Supporting Information

Shaw et al. 10.1073/pnas.1016789108

SI Methods

Compound Set Enrichment Analysis (CSEA). CSEA begins with a ranked list of small molecules, in our case ranked by signal-to-noise score (S2N), but in general other metrics can be used that reflect the correlation between the compound's cellular phenotype and the class distinction between mutant and wild-type cell lines. Given a prespecified set of compounds *S* (defined by a shared attribute; e.g., acting in the same metabolic pathway or belonging to the same drug class), CSEA asks if members of set *S* are randomly distributed throughout the ranked list or are enriched at the top or bottom (as would be expected if members of set *S* can discriminate between mutant and wild-type classes). An enrichment score (ES) is calculated by walking down the ranked list and increasing a running-sum statistic whenever a member of set *S* is encountered and decreasing the running-sum statistic whenever a compound that is not in set *S* is encountered. The greater the correlation between the effect of a compound in set *S* and the mutant vs. wild-type class distinction (i.e,. the greater the absolute magnitude of S2N), the greater the increase in the running-sum statistic. The ES is defined as the greatest deviation from zero achieved by the running-sum statistic while walking through the entire ranked list; the ES is a weighted Kolmogorov–Smirnov-like statistic. The normalized ES is derived from the ES and adjusts for variation in compound set size (1).

Lymphoblast Cell Line (LCL) Characterization, Small-Molecule Screen, and ATP Assay. All LCLs were subjected to DNA sequencing to confirm the presence of the heterozygous C-to-T substitution in codon 268 exclusively in individuals with documented maturity onset diabetes of the young type 1 and not in unaffected individuals. Genomic DNA was purified using the Gentra Puregene DNA Purification Kit (Qiagen). A 315-bp fragment was amplified from exon 7 using the primers: 5′-GCA CCA GCT ATC TTG CCA AC -3′ (forward) and 5′-AGG AGA AGT CTG GCA GAG CG -3′ (reverse), confirmed by agarose gel electrophoresis, and purified by Qiaquick PCR purification kit (Qiagen). Sequencing primers were as follows: 5′- ACT AGA GGA GAG GGG TCA AC-3′ (forward) and 5′-CGT TCT GGA GAG AGA GTC AG-3′ (reverse). HNF4α expression was assessed by quantitative RT-PCR (incorporating "no RT" and "no template" controls) using a commercially available TaqMan primer/probe set that spans the exon 3/exon 4 boundary (Applied Biosystems Hs00230853_m1) and a control GAPDH TaqMan primer/set (Applied Biosystems Hs02786624_g1) according to manufacturer instructions. RNA was isolated from LCLs using Trizol (Invitrogen), and cDNA was prepared using the MessageSensor RT Kit (Ambion) according to manufacturer instructions. Relative EBV copy number was assessed by quantitative RT-PCR using a 66-bp fragment at the EBV DNA polymerase locus and an RNase P control (Applied Biosystems # 4316844) using previously published protocols (2). LCL growth rates were determined by counting growth-phase LCLs in triplicate for 4–5 d as previously published (2).

shRNA Knockdown in Murine β-Cells. Murine β-cells (MIN6 cells) were cultured in DMEM containing: 25 mM glucose, supplemented with 15% FBS (ATCC), penicillin (50 international units/mL), streptomycin (50μg/mL), and 27.5 μM β-mercaptoethanol under humidified conditions of 5% CO2 and 95% air at 37 °C. Upon stable infection with shRNA lentiviral infection particles, DMEM was supplemented with 1.25 μg/mL puromycin.

Lentiviral particles were obtained from The RNAi Consortium at the Broad Institute and used according to recommended protocols (http://www.broadinstitute.org/rnai/public/resources/protocols). Lentivirus stocks were prepared from a hairpin-pLKO.1 vector containing a puromycin resistance gene, and a hairpin targeting HNF4α; lentivirus was also prepared from empty vector as a control. For lentiviral infection, cells were seeded at 30,000 cells/well in 96-well plates and cultured in DMEM until 70% confluence (~48 h). Media was replaced with 100 μL antibiotic-free DMEM supplemented with 8 μg/mL protamine sulfate and 8 μL of virus/well. Plates were spun for 30 minutes at 825 × *g*, 37 °C. Following 24 h, media was removed and replaced with DMEM containing 1.25 μg/mL puromycin. Selection continued for 5 d; all wells infected with the same hairpin were pooled and further expanded under continued selection.

HNF4α knockdown was assessed by quantitative real time PCR using TaqMan gene expression assays targeted to mouse HNF4α (FAM/MGB probe; Applied Biosystems Mm00433964_m1), using β-Actin (VIC/MGB probe) as an endogenous control, according to manufacturer instructions. shRNA hairpin NM_008261.2-1313-s1c1, comprised of the target sequence GCACCAATGTCATTGTTGCTA, demonstrated the highest knockdown [70–80% by quantitative RT-PCR and Western blotting using 1:800 dilution of HNF4α antibody (Santa Cruz Biotechnology: sc-6556)] and was selected for further studies.

Other Supporting Information Files

Dataset S1 (XLS)

1. Subramanian A et al. (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 102:15545–15550.

2. Choy E et al. (2008) Genetic analysis of human traits in vitro: Drug response and gene expression in lymphoblastoid cell lines. *PLOS Genet* 4:e1000287.

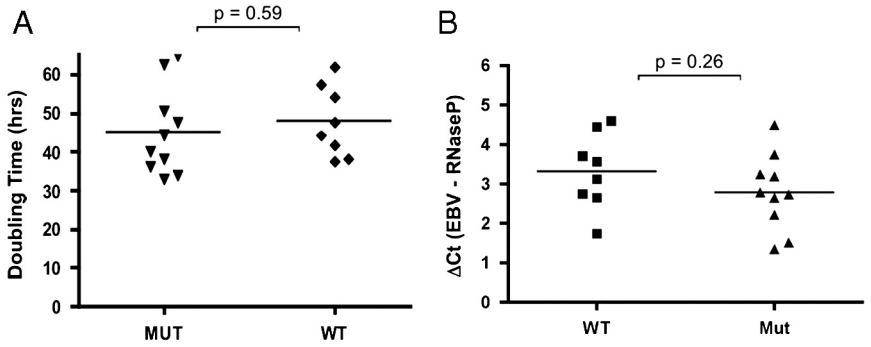

Fig. S1. No significant stratification of doubling time or EBV copy number based on mutation status was observed. Doubling time (*A*) and EBV copy number (*B*) are shown for LCLs that are either mutant or wild-type for HNF4α. EBV copy number is expressed as the difference in C_t between a fragment at the EBV DNA polymerase locus (see *SI Methods*) and endogenous RNase P expression.

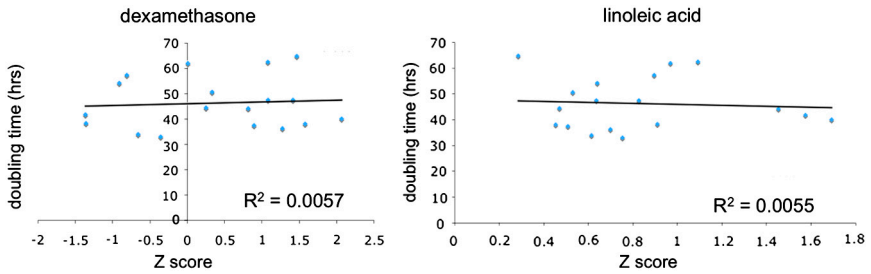

Fig. S2. Small-molecule–induced *Z* scores for ATP assay do not correlate with the doubling time of the cell line. Representative graphs shown are for dexamethasone (*Left*) and linoleic acid (*Right*).

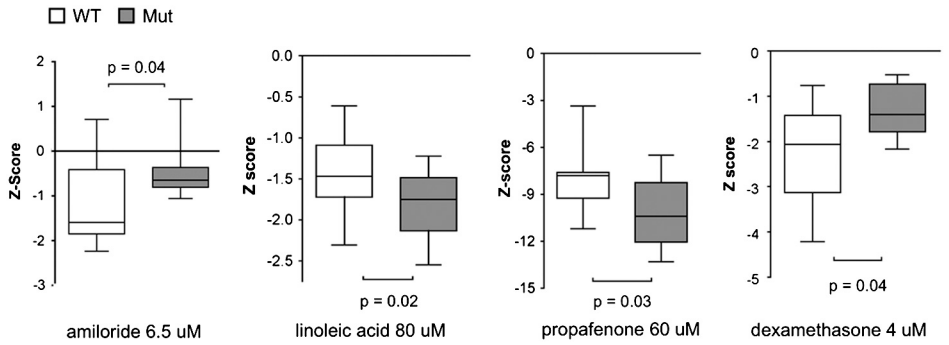

Fig. S3. Representative compounds from top-ranked compound sets induce distinct ATP measurements in LCLs that are wild-type (WT, white bars) vs. mutant (mut, gray bars) at HNF4α. Data from individual LCLs are plotted as a "box and whisker" plot; the upper and lower bounds of the box represent the 25th and 75th percentile, with a line at the median value; the whiskers extend to the highest and lowest observed values.

Table S1. Coefficient of variation (CV, expressed as percent) for 384-well ATP assay in all LCL lines screened, following either 24 h or 48 h incubation of cells at 37 °C and 5% CO_2

LCL line	CV, % (24 h)	CV, % (48 h)
1240	5.86	7.10
1242	7.62	7.81
1243	5.18	5.44
1244	6.56	6.55
1246	6.66	6.11
1247	4.44	12.34
1498	5.67	8.20
1956	5.87	5.58
8008	5.26	8.35
8106	5.90	9.73
8107	7.36	5.43
8392	4.49	7.02
8393	4.66	6.31
8756	4.95	5.44
10036	4.74	8.12
11452	7.96	6.44
11493	4.80	7.74
11494	5.58	6.03

Table S2. Compound sets used for enrichment analysis

ace inhibitors, plain
acetic acid derivatives and related antiinflammatories
actin cytoskeleton signaling
adrenergic and dopaminergic agents
alkylating agents all
alpha-adrenoreceptor antagonists
ALS signaling
aminoalkyl ether antihistamines
aminoglycoside antibacterials
aminoquinoline and methanolquinolines antimalarials
amphenicols
angiotensin ii antagonists, plain
anthracyclines and related substances
antiarrhythmics, class IA
antiarrhythmics, class IB
antiarrhythmics, class IC
anticholinergic opthalmic
antihistamines
antihistamines for systemic use all
antithyroid
arginine and proline metabolism
benzamides
benzimidazole derivatives
beta-blocking agents, nonselective
beta-blocking agents, selective
beta-adrenoreceptor agonis
beta-blocking agents all
beta-lactamase resistant penicillins
biguanides
butyrophenone derivatives
calcium signaling drugs
camp signaling
carbonic anhydrase inhibitors
cardiac B-adrenergic signaling
centrally acting antiobesity products
corticosteroids atc
coxibs
curare alkaloids
digitalis glycosides
dihydropyridine CA channel blockers
dopamine agonists
dopamine receptor signaling
eicosanoid signaling
ergot alkaloids
ERK MAPK signaling
estrogen receptor signaling
estrogens natural and semisynthetic estrogens
estrogens, all
fatty acid metabolism
fenamate antiinflammatories
fibrates
first-generation cephalosporins
fluoroquinolones
G2_M checkpoint
GABA receptor signaling
glutamate metabolism
H2-receptor antagonists
histidine metabolism
HMG CoA-reductase inhibitors
hydantoin antiepileptic
IL4 signaling
IL6 signaling
imidazole and triazole antifungals antiprotozoals
imidazoline receptor agonists
imidazoline vasodilators
inositol phosphate metabolism
insulin signaling
linoleic acid metabolism
local anesthetic—amides
local anesthetic—esters of aminobenzoic acid
long term potentiation
macrolides
methionine metabolism
monoamine oxidase inhibitors, nonselect
nitrofuran derivatives
nitrogen metabolism
nitrogen mustard analogues
nitric oxide cardiovascular signaling
nonselective monoamine reuptake inhibitor antidepressants
nonselective phenylalkylamine calcium channel blockers
nucleosides and nucleotides antivirals
omega-3 fatty acids
omega-6 fatty acids
one carbon pool by folate
opium alkaloids and derivatives
other antihistamines for systemic use
other potassium-sparing agents
other quaternary ammonium muscle relaxants
other quinolones
oxicam antiinflammatories
parasympathomimetics opthalmologic
penicillins with extended spectrum
phenothiazines with aliphatic side-chain
phenothiazines with piperazine structure
phenothiazines with piperidine structure
phenylalanine metabolism
phosphodiesterase inhibitors
piperazine derivative antihistamines
platelet aggregation inhibitors
PPAR signaling
pregnadien progestogens
pregnen (4) progestogens
progestogens
prolactine inhibitors
propionic acid antiinflammatories
propulsives
prostaglandins
purine metabolism
pyrimidine analogues
pyrimidine metabolism
quinoline antibiotics
salicylic acid and derivatives
second-generation cephalosporins
selective beta-2-adrenoreceptor agonists
selective serotonin reuptake inhibitors
serotonin (5HT3) antagonists
serotonin receptor signaling
sterol biosynthesis
stilbene lignin coumarine biosynthesis
substituted alkylamine antihistamines
substituted ethylene diamine antihistamines
sulfonamide diuretics
sulfonamides
sulfonamides loop diuretics
sulfonamides, urea derivatives
sympathomimetic decongestants
synaptic long term potentiation
synthetic anticholinergics, esters with tertiary amines
synthetic anticholinergics, quaternary amines
tertiary amine anticholinergic
tetracycline antibiotics
tetrahydropyrimidine derivatives
thiazides, plain
thiazolidinediones
thioxanthene derivatives
third-generation cephalosporins
thyroid hormones
toll-like receptor signaling
tryptophan metabolism
tyrosine metabolism
vitamin K antagonists
xanthines
xenobiotic metabolism signaling

Table S3. Individual compounds comprising top-ranked compound sets

Coxib antiinflammatories	Potassium-sparing diuretics	Glucocorticoids	Imidazole/triazole antiinfectives	HMG CoA-reductase inhibitors	Acetic acid antiinflammatories	Oxicam antiinflammatories	Class Ic antiarrhythmics	Linoleic acid	Omega-6 fatty acids
rofecoxib	amiloride hydrochloride dihydrate	dexamethasone	fluconazole	rosuvastatin	zomepirac sodium salt	tenoxicam	propafenone	arachidonic acid (20:4, *n* – 6)	docosapentaenoic acid
celecoxib	amiloride hydrochloride	betamethasone valerate	sulconazole nitrate	simvastatin	diclofenac	piroxicam	flecainide	dihomo-gamma-linolenic acid	arachidonic acid (20:4, *n* – 6)
valdecoxib	amiloride	hydrocortisone	oxiconazole nitrate	lovastatin	etodolac	meloxicam	propafenone hydrochloride	gamma-linolenic acid (18:3, *n* – 6)	adrenic acid (22:4, *n* – 6)
	triamterene	clocortolone pivalate	ornidazole	fluvastatin sodium salt	tolmetin sodium		flecainide acetate	9(S)-HPODE	dihomo-gamma-linolenic acid
		fluocinolone acetonide	metronidazole	atorvastatin calcium	bufexamac			13-(S)-HODE	gamma-linolenic acid (18:3, *n* – 6)
		hydrocortisone hemisuccinate	sertaconazole nitrate		zomepirac sodium			13-(S)-HPODE	linoleic acid
		hydrocortisone acetate	butoconazole nitrate		acemetacin			linoleic acid	eicosadienoic acid (20:2, *n* – 6)
		methylprednisolone	miconazole		aceclofenac			9(S)-HODE	
		amcinonide	clotrimazole		ketorolac tromethamine				
		triamcinolone diacetate	econazole nitrate		tolmetin sodium salt dihydrate				
		prednisolone	isoconazole		sulindac				
		dexamethasone sodium phosphate	bifonazole		diclofenac sodium				
		fluticasone propionate	ketoconazole						
		prednisolone acetate	tiabendazole						
		betamethasone	flutrimazole						
		fluocinonide	tinidazole						
		triamcinolone	miconazole nitrate						
		hydrocortisone sodium phosphate							
		prednicarbate							
		budesonide							
		hydrocortisone butyrate							
		cortisone							
		mometasone furoate							
		6alpha-methylprednisolone acetate							
		triamcinolone acetonide							
		hydrocortisone base							
		diflorasone Diacetate							
		flunisolide							
		fluorometholone							
		medrysone							
		alclometasone dipropionate							
		clobetasol propionate							
		halcinonide							
		rimexolone							
		dexamethasone acetate							
		prednisone							
		cortisone acetate							

Table S4. Two-way ANOVAs for compound treatment and LCL mutation status

	Interaction (compound × mutation status)		DMSO vs. compound		WT vs. HNF4α k/d	
Compound	2.5 mM glucose	12.5 mM glucose	2.5 mM glucose	12.5 mM glucose	2.5 mM glucose	12.5 mM glucose
Linoleic acid	<0.0001	0.5734	<0.0001	0.1062	<0.0001	<0.0001
Propafenone	0.0370	0.0461	0.3219	<0.0001	0.0019	0.0002
Amiloride	<0.0001	0.2564	<0.0001	0.0494	0.0004	0.2618
Simvastatin	0.6469	0.0008	<0.0001	<0.0001	0.5952	0.0180

P values are listed for the interaction between factors (i.e., effect of the small molecule depends on mutation status), the main effect of compound treatment (DMSO vs. compound), and the main effect of mutation status (wild-type vs. HNF4α knockdown). Values are given for two-way ANOVA at low (2.5 mM) glucose and two-way ANOVA at high (12.5 mM) glucose.