), peroxisome proliferative-activated receptor-alpha (PPAR-α), sterol regulatory element-binding protein (SREBP)-1c and SREBP-2 (Fig. 3, rectangles). The shortest links to other pathways from Gyk and NCA TFs identified by PathwayAssist and highlighted by diamonds include angiotensinogen (AGT), glycerol-3-phosphate dehydrogenase 2 (mitochondrial) (GPD2), nuclear receptor subfamily 1, group H, member 4 (NR1H4), p21 (CDKN1A)-activated kinase 1 (PAK1), PPAR gamma coactivator 1, alpha (PPARGC1α), PPAR-γ, glucagon (GCG), Sp1 TF (SP1) and tumor necrosis factor (TNF) (Fig. 3). Cell processes involved in these connections include amino acid transport, cell survival, glycolysis,
oxidative phosphorylation, pathogenesis, proliferation, proteolysis, regulation of signal transduction and secretion (Fig. 3, hexagons). This pathway analysis illustrated links among the organic acid metabolism genes, TFs and Gyk and reveals the potential significant network effects of Gyk ablation.

Network component analysis
NCA was used to deduce quantitatively TFAs of SREBP-1c, SREBP-2, ChREBP, HNF-4α and PPAR-α from Gyk KO and WT mouse liver microarray data. NCA revealed a positive TFA for SREBP-1c, ChREBP, HNF-4α, and PPAR-α, indicating that these TFs were in a more activated state in the Gyk KO mouse and thus increased the transcription of the genes they regulate more than in the WT mice (Fig. 4). SREBP-2 had a negative TFA indicating that it had stronger transcriptional activity in the WT mice (Fig. 4).

DISCUSSION
We examined expression levels of approximately 45,000 genes from Gyk KO and WT mouse livers to develop a better understanding of GKD and to identify other molecular networks that are important to GKD pathogenesis. HC involving the 4387 most varying genes shows that KO samples group together and are distinct from WT samples indicating that Gyk KO mice have globally distinct gene expression. Therefore, decreasing Gyk gene expression levels has global transcriptional and presumably proteomic and metabolomic effects.

Interestingly, the expression levels corresponding to the three different Affymetrix Gyk probe sets varied from 3- to 16-fold decreases. Probe set [1445242_at] showed a 3-fold decrease in Gyk gene expression. This probe set contained 62 bp of DNA with significant sequence homology to conserved sequence from several mouse chromosomes as well as hypothetical glycosyl transferases group 1 and a novel protein similar to C6orf32, KIAA0386 and DIFF48 (Supplementary Material, Fig. S2). This sequence homology in the 5' end of the Gyk gene explains why there was less perturbation of this Gyk probe set compared with the other two Gyk probe sets. This finding is consistent with other groups who have noted that microarray data are complicated by cross-hybridization of short contiguous sequences and highlights

Figure 3. PathwayAssist analysis. Analysis of links between GK and other metabolic pathways including organic acid metabolism and glycerol metabolism. Gyk is represented by a triangle. Organic acid genes are represented by ovals. TFs are represented by rectangles near the bottom of the diagram. The shortest links to other pathways from Gyk and NCA analysis TFs are represented by diamonds. Cell processes are represented by hexagons. CBS, CRAT, CTH, EHHADH, LPL, NAGS, SLC7A2, TAT and TP1, ChREBP, HNF-4α, PPARα, SREBP-1c, SREBP-2, AGT, GPD2, nuclear receptor subfamily 1, group H, member 4 (NR1H4), p21 (CDKN1A)-activated kinase 1 (PAK1), PPARGC1a, PPARγ, GCG, SP1, TNF.
the importance of using more than one probeset to determine the level of expression of a gene (10–12).

RT–PCR confirmed that Gyk KO mice have increased levels of Tat and decreased levels of Gyk (8). Gpd2 (mitochondrial), the next enzyme in glycerol metabolism and a modifier of Gyk (13), was significantly decreased. Mouse FAD-linked Gpd can oxidize G-3P to DHAP (14). Gpd-2 KO mice have hyperglycerolemia, glyceroluria, hypoglycemia, elevated plasma-free fatty acids, ketonuria and decreased Ucp-1, and become growth-retarded and die within a week of post-natal life (15). This phenotype is strikingly similar to the Gyk KO mouse phenotype suggesting common physiological and network partners.

EASE analysis was used to find enriched biological groups and functions within the differentially expressed genes and identified 56 increased ($p < 0.05$) and 37 decreased ($p < 0.05$) gene functional categories. These groups represent biological themes and associated networks and suggest that proteomic networks are globally disrupted as a result of Gyk ablation. Glycerol/glycerol ether metabolism, cell death, cell-cycle arrest, steroid dehydrogenase activity and organic acid metabolism are among the over-represented functional categories in Gyk KO mice.

Genes involved in organic acid metabolism and organic acid transport were significantly up-regulated in the Gyk KO mouse. Of these, Crat is a key metabolic enzyme in mitochondria, peroxisomes and endoplasmic reticulum, as it catalyzes the reversible transfer of acyl groups from an acyl-CoA thioester to carnitine and regulates the ratio of acyl-CoA/CoA in these subcellular compartments (16). LPL has the dual functions of triglyceride hydrolase and a ligand for receptor-mediated lipoprotein uptake. Murine NAGS is a mitochondrial enzyme that catalyzes the formation of N-acetylglutamate, an essential allosteric activator of carbamoyl phosphate synthase I, the first enzyme in the urea cycle (17), and there is an association between elevated levels of dietary protein and increased synthesis of NAGS (18). Elevated expression of NAGS could indicate that Gyk KO mice rely on alternative sources of energy, specifically protein catabolism. Cytochrome P450, family 39, subfamily a, polypeptide 1 (Cyp39a1) is an endoplasmic reticulum protein involved in the conversion of cholesterol to bile acids, and its substrates include the oxysterols 25-hydroxycholesterol, 27-hydroxycholesterol and 24-hydroxycholesterol (19). Altered expression of these organic acid functional group genes in Gyk KO mice illustrates the broad spectrum of metabolic networks affected by Gyk ablation.

Genes that are up-regulated in the glycerol ether metabolic functional group in Gyk KO mice include microsomal triglyceride transfer protein (Mttp) and phosphatidic acid phosphatase type 2A (Ppap2a). Mttp is expressed in the murine liver and is required for apoB secretion in liver cells (21). Increased gene expression of Mttp in Gyk KO mice suggests that these mice have increased levels of apoB and may be compensating for the lack of Gyk. Ppap2a, part of the lipid phosphate phosphatase (LPP) family, converts phosphatidic acid to diacylglycerol and functions in de novo synthesis of glycerolipids, although all of the substrates for LPP are not known (22,23). This integral membrane glycoprotein is a surface enzyme that plays a role in the hydrolysis and uptake of lipids from extracellular space (24). Therefore, Ppap2a may be facilitating uptake of glycerol lipids and attenuating the hyperglycerolemia in Gyk KO mice.

Gyk, Gpd2, and monoaoylglycerol O-acyltransferase (Mogat-2) are down-regulated in the glycerol ether
metabolism category. With decreased Gpd2, KO mice would not likely benefit from the glycerol-phosphate-shuttle for energy generation in the liver (1,25,26), which may contribute to the failure to thrive in the Gym KO mice. Murine Mogat-2 is involved in dietary fat uptake, lipid synthesis and storage (27). Mogat-2 possesses acyl-CoA:diacylglycerol acyltransferase (Dgat) activity which could provide an additional pathway for triacylglycerol synthesis (28). Furthermore, Dgat has been shown to be a modifier of Gym (Martinez and McCabe, unpublished data), and we speculate that Mogat-2 may have a modifier role in glycerol metabolism due to its Dgat enzymatic activity. It also is intriguing that Mgtat2 is down-regulated and Ppap2a is up-regulated, consistent with a more focussed role for Mgtat2 on glycerol lipid metabolism (triacylglycerol synthesis) and coordinated dysregulation of enzymes involved in the glycerol metabolic pathway. One interpretation of our data on the glycerol ether metabolic group is that with Gym deleted, the KO mice do not have the need to expend energy on the synthesis of high levels of the enzymes that are so closely related in the same metabolic network. This would also suggest that Gym may be a critical node for this part of the network (6).

It is particularly interesting that expression levels of genes involved in regulation of cell cycle are significantly decreased. These genes included cyclin D1 (Ccnid1), which is involved in promoting cell-cycle G1/S transitions, and cyclin A2 (Ccna2), which is involved in promoting G1/S and G2/M transitions (29). Cyclin-dependent kinase inhibitor 2C (Cldkn2c) functions as a cell growth regulator that controls cell-cycle G1 progression (30). Growth arrest and DNA-damage-inducible 45 alpha (Gadd45a) is a member of a group of genes with transcript levels increased, following the stressful growth-arrest conditions and an activity that involves mediating activation of the p38/JNK pathway via MTK1/MEKK4 kinase (31). Signal transducer and activator of transcription 1 (Stat1) mediates the expression of a variety of genes that are thought to be important for cell viability in response to different cell stimuli and pathogens (32). Cullins are the major components of a series of ubiquitin ligases that control the degradation of a broad range of proteins (33,34). It is intriguing that cell-cycle genes are down-regulated at the time when the mice have growth retardation and may indicate a role for these cell-cycle proteins in GKD pathogenesis. These results are also intriguing given our recent finding that GK binds to histones and may play a role in apoptosis (13).

Aquaporin (Aqp) activity is differentially altered in Gym KO mice. Aqp1 and 8 gene expression is down-regulated, whereas Aqp3 is up-regulated. Aqp3 is glycerol-permeable (35) and Aqp3 gene expression may be elevated over water channel Aqp1 and 8 gene expression to ameliorate the hyperglycerolemia in Gym KO mice.

PathwayAssist software analysis provided indirect links between GK and proteins including organic acid metabolism, TFs and other critical metabolic pathways and biological processes. The involvement of these transcripts, proteins and small molecules identified by PathwayAssist software and our expression analyses indicates a broader involvement of Gym in biological networks than we would have predicted. Some of these networks may suggest as yet unrecognized ‘moonlighting’ functions of Gym (36).

In addition to its enzymatic role of converting glycerol into G-3P, Gym has moonlighting activities (36) including interaction with the voltage-dependent anion channel, a possible role in apoptosis, involvement in binding to and nuclear uptake of activated, liganded glucocorticoid-receptor complex, binding to nuclear histones and stabilization of lysosomes (1,13,37–43). Disruption of the enzymatic and moonlighting activities as well as network links to other pathways could explain the global effect of Gym ablation, resulting in death at 3–4 dol. For example, the GK promoter contains a functional binding site for HNF-4α, suggesting that HNF-4α is a modifier of Gym (44). Furthermore, both GK and HNF-4α are implicated in dysregulation of glucose metabolism and diabetes (44–46). These links may provide additional insights to the role of GK in insulin sensitivity.

Ablation of Gym expression has a broad ranging effect on gene expression. Alterations in gene expression of organic acid metabolism are most notable because Gym KO mice and human patients with symptomatic GKD have metabolic acidemia. We have noted previously that if different proteins are involved in the pathogenesis of the same disorder, then they are involved in the same pathogenic network, even if the connection may not be obvious (47). These links to organic acid metabolism may provide additional insights to the metabolic acidosis seen in patients.

We applied NCA (48) to these Gym KO liver expression data and this is the first application of NCA to an investigation of mammalian tissues. Unlike principal component analysis, which assumes orthogonality of the components, NCA uncovers hidden regulatory signals from outputs of networked systems based on TF-promoter interaction information. The TFA signal from NCA cannot be captured in TF expression values (microarray data) alone, as many TFs require post-translational modification or ligand-binding for activation or interaction with different TFs under different conditions leading to variable TF strengths. As individual genes are commonly regulated by multiple TFs, direct interpretation from gene expression data alone cannot deconvolute the simultaneous contribution of various TFs. NCA was designed to achieve such a task based on specific mathematical properties satisfied by the network and has previously been used to study Escherichia coli carbon source transition from glucose to acetate (49) and yeast cell-cycle regulators (50). To date, NCA has only been used in bacteria and yeast.

NCA was used to determine quantitatively the effects of a Gym KO on five key transcriptional regulators of sugar and fatty acid metabolism and revealed that SREBP-1c, ChREBP, HNF-4α and PPAR-α were activated in the Gym KO mice compared with the WT mice. In contrast, SREBP-2 was less activated in the KO mice. Interestingly, the microarray data revealed that HNF-4α was increased 2-fold and that there was no significant difference in the expression levels of SREBP-2, ChREBP, PPAR-γ or SREBP-1c. The discrepancy of the TFA of these TFs and their gene expression levels shows that it is not expression level alone that is important for the TFs to be active and highlights the importance of NCA of microarray data to elucidate transcriptional networks. The increased activity of SREBP-1c, HNF-4α and PPAR-α reveals that regulatory networks in the Gym KO mouse...
liver are compensating for Gyk deletion by increasing the transcription of liver-specific genes involved in carbohydrate and fat metabolism. ChREBP is a known activator of pyruvate kinase and lipogenic genes (51,52), and HNF4-α has been shown to regulate GK (44). SREBP-1c is a ligand-activated TF and established transcriptional activator of genes involved in lipid synthesis (53). Interestingly, SREBP-2, another SREBP isoform that is a key regulator of cholesterol synthesis, had a negative TFA in the KO mice suggesting that it has less transcriptional activity and did not increase the expression of genes under its control. This decrease in SREBP-2 TFA is particularly interesting given the up-regulation of Cyp39a1.

The biological significance of these NCA results is understandable. PPAR-α is a ligand-activated regulator of glycerol metabolism and was found to have a positive TFA. PPAR-α has been shown to induce expression of Aqp3, a hepatic glycerol transporter, and Gpd1l (54) and Modl (55), both participants in the conversion of glycerol to glucose. We found Aqp3, Gpd1l and Modl to be up-regulated in the microarray analysis and thus are consistent with this NCA result. We can interpret a positive PPAR-α TFA as a biological response to ameliorate the hyperglycerolemia in Gyk KO mice. HNF-4α is a key mediator in liver lipid homeostasis and also a regulator of GK (44,46). Two HNF-4α target genes, Acox1 and Mttp, were up-regulated again consistent with this NCA result. Acox1 is involved in lipid and fatty acid metabolism, whereas Mttp is essential for assembly of lipoproteins. G-3P represents a branch point between gluconeogenesis and lipid synthesis, and thus we would predict that TFs such as HNF-4α involved in hepatic lipid and carbohydrate metabolism will be significantly perturbed in Gyk KO. This prediction is supported by the positive TFA found for HNF-4α. These results demonstrate the applicability of NCA to provide increased information from and understanding of high throughput microarray expression data.

In conclusion, Gyk KO mouse hepatic mRNA levels are altered in a broad range of functional categories, some of which are obviously relevant to the biochemical aberrations in human GKD. Investigations of altered gene expression also allow the opportunity to identify modifier genes in this GKD murine model that will provide valuable insights into GKD pathogenesis, the lack of genotype–phenotype correlation and possible novel therapeutic interventions all of which can be investigated further in model organisms.

MATERIALS AND METHODS

Animal model

Gyk KO mice were provided by the courtesy of W.J. Craigen, Baylor College of Medicine (8). The mice were generated using 129/SvJ embryonic stem cells and bred on a C57B1/6J mouse (8). At UCLA, the mice were inbred. The WT controls are male littermates of the KO male mice.

Mice were genotyped post-sacrifice (7). Male WT and KO mouse pups were sacrificed on dol 3 and each liver was harvested, placed in TRizol Reagent (Invitrogen Life Technologies, Carlsbad, CA, USA), quickly frozen in liquid nitrogen and stored at −80°C until further use. Dol 3 was chosen because the mice are phenotypically symptomatic with statistically different parameters for hypoglycemia, acidosis; low bicarbonate and decreased base excess (7). On dol 2, they are not significantly different from WT in all of these important clinical phenotypes.

RNA isolation

Total RNA from four KO and four WT livers was isolated individually using TRizol Reagent (Invitrogen Life Technologies) and purified using RNeasy MiniElute Cleanup Kit (Qiagen Inc., Chatsworth, CA, USA).

cDNA synthesis and hybridization

Eluted total RNAs were quantified and adjusted to a final concentration of 1.25 µg/µl. Total RNA quality was assessed using an RNA Lab-On-A-Chip (Caliper Technologies Corp., Mountain View, CA, USA) and Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA). cDNA was synthesized from the poly(A)⁺ mRNA present in the isolated total RNA using the SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen Life Technologies) and poly(T)-nucleotide primers. Biotin-tagged cRNA was generated using the BioArray High-Yield RNA Transcript Labeling Kit (T7) (Enzo Diagnostics, Inc., Farmingdale, NY, USA) and fragmented to an average strand length of 100 bases (range 35–200 bases) following prescribed protocols (Affymetrix GeneChip Expression Analysis Technical Manual). Ten micrograms of cRNA was hybridized at 45°C with rotation for 16 h (Affymetrix GeneChip Hybridization Oven 640) to an Affymetrix Mouse Genome 430 2.0 GeneChip as per the manufacturer’s instructions (Affymetrix, Inc.). The GeneChip arrays were washed and then stained with streptavidin–phycoerythrin on an Affymetrix Fluidics Station 450, followed by scanning on a GeneChipScanner 3000.

MicroArray analysis

Results were quantified and analyzed with GeneChip Operating Software (GCOS) 1.2 software (Affymetrix, Inc.) using default values (scaling, target signal intensity = 500; normalization, all probe sets; parameters, all set at default values). We used Affymetrix GeneChip Mouse Expression set 430 2.0 Array which uses mouse maintenance genes for normalization. Details are available on the Affymetrix website (www. affymetrix.com/support/technical/mask_files.affx). The GCOS data were analyzed with dChip (56), EASE (NIH) (57) and PathwayAssist (Stratagene, La Jolla, CA, USA) computer programs. All microarray data were analyzed using the dChip (56) or R software (58). GEO accession number GSE 3843.

Hierarchical clustering

We used average linkage HC with the Euclidean distance to cluster the microarray samples on the basis of their gene expression profiles. To eliminate noise, the most highly varying genes were selected by restricting the unsupervised analysis to genes with coefficients of variance between 0.34 and 2 and a present percent call of ≥20%. This resulted in 4387 genes, which will be referred to as the most varying
genes. Microarrays (columns) were standardized for sample clustering. The same gene set was used to visualize the samples in a non-metric MDS plot as implemented in the R function isoMDS. The closer the two samples were in the MDS plot, the more similar were their gene expression profiles according to the Euclidean distance of the standardized expression profiles. The WT and the KO samples each clustered together, which indicated that these groups had globally distinct gene expression profiles.

Gene filtering

We used both a $t$-test and a fold-change criterion to screen for differentially expressed genes between the WT and KO mice. Specifically, we used the following criteria for gene list A: (i) two sample $t$-test, $p < 0.05$, (ii) absolute value of fold change $> 2.0$, (iii) absolute difference between baseline (WT) and experimental (KO) $> 100$ and (iv) percent of present call in samples $\geq 20\%$ to select differentially expressed genes. This resulted in 103 genes/110 probe sets differentially expressed and used for the HC dendogram. For gene list B, we used slightly less stringent criteria: (i) two sample $t$-test, $p < 0.05$, (ii) absolute value of fold change $> 1.4$, (iii) absolute difference between baseline (WT) and experimental (KO) $> 100$ and (iv) percent of present call in samples $\geq 20\%$ to select differentially expressed genes. This resulted in 477 genes/529 probe sets differentially expressed and used for EASE analysis.

Gene grouping/annotation

The EASE analysis (57) was used to find functional categories that were over-represented or enriched in the filtered gene list. The analysis was performed by using the over-representation function and the one-tailed Fisher’s exact probability for over-representation. EASE facilitates the biological interpretation of gene lists derived from the results of microarray, proteomic and SAGE experiments. EASE provides statistical methods for discovering enriched biological themes within gene lists and generates gene annotation tables.

PathwayAssist analysis

Differentially expressed genes in the microarray data were analyzed using PathwayAssist analysis software (version 3.0, Stratagene). The Pathway was built by searching for connections of genes with common regulators and finding the shortest paths between nodes. PathwayAssist pathway analysis software helps to interpret experimental results in the context of pathways, gene regulation networks and protein interaction maps. Using curated and automatically created databases, PathwayAssist identifies relationships among genes, small molecules, cell objects and processes, builds networks and creates pathway diagrams.

Quantitative real-time-PCR analysis

cDNA was synthesized with Superscript III (Invitrogen Life Technologies) from DNA-free Dnase I (Ambion Inc.)-treated RNA. Target probes were labeled with a fluorescent reporter dye, and primers were designed using Primer Express (Applied Biosystems Inc.) and synthesized at Eurogentec (San Diego, CA, USA). Quantitative real-time-PCR (QRT-PCR) was performed using 18S and $\beta$-actin endogenous controls with dye emission detected by the ABI Prism 7700 Sequence Detection System (Applied Biosystems Inc.).

Network component analysis

NCA determines the TFA of a given activated TF by evaluation of the expression level of genes known to be affected by that TF (48). NCA was used to identify the effects of a Gyk KO on TFs related to glycolysis and lipid homeostasis. Mathematically, the TFA is defined as $\log(TFA-\text{KO}/TFA-\text{WT})$, and therefore a positive value means the corresponding TF transmits a stronger signal in the Gyk KO than the WT mouse, whereas a negative value has the opposite meaning. The analysis was performed by constructing a connectivity matrix encompassing the regulatory network of five TFs (SREBP-1c, SREBP-2, ChREBP, HNF-4$\alpha$ and PPAR-$\alpha$) and 89 genes, including those from glycolysis, fatty acid metabolism and gluconeogenesis. Gene–TF interactions were compiled from the literature and are reported in the Supplementary Material. All of the analyses were performed with Matlab software modules (The Mathworks, Natick, MA, USA).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

We wish to thank William J. Craigen, M.D., Ph.D. for providing us with the Gyk KO mice used to establish our colony. This work was supported by NIH grants K08 DK60055 (K.M.D.), R01 GM67929 (K.M.D.) and R01 HD22563 (E.M.C.).

Conflict of Interest statement. The authors have no conflicts of interest.

REFERENCES


