Establishing a Central Resource of Data from Genome Sequencing Projects: Target Validation Landscape

Jeff Trimmer, Ph.D.

June 5-6, 2012

Target Validation Landscape

- ***Drug target validation:***
  - How could the database be used for drug target validation (such as identifying people with putative loss-of-function variants for more phenotyping)?
  - What type of data are needed?
  - How important is the ability to re-contact participants?
  - How would study designs for drug target validation differ from ones for disease studies?

Target Validation “Definitions”

- ***A Valid Target is*:** A target that when modulated pharmacologically, provides meaningful efficacy and acceptable safety for specific human disease in long-term clinical usage.
- ***Target Validation is*:** The process of demonstrating in a clinical trial that engaging the target provides statistically meaningful therapeutic benefit with acceptable safety for a given indication.
- ***Target Qualification is*:** Preclinical or limited clinical studies prior to well-powered clinical trials, that establish the scientific validity and safety of a drug target; it is part of the continuum of target validation**.**
- ***Target Identification is*:** The generation of scientific evidence that a manipulatable able target is involved in some significant way in a disease process

Joint NIH-Industry Target Validation Workstreams Workshop, May 31st & June 1st, Cambridge, MA

An Example

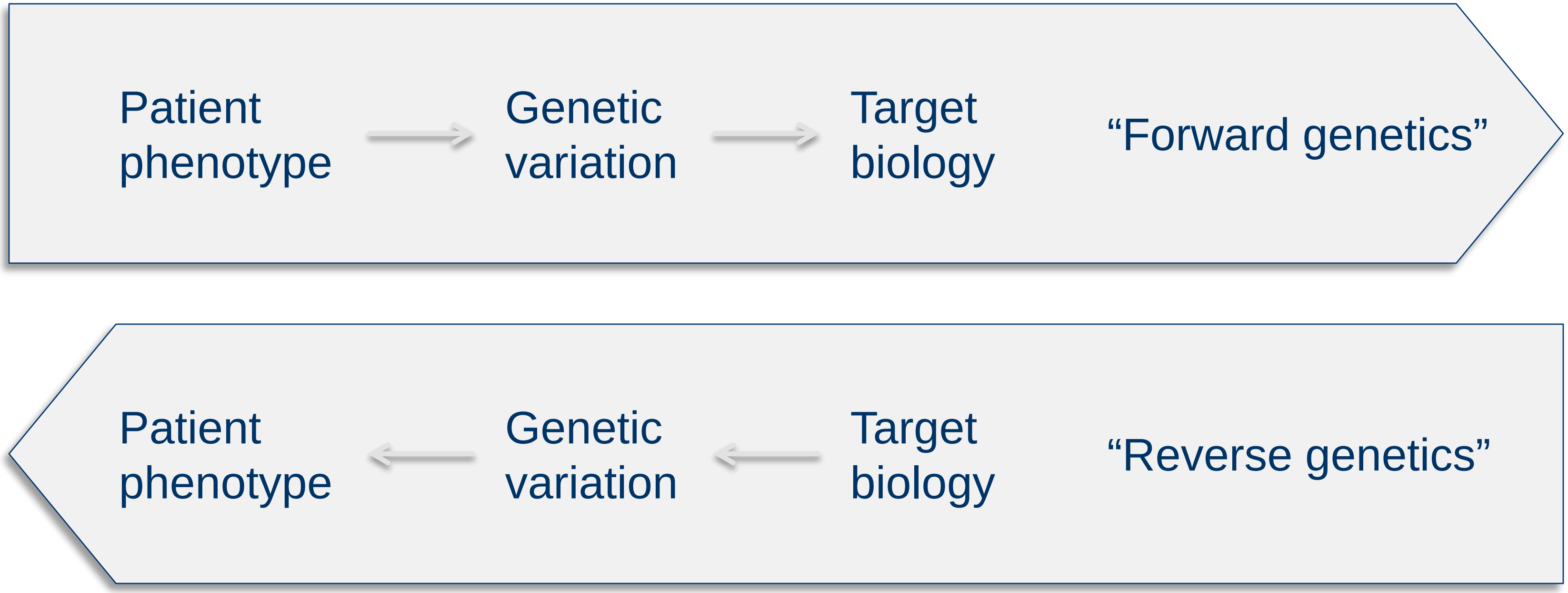
Pfizer BROAD INSTITUTE MGH 1811

The primary goals of our collaboration are to validate the human *in vivo* relevance of potential drug targets via human genetics…

…and to develop and apply novel methods to bridge from patients in the human population to tractable biology and drug targets *in vitro*

Unbiased vs. Biased Target Validation

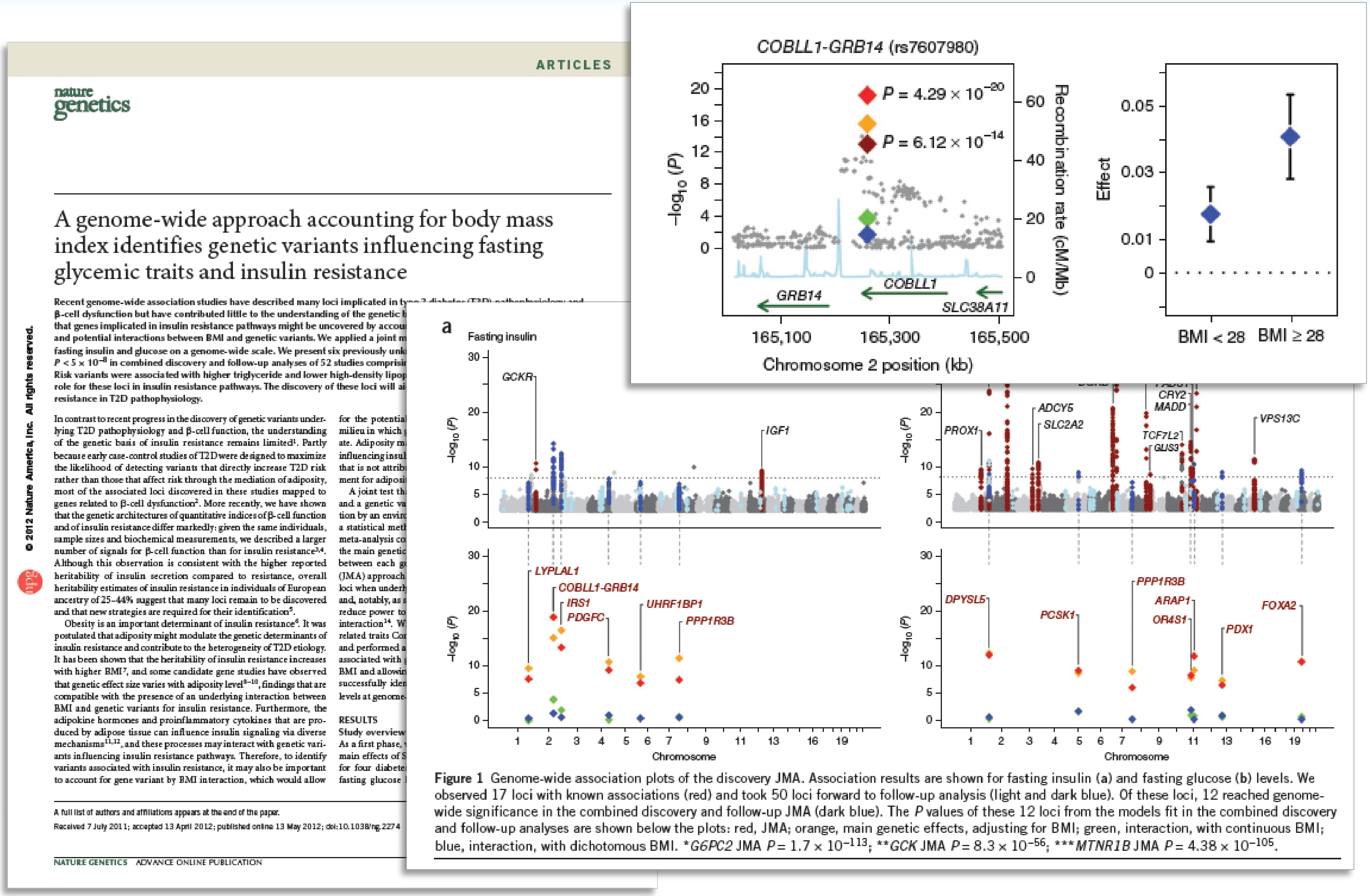
- **“Can you find me a new target backed by human genetics?” vs. “Is my target supported by human genetic evidence?**



Patient phenotype → Genetic variation → Target biology “Forward genetics”

Patient phenotype ← Genetic variation ← Target biology “Reverse genetics”

Integration of Multiple Databases May Enable Further Characterization of the Etiology and Causality of Disease



ARTICLES

nature genetics

A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance

Recent genome-wide association studies have described many loci implicated in type 2 diabetes (T2D) pathophysiology and
β-cell dysfunction but have contributed little to the understanding of the genetic b
that genes implicated in insulin resistance pathways might be uncovered by accou
and potential interactions between BMI and genetic variants. We applied a joint m
fasting insulin and glucose on a genome-wide scale. We present six previously unk
$P < 5 \times 10^{-8}$ in combined discovery and follow-up analyses of 52 studies comprisi
Risk variants were associated with higher triglyceride and lower high-density lipop
role for these loci in insulin resistance pathways. The discovery of these loci will ai
resistance in T2D pathophysiology.

In contrast to recent progress in the discovery of genetic variants underlying T2D pathophysiology and β-cell function, the understanding of the genetic basis of insulin resistance remains limited[1]. Partly because early case-control studies of T2D were designed to maximize the likelihood of detecting variants that directly increase T2D risk rather than those that affect risk through the mediation of adiposity, most of the associated loci discovered in these studies mapped to genes related to β-cell dysfunction[2]. More recently, we have shown that the genetic architectures of quantitative indices of β-cell function and of insulin resistance differ markedly: given the same individuals, sample sizes and biochemical measurements, we described a larger number of signals for β-cell function than for insulin resistance[3,4]. Although this observation is consistent with the higher reported heritability of insulin secretion compared to resistance, overall heritability estimates of insulin resistance in individuals of European ancestry of 25–44% suggest that many loci remain to be discovered and that new strategies are required for their identification[5].

Obesity is an important determinant of insulin resistance[6]. It was postulated that adiposity might modulate the genetic determinants of insulin resistance and contribute to the heterogeneity of T2D etiology. It has been shown that the heritability of insulin resistance increases with higher BMI[7], and some candidate gene studies have observed that genetic effect size varies with adiposity level[8–10], findings that are compatible with the presence of an underlying interaction between BMI and genetic variants for insulin resistance. Furthermore, the adipokine hormones and proinflammatory cytokines that are produced by adipose tissue can influence insulin signaling via diverse mechanisms[11,12], and these processes may interact with genetic variants influencing insulin resistance pathways. Therefore, to identify variants associated with insulin resistance, it may also be important to account for gene variant by BMI interaction, which would allow

for the potential
milieu in which
ate. Adiposity m
influencing insul
that is not attrib
ment for adiposi

A joint test th
and a genetic va
tion by an envir
a statistical meth
meta-analysis co
the main genetic
between each g
(JMA) approach
loci when underl
and, notably, as
reduce power to
interaction[14]. W
related traits Co
and performed a
associated with
BMI and allowi
successfully ide
levels at genome-

RESULTS

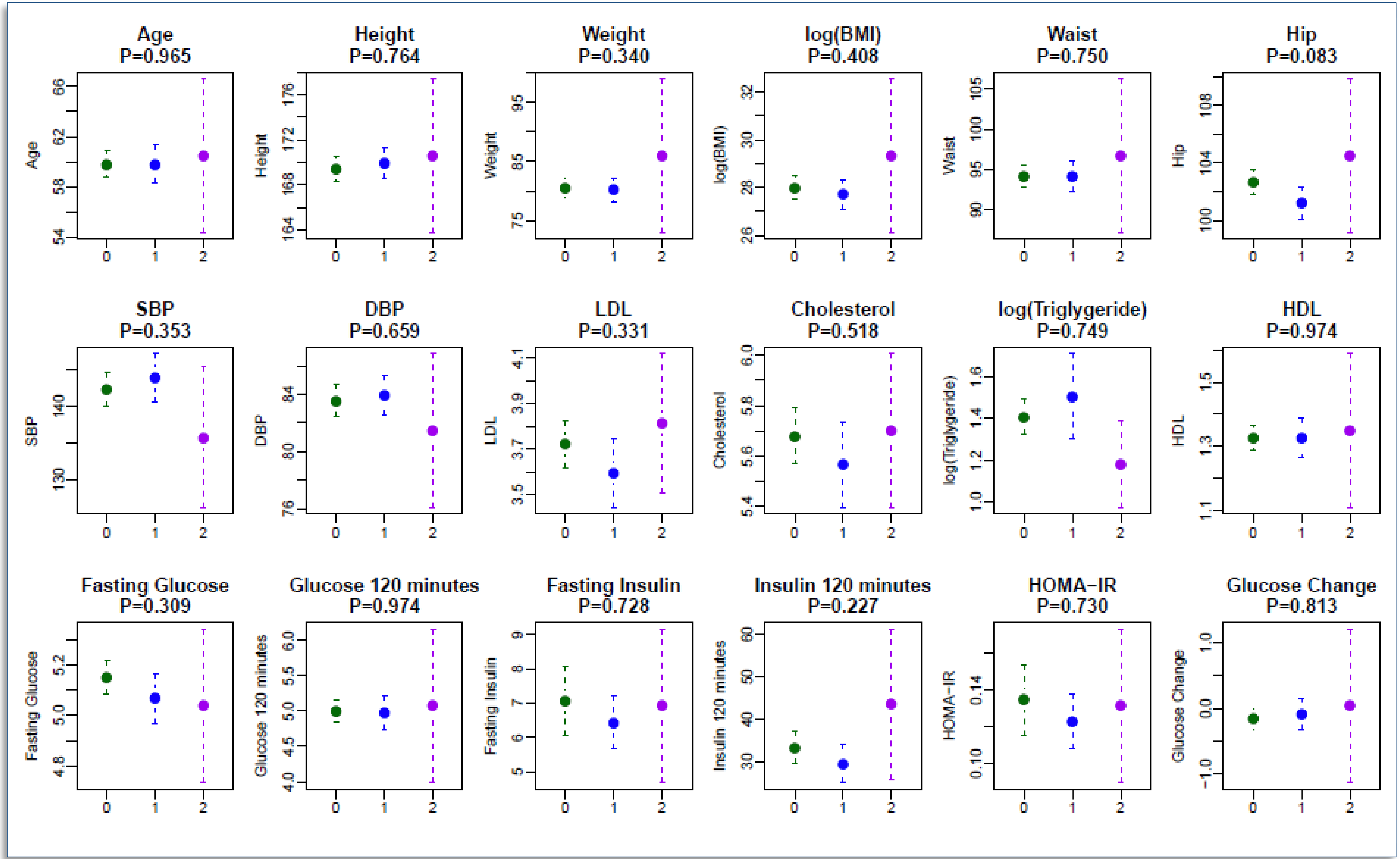
Study overview

As a first phase,
main effects of S
for four diabete
fasting glucose

A full list of authors and affiliations appears at the end of the paper.

Received 7 July 2011; accepted 13 April 2012; published online 13 May 2012; doi:10.1038/ng.2274

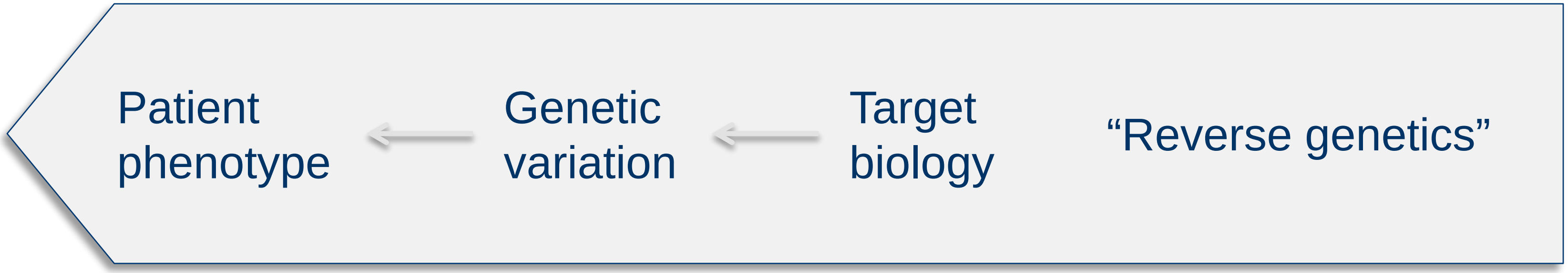
Figure 1 Genome-wide association plots of the discovery JMA. Association results are shown for fasting insulin (a) and fasting glucose (b) levels. We observed 17 loci with known associations (red) and took 50 loci forward to follow-up analysis (light and dark blue). Of these loci, 12 reached genome-wide significance in the combined discovery and follow-up JMA (dark blue). The *P* values of these 12 loci from the models fit in the combined discovery and follow-up analyses are shown below the plots: red, JMA; orange, main genetic effects, adjusting for BMI; green, interaction, with continuous BMI; blue, interaction, with dichotomous BMI. **G6PC2* JMA $P = 1.7 \times 10^{-113}$; ***GCK* JMA $P = 8.3 \times 10^{-56}$; ****MTNR1B* JMA $P = 4.38 \times 10^{-105}$.



Age
P=0.965
Age
54 58 62 66
Height
P=0.764
Height
164 168 172 176
Weight
P=0.340
Weight
75 85 95
log(BMI)
P=0.408
log(BMI)
26 28 30 32
Waist
P=0.750
Waist
90 95 100 105
Hip
P=0.083
Hip
100 104 108
SBP
P=0.353
SBP
130 140
DBP
P=0.659
DBP
76 80 84
LDL
P=0.331
LDL
3.5 3.7 3.9 4.1
Cholesterol
P=0.518
Cholesterol
5.4 5.6 5.8 6.0
log(Triglygeride)
P=0.749
log(Triglygeride)
1.0 1.2 1.4 1.6
HDL
P=0.974
HDL
1.1 1.3 1.5
Fasting Glucose
P=0.309
Fasting Glucose
4.8 5.0 5.2
Glucose 120 minutes
P=0.974
Glucose 120 minutes
4.0 4.5 5.0 5.5 6.0
Fasting Insulin
P=0.728
Fasting Insulin
5 6 7 8 9
Insulin 120 minutes
P=0.227
Insulin 120 minutes
30 40 50 60
HOMA−IR
P=0.730
HOMA−IR
0.10 0.14
Glucose Change
P=0.813
Glucose Change
−1.0 0.0 1.0
0 1 2

Reverse Genetics: Is My Target Supported by Genetic Evidence?

- **Start with biology of interest to drug development**
- **Identify which genes have loss of function (or other strong) mutations based on sequencing**
- **Call in genotype carriers and perform in-depth phenotyping to evaluate effect of perturbation**



Patient phenotype ← Genetic variation ← Target biology

“Reverse genetics”

Gene Symbol	Original Target Hypothesis	Category
CPT1A	CPT1	Fatty acid metabolism
CPT1B	CPT1	Fatty acid metabolism
CPT1C	CPT1	Fatty acid metabolism
CPT2	CPT1	Fatty acid metabolism
Acaca	Acaca-Mm01304277_m1	Insulin signalling
Fasn	Fasn-Mm00662319_m1	Insulin signalling
G6pc	G6pc-Mm00839363_m1	Insulin signalling
Pck1	Pck1-Mm01247058_m1	Insulin signalling
Ppargc1a	Ppargc1a- Mm01208835_m1	Insulin signalling
Srebf1	Srebf1-Mm00550338_m1	Insulin signalling
LIPE	HSL	Insulin signalling
PRKAA1	AMPK	Insulin signalling
PRKAA2	AMPK	Insulin signalling
PRKAB1	AMPK	Insulin signalling
PRKAB2	AMPK	Insulin signalling
PRKAG3	AMPK	Insulin signalling
DGKA	DGKA	Metabolism of triglycerides
DGKB	DGKB	Metabolism of triglycerides
DGKD	DGKD	Metabolism of triglycerides
DGKE	DGKE	Metabolism of triglycerides
DGKG	DGKG	Metabolism of triglycerides
DGKI	DGKI	Metabolism of triglycerides
DGKQ	DGKQ	Metabolism of triglycerides
DGKZ	DGKZ	Metabolism of triglycerides
PPAP2A	PPAP2A (LPP1)	Metabolism of triglycerides
PPAP2B	PPAP2B (LPP2)	Metabolism of triglycerides
PPAP2C	PPAP2C (LPP3)	Metabolism of triglycerides