View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Mouse phenogenomics: The fast track to "systems metabolism"

Review

Carmen A. Argmann,[1] Pierre Chambon,[2] and Johan Auwerx[1,2,*]

[1] Institut de Génétique et de Biologie Moléculaire et Cellulaire, Centre National de la Recherche Scientifique/Institut Nationale de la Santé et de la Recherche Médicale/Université Louis Pasteur, 67404 Illkirch, France
[2] Institut Clinique de la Souris, 67404, Illkirch, France
*Correspondence: auwerx@igbmc.u-strasbg.fr

With the completion of the many genomes, genetics is positioned to meet physiology. In this review, we summarize the coming of "systems metabolism" in mammals through the use of the mouse, as a model system to study metabolism. Building on mouse genetics with increasingly sophisticated clinical and molecular phenotyping strategies has enabled scientists to now tackle complex biomedical questions, such as those related to the pathogenesis of the common metabolic disorders. The ultimate goal of such strategies will be to mimic human metabolism with the click of a mouse.

Mice are valuable experimental models for biomedical research, as they are nearly genetically identical to humans, with many biochemical and physiological pathways being conserved. Compared to other vertebrate models, the genetic technologies required to introduce mutations and interrogate the mouse genome are the most developed. To scrutinize the polygenic networks underlying complex diseases, however, mouse resources that are optimized to study the actions of isolated genetic loci on a fixed background will be insufficient on their own. For example, predisposition to the metabolic syndrome is inherited in a non-Mendelian fashion stressing genetic heterogeneity and multigenetic pathogenesis (Nandi et al., 2004). With the reawakening as to the extraordinary genetic resources and phenotypic diversity archived in extant inbred strains, however, a foundation is in place for tracking down these complex traits and quantitative trait loci (QTL). Moreover, environmental factors can impact on the manifestations of the genotype (incomplete penetrance) or produce a phenotype that mimics the genetically produced phenotype (phenocopy; Paigen, 2003). Problems due to inherent genetic and environmental heterogeneity can be minimized in the mouse as the genetic background, and environmental factors can be rigorously controlled. These features combined earn the mouse the status of the preferred model to study complex biomedical problems, with the expectation that they will help to elucidate the pathogenesis of common metabolic diseases such as obesity, type 2 diabetes (T2DM), and atherosclerosis. In this review, we summarize how the judicious use of mouse models can propel metabolic research toward "systems metabolism," ultimately helping to better understand human metabolism.

Genetically engineered mouse models (GEMMs) and reverse genetics

Reverse genetic strategies have been truly powerful in understanding molecular pathways, since the phenotypes induced by gene ablation or knockout experiments represent the most effective route to acquire information on gene function (Figure 1). The values and inherent limitations of simple gene-driven strategies of "one gene, one protein, one phenotype" are elegantly portrayed by the efforts to dissect the genetics of insulin resistance and β cell dysfunction. Many GEMMs have been created using embryonic stem (ES) cell technology to evaluate candidate "diabetogenes" (Nandi et al., 2004), a strategy most successful at elaborating complex "wiring" diagrams that trace the path from environmental stimuli through cell surface receptors, downstream targets, to gene expression changes. However, many of the phenotypes of the homozygous null mutations were extreme and/or did not model the complexity of the metabolic syndrome. For example, IR knockout ($IR^{-/-}$) mice died because of developmental effects (Accili et al., 1996), which precluded analysis of adult mice. Likewise, $GLUT4^{-/-}$ mice exhibited only moderate insulin resistance and were not overtly diabetic, suggesting compensatory mechanisms (Katz et al., 1995). Monogenic GEMMs furthermore ignore the polygenic nature of metabolic diseases, resulting from genetic and environmental factors impacting at multiple levels in signaling cascades. Oligogenic mouse models remedied some of these shortcomings. For example, IR and IRS1 compound heterozygote mice developed severe insulin resistance in muscle and liver (Kido et al., 2000), a phenotype not observed in either mouse model alone. This demonstrated that each predisposing allele, although having a modest effect alone, plays a prominent role in a sensitized background. Such oligo- or polygenic mouse models also demonstrated that interactions among genetic loci could either create strong positive or negative (protective) effects in the predisposition of diabetes.

Realization that metabolic diseases do not only involve many genes but also many organs necessitated engineering the mouse genome with spatial and temporal controls (Figure 1). This requires the generation of premutant mice whereby the allele of interest is flanked by recognition sites for DNA recombinases such as Cre (loxP sites) or Flp (Frt sites; Branda and Dymecki, 2004). When such premutant mice are bred with transgenic mice that express the corresponding DNA recombinase in a tissue-specific fashion, the gene of interest is inactivated only in that particular tissue. An added sophistication of spatially controlled mutagenesis is the inclusion of temporal control, which is achieved by using ligand-activated chimeric recombinases composed of the fusion of the recombinase with the ligand binding domain of a nuclear receptor (Metzger and Chambon, 2001).

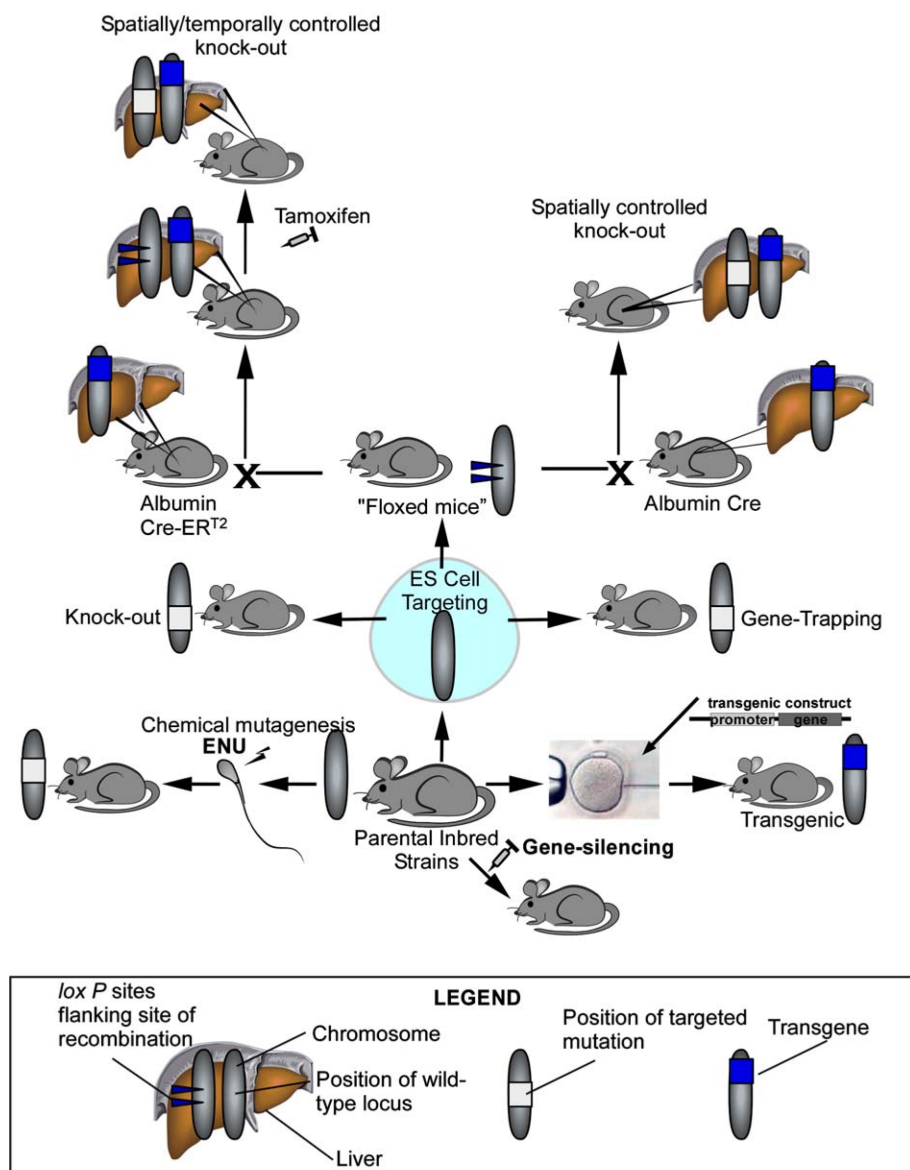


Figure 1. Genetically engineered mouse models (GEMMs)

Gene targeting can be restricted in a tissue and/or temporal manner through generation of "premutant" mice (e.g., "floxed" or flanked with loxP sites) which are then bred with transgenic mice that express the corresponding DNA recombinase (e.g., Cre), allowing for gene deletion to occur in a given tissue (right top panel) and at a given time (left top panel). Gene-trap mutagenesis involves the random insertion of gene-trap vectors into the genome, whereas ethylnitrosourea (ENU) is used as an in vivo mutagen to generate point mutations. Gene silencing is possible through RNAi, a method which uses siRNAs or shRNAs, which can be delivered either by direct administration or via transgenic, homologous recombination, or viral delivery technologies, respectively.

Employing spatially and temporally controlled mutagenesis revealed that metabolic perturbation in one tissue encroaches on the metabolism of another. This tissue communication and crosstalk allows for metabolic flexibility, with respect to the use or production of metabolic fuel. Understanding this may ultimately reveal the genetic determinants that limit the degree of metabolic flexibility, thus promoting metabolic diseases in humans (Bickel, 2004). One example of a spatially controlled mouse mutant included the muscle-specific $IR^{-/-}$ mice, which did not develop systemic insulin resistance; a surprising observation considering that the muscle is the primary insulin-dependent glucose-disposing tissue (Bruning et al., 1998). These mice escaped insulin resistance by compensating with insulin sensitization of the adipose tissue, shifting glucose utilization from muscle toward adipose tissue. An equivalent example on the "lipid side" was the liver-specific ablation of the escort protein, sterol regulatory element binding protein (SREBP) cleavage-activating protein (Kuriyama et al., 2005). The resulting selective reduction in SREBP-1 activation, which diminished liver fatty acid synthesis, was compensated for on a molecule-for-molecule basis by adipose tissue such that total body fatty acid synthesis remained unchanged.

The advantage of adding a temporal dimension to spatial-specific gene ablation was illustrated by studies with peroxisome proliferator-activated receptor γ 2 (PPARγ2). Mice with a selective disruption of the PPARγ2 gene in adipose tissue were lipodystrophic, underscoring its predominant role in adipogenesis (He et al., 2003; Koutnikova et al., 2003). Although these mice revealed that the specific absence of PPARγ in fat impacts "nonadipose" tissues such as bone (Cock et al., 2004), they did not expose the vital role of PPARγ in differentiated adipocytes. Selective ablation of PPARγ2 in adipocytes of adult mice, by an inducible Cre recombinase, induced mature adipocytes to die within a few days; however, these cells were replaced days later with newly differentiated PPARγ-expressing adipocytes (Imai et al., 2004), revealing the importance of PPARγ in adipocyte survival. Tissue-specific overexpression of transgenes can also be achieved through several strategies, with or without

temporal control, based on the use of tissue-specific promoters. For instance, when FOXC2 was overexpressed in adipose tissue, mice remained lean and insulin sensitive despite a high-fat diet (Cederberg et al., 2001). This was in part due to increased metabolic efficiency as white adipose tissue acquired brown fat-like features.

Genetic engineering strategies that generate point mutations or chromosomal alterations that mimic human mutations are also relevant for modeling the genetic structure of human populations in the mouse. For example, the P467L mutation in the ligand binding domain of PPARγ in humans causes severe insulin resistance and hypertension; when this mutation was recreated in mice a critical role for PPARγ in blood pressure control became evident (Tsai et al., 2004). Likewise, a S112A knockin mouse confirmed a modulatory role for phosphorylation on PPARγ function (Rangwala et al., 2003). The transchromosomic mouse line, Tc1, which is a model of trisomy 21 and Down's syndrome in humans, demonstrated the success of the mouse even for dissecting complex human aneuploidies (O'Doherty et al., 2005).

In order to increase the efficacy of functional gene annotation, high throughput mutagenesis methods, such as gene trapping and gene silencing by RNAi, are now becoming commonplace. Gene-trap mutagenesis enables the random generation of loss-of-function mutations and many of the vectors allow for the simultaneous evaluation of the expression of the trapped gene. Several large-scale gene-trap efforts are underway to generate a public resource of mutagenized ES cells (Austin et al., 2004; Auwerx et al., 2004). However, to move toward saturation mutagenesis of the mouse genome and to isolate mutations in genes not expressed in ES cells or that are refractory to gene trapping other techniques such as classical gene-targeting and targeted-trapping strategies will still be required (Skarnes, 2005). Although the time-consuming validation to verify effective gene inactivation upon gene-trapping may pose a bottleneck when these trapped ES cells are converted into mice, this technique holds promise to define new genes involved in metabolism.

RNA interference (RNAi) is an alternative method to silence gene expression that is accomplished either with sequence-specific short interfering RNAs (siRNA) or through the transcription of short hairpin RNAs (shRNA) (Dykxhoorn et al., 2003; Figure 1). Although RNAi has become the experimental tool of choice for genome-scale analyses of gene function in mammalian cells, siRNA and shRNA technologies when combined with viral delivery also have potential in vivo. One domain, in which RNAi-generated mouse models may be valuable, concerns genes that yield very stable mRNAs and are difficult to inhibit with gene-based strategies. siRNAs also show promise as a therapeutic approach to silence disease-causing genes, particularly those that encode so-called "nondruggable" targets. One example has been the use of chemically modified siRNAs to silence endogenous apolipoprotein B (apoB) mRNA, in liver and jejunum, resulting in decreased plasma apoB and cholesterol levels in treated mice (Soutschek et al., 2004). It is, however, unlikely that RNAi technology will replace gene targeting or trapping to generate mouse models, since despite numerous advantages, many shortcomings still exist. For example, siRNA delivery must be optimized to improve subcellular localization, decrease genotoxic effects (IFN response), diminish off-target effects, and increase consistency of silencing. Validation of RNAi-generated mouse models is also problematic, as issues of tissue specificity, degree and duration of silencing, and toxicity are critical.

Forward genetic strategies

Forward genetics is centered on phenotype-driven gene discovery that exploits mouse models with subtle genetic variations resulting from allelic variants. The enthusiasm for these models over reverse genetic strategies has been the random nature of the mutations, which covers the full gamut from point mutations to deletions and to rearrangements and thus produces a full spectrum of changes in protein structure and function. Initially, a large number of these mouse models resulted from spontaneous mutations that yielded striking phenotypes, such as the obesity-associated metabolic problems, caused by mutations in the genes encoding leptin (Halaas et al., 1995) and its receptor (Chen et al., 1996). In an effort to provide a broader supply of spontaneous-like mutations, strategies including chemical and radiation mutagenesis of sperm or ES cells have been adopted, with ethylnitrosourea (ENU) being the mutagen of choice (O'Brien and Frankel, 2003; Figure 1). Although with ENU mutagenesis the generation of mutant mice is fast, the real success of this technique rests in the implementation of validated phenotyping resources relevant to the trait/disease of interest. Once a mutant phenotype has been identified, the underlying gene must then be identified to make the connection between gene and function. This task remains timely and costly, as mutants must be backcrossed to other inbred strains to identify cosegregating markers. Furthermore, random mutagenesis screens to identify recessive mutations have until recently been rather complex and costly. This may change with the use of ES cells that are deficient in the Bloom's syndrome gene, which expedites conversion of heterozygous into homozygous mutants (Guo et al., 2004; Yusa et al., 2004). So far, the value of ENU mutants has mainly been limited to the identification of dominant mutations that give rise to monogenic diseases. For example, a large-scale ENU mutagenesis screen for dominant traits identified 12 mutations in the glucokinase gene, a maturity onset diabetes of the young type 2 (MODY2) gene (Inoue et al., 2004; Toye et al., 2004). Since glucokinase$^{-/-}$ mice are embryonic lethal, this collection of glucokinase mutants is useful for dissecting the pathogenesis of MODY2.

Genetic reference populations (GRPs)

Perhaps the most "refreshing" mouse resource for investigating complex diseases is the construction of mouse crosses using inbred mice and the subsequent QTL mapping. Inbred mice have an inherent wealth of variation due to past spontaneous mutation events, which have been preserved through systematic and uninterrupted brother-sister matings (Paigen, 2003). Inbred mice are appealing since they are genetically identical within a strain but are diverse between strains. Crosses usually expose subtle variations affecting many traits of biomedical interest. The Complex Trait Consortium was formed to systematically exploit large-scale resources with the ultimate aim to partition these strain differences in hundreds of quantitative traits to sets of biologically related genes or loci with the ultimate aim the identification of genomic regions and genes associated with these traits (Threadgill et al., 2002). However, mapping QTLs based solely on the parental strain phenotypes will not be effective enough in that many genetic contributors to complex diseases will be missed (Li et al., 2005). This is because the power

and accuracy to detect QTLs depend on multiple factors, including genetic diversity of the parents, marker density, trait heritability, and size of cross. In any single or double cross, recombination events and mapping resolution is thus limited, as we can only detect loci that show allelic variation between two strains. Such complex mouse genetic approaches, however, are continually evolving to offer the necessary variability, power, or resolution. For instance, the use of multiple crosses, which improves the power and resolution, is nicely demonstrated when data from a genome-wide analysis of HDL QTLs collected from multiple crosses (four single crosses involving five strains) was computationally combined in a metacross analysis (Li et al., 2005). This metacross analysis identified four significant QTLs, including some that were either significant or nearly so in each single cross and substantially narrowed some of the QTL regions.

Another advancement has been the generation of large-scale GRPs of recombinant inbred (RI) mouse lines that have a higher rate of polymorphism between strains. GRPs are genetically well-defined mouse strains derived by systematically inbreeding, for at least 20 generations, the F2 from a cross of two distinct inbred strains and ultimately representing an ∼50:50 genetic mixture of the progenitor strains (Churchill et al., 2004; Williams et al., 2001; Figure 2). Thus, the strains that make up a GRP are not mutants or engineered mice; they are normal lines that have a level of variation more similar to that among human populations. Importantly, GRPs are a renewable resource which only needs to be genotyped once since the strains are genetically identical. However, at present there are too few GRPs with limited QTL power. For example, a recent study compared QTL analyses for aortic fatty lesion size and plasma lipid levels among high-fat-fed parental mice (C57BL/6 and DBA, >10 animals/group), BXD RI strains (18 strains, two to four animals/group), and B6D2F2 mice (∼140 animals) (Colinayo et al., 2003). No differences in plasma lipids or lesion formation were observed between the two parental strains. In contrast, there were significant differences among the 18 BXD RI strains when the same traits were evaluated, confirming the presence of multigenic determinants. However, QTL analysis did not reveal significant linkages for either of these traits. In contrast, in the B6D2F2 population there were significant changes in lipids and aortic lesion size, enabling the mapping of three loci with suggestive linkage to aortic lesion size. Such results show the limitations of current RI lines and spurred interest to create as many as 1000 new RI mouse lines with hopes to cover much of the genetic variation present in natural populations (Churchill et al., 2004; Threadgill et al., 2002). From these RI strains, it will be possible to generate an unlimited combinatorial diversity using the F1 progeny to generate recombinant inbred crosses (RIX).

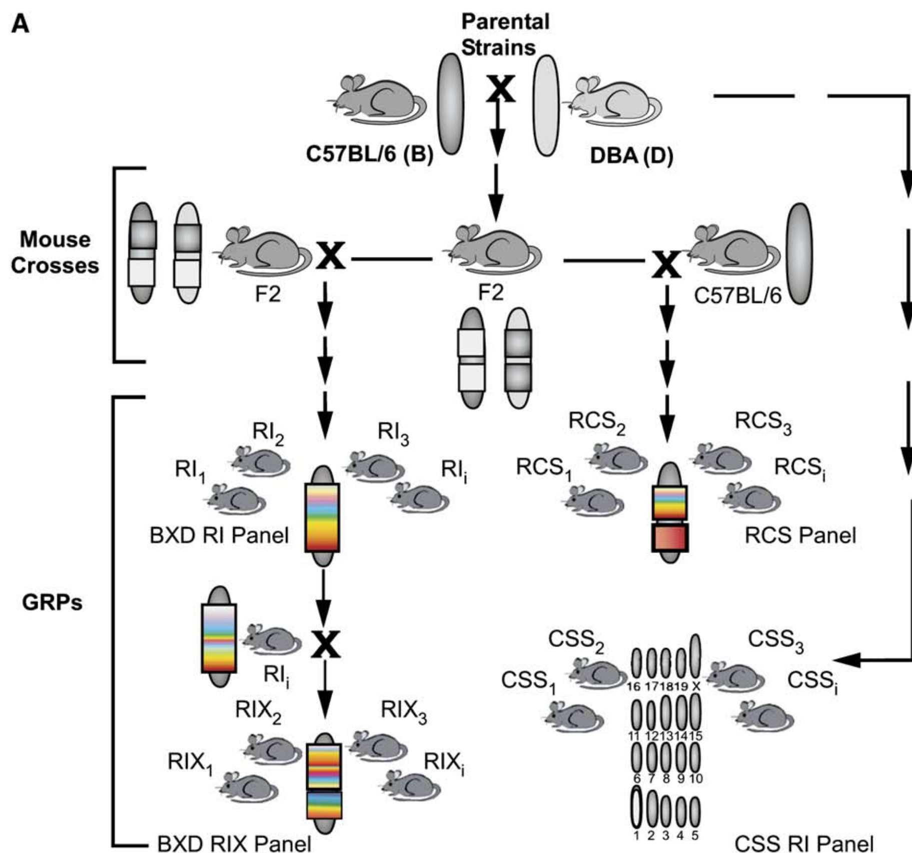
Further sophistication of RI strains include recombinant congenic strains (RCS), whereby two parental inbred strains are initially crossed, but the resulting progeny are then backcrossed with one of the parental strains prior to sib-pair matings (Figure 2). Instead of the mix of progenitor genomes approaching half and half, the genome of RCS is mainly representative of one parent. One example includes the mice generated by combining independent diabetes risk-conferring QTLs from two unrelated parental strains, the New Zealand obese (NZO) and nonobese nondiabetic mice (Reifsnyder and Leiter, 2002). These mice lack the extreme metabolic phenotype characteristic of the parental NZO mice and are more akin to T2DM in that the diabesity is differentially expressed among the RCS. Importantly, allelic variants controlling these differences could be characterized. RCS are mainly beneficial when a single gene explains the majority of variance in a trait, otherwise RI strains have the advantage.

Chromosome substitution strains (CSS) are another useful and permanent resource. A CSS panel partitions the variation between two strains by chromosome (Figure 2; Singer et al., 2004). CSS are constructed by successive backcrosses between a donor strain and a host strain. Progeny carrying a nonrecombinant copy of a given chromosome are selected and backcrossed to produce progenitors that are homosomic for the desired chromosome. A complete CSS panel (22 strains, one for each autosome, the sex chromosomes, and the mitochondria) has been derived from the inbred A/J and C57BL/6J strains as donor and host, respectively. These CSS lines were used to dissect genetic factors affecting 53 complex traits related to sterol levels, diet-induced obesity, anxiety and amino acid levels, and readily provided evidence for 150 QTLs. Eight of the inferred QTLs were then analyzed by fine-mapping strategies, which identified in each of the cases a QTL in a specific location on the substituted chromosomes (Singer et al., 2004).

Genetical genomics and phenogenomics

The power to decipher the molecular networks underlying the multifactorial origin of complex diseases is further boosted by applying "omics" approaches to mouse models. When combined, the omics strategies verge on systems biology in such a way that functional genomics, proteomics, and metabolomics parallel the hierarchy of transcription, translation, and production of small molecules. One example is "genetical genomics" or "expression genetics," which pairs traditional genetic mapping techniques with advanced microarray technology (see below; Damerval et al., 1994; Jansen and Nap, 2001; Klose et al., 2002; Brem et al., 2002; Chesler et al., 2003; Schadt et al., 2003).

With all the successes in genetic manipulation of the mouse genome, the real bottleneck to systems metabolism is mouse phenogenomic strategies, which rely on efficient and vigilant mouse phenotyping. It is now recognized that the phenotype of a mouse is highly dependent on both genetic and environmental factors. Thus, all experiments must begin with careful consideration of strain background, as this influences almost all variables and biases the resulting phenotype (Paigen et al., 1985; Seong et al., 2004). This becomes highly significant when one considers the huge number of studies performed in gene-targeted or gene-trapped mice on a nonhomogeneous genetic background. If the background strain of the phenotyped mouse is different from the one used to generate the GEMM, mice must be backcrossed for multiple generations (>9) to avoid confounding factors contributed by flanking donor chromosomal DNA. Alternatively, flanking donor chromosomal DNA can be monitored by using a marker-assisted selection protocol and "speed backcrossing" to generate congenic strains in half the time compared to "traditional backcrossing" (Collins et al., 2003). Although confounding factors in the analysis of mice with single gene defects, strain-specific differences are a treasure trove when one aims to study genotype/phenotype correlations of complex traits.



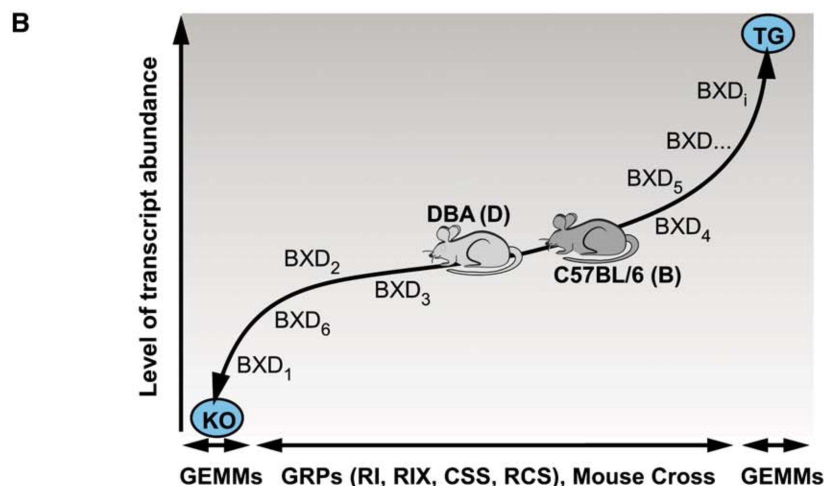


Figure 2. Recombinant inbred strains (RI) and mouse crosses

A) RI strains (e.g., BXD panel) are derived by systematic inbreeding of F2 mice, from a cross of two distinct inbred strains, such as DBA (D) and C57BL/6 (**B**), for multiple generations, as indicated by arrows. Their genome is a mosaic of the two parental strains. RIX mapping involves generating F1 crosses between individual RI strains from a particular RI panel such as BXD that results in a higher density of recombination. RCS are a variation on RI. Following the initial outcross of mice of two inbred strains, F2 are backcrossed to mice of one of the parental strains for one or two generations prior to sibling matings. Whereas RI strains represent a 50:50 genetic mixture of the progenitor strains, the genome of RCS are derived predominantly from one parent. A CSS panel usually consists of 22 mouse strains, each of which carries a single chromosome substituted from the donor strain, onto the host background. In this theoretical example, DBA is the donor and C57BL/6 is the host. They are constructed by successively backcrossing to the host strain progeny identified in each generation as carrying a nonrecombinant copy of the desired donor chromosome until the progeny are heterosomic for the desired chromosome on an otherwise host background. These progeny are then intercrossed to produce progeny homosomic for the desired chromosome.

B) In this theoretical example, expression levels of a given gene are confined between the extreme changes in expression seen in the knockout (KO) and transgenic (TG) lines of this gene. The level of genetic variation in GRPs, mouse crosses, and CSS results in changes in gene expression level that are subtler and cover a wider spectrum (e.g., BXD panel).

In addition to genetic factors, environmental factors have a major impact on the outcome of phenotyping tests. Some of these factors have become experimental tools to challenge metabolic pathways (e.g., diet and exercise). Many others, however, have become standard "housekeeping" items, in that their order and consistency must be maintained. Typical housekeeping items include the animal housing and handling conditions, such as the number of animals per cage, diurnal rhythm, the blood collection procedure, age, and gender of the mice (Champy et al., 2004). Cage-housing density significantly impacts many metabolic variables, as illustrated by increased glucose levels in animals caged in pairs as compared to those housed in groups of four (Champy et al., 2004). Finer housekeeping details, such as the timing of blood collection, should also be critically controlled since metabolic parameters are affected by nychthemeral variation in activity and food intake (higher at night) and hormone levels (Champy et al., 2004). Even though it is clear that the potential variables in phenotyping should be driven down so that mainly gene-elicited, environmentally defined mutant phenotypes are measured, it is surprising that only recently the impact of these important confounding factors has been realized. These shortcomings led the Eumorphia program to standardize and validate many phenotyping tests, linking them to standardized operating procedures

Table 1. Examples of metabolic phenotyping tests and their standard operating procedures (SOPs)

	Phenotype test	SOP website link[a]
Energy expenditure	Body weight, food, and water intake	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/4_005_0.pdf
	Body composition	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/4_002_0.pdf
	Cold test	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/4_001_0.pdf
Serum parameters	Cholesterol	
	Triglycerides	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/3_002_0.pdf
	Free fatty acids	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/3_003_0.pdf
	Lipoproteins	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/3_004_0.pdf
	Apolipoproteins	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/3_A01_0.pdf
	Bile Acids	
Glucose tolerance	Intraperitoneal glucose tolerance test	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/4_004_0.pdf
	Meal tolerance test	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/4_007_0.pdf
	Oral glucose tolerance test	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/4_008_0.pdf
Ex vivo	General histology techniques	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/15_001_0.pdf
	H&E staining	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/15_016_0.pdf
	Oil Red O staining	http://www.empress.har.mrc.ac.uk/EMPReSS/pdf/15_026_0.pdf

[a] These protocols and others can be found on the EMPReSS website (http://www.empress.har.mrc.ac.uk/EMPReSS/servlet/EMPReSS.Frameset) (Green et al., 2005). A list of authors and contributors to EMPReSS SOPs and phenotype data is found at http://www.empress.har.mrc.ac.uk/EMPReSS/area/authors.pdf.

(SOPs), to ensure comparable test outcomes between different laboratories and ensure more robust and high-quality phenotypes (Tables 1 and 2; Green et al., 2005). Despite these efforts, however, obtaining exactly the same outcomes wherever and whenever testing takes place is still challenging given the inevitable variability in biological systems as a consequence of "noise" in gene expression, even under supposedly identical conditions within populations (Raser and O'Shea, 2005).

Subtle metabolic phenotypes, as revealed by meticulous and standardized phenotyping, are perhaps the most relevant since they reflect the insidious nature of common metabolic diseases in men. For example, the liver X receptor (LXR) was initially suspected to control cholesterol homeostasis through QTL-mapping studies that linked Cyp7a1, the rate-limiting enzyme in bile acid synthesis and an LXR target gene, with the control of serum HDL levels and cholesterol absorption (Machleder et al., 1997; Schwarz et al., 2001). However, the full "homeostatic potential" of these genes was not fully appreciated until $LXR^{-/-}$ mice, which are seemingly normal on regular chow, were reported to accumulate cholesterol in their livers when fed a cholesterol-rich diet (Peet et al., 1998). This was explained by their inability to stimulate bile acid synthesis and cholesterol excretion in response to cholesterol loading, underscoring the crucial dependence of Cyp7a1 regulation on LXR (Peet et al., 1998) and its competence factor the nuclear receptor, liver receptor homolog-1 (LRH-1; Goodwin et al., 2000; Lu et al., 2000). As bile-acid-pool size increases, a feedback mechanism is evoked, which involves the activation of another nuclear receptor, farnesoid X receptor (FXR), leading to the induction of the short heterodimer protein (SHP; Sinal et al., 2000). Elevated SHP then inactivates LRH-1 and LXR, leading to promoter-specific repression of Cyp7a1, thus completing an elaborate autoregulatory loop, which maintains hepatic cholesterol homeostasis (Brendel et al., 2002; Goodwin et al., 2000; Lu et al., 2000). However, only when $SHP^{-/-}$ mice were pharmacologically challenged with synthetic or natural FXR agonists was it shown that it takes even more than FXR, SHP, and LRH-1 to inhibit bile-acid synthesis (Kerr et al., 2002; Wang et al., 2002). Using the same

Table 2. Mouse genome-related internet resources

Description	Website
Edinburgh Mouse Atlas (emap) and gene expression database (emage)	http://genex.hgu.mrc.ac.uk
EMPReSS—European mouse phenotyping resource of standardized screens (SOPs)[a]	http://www.empress.har.mrc.ac.uk/EMPReSS/servlet/EMPReSS.Frameset
EU Eumorphia program[a]	http://www.eumorphia.org
European Mouse Mutant Archive (EMMA)—transgenic mouse repository	http://www.emma.rm.cnr.it
Festing's Inbred Strain Characteristics	http://www.informatics.jax.org
Insitut Clinique de la Souris (ICS)	http://www-mci.u-strasbg.fr/
International Mouse Strain Resource—aims to list all publicly available mouse strains[a]	http://www.informatics.jax.org/imsr/index.jsp
Mouse Genome Informatics—provides integrated access to data on the genetics, genomics, and biology of the laboratory mouse	http://www.informatics.jax.org
Mouse Knockout and Mutation Database—searchable database for phenotypic information related to GEMMs	http://www.biomednet.com/db/mkmd
Mymouse.org—aims to facilitate international collaborative research among mouse users	http://www.mymouse.org
Online Mendelian Inheritance in Man (OMIM)—catalogs human phenotypes and genotypes and relevant mouse models	http://www.ncbi.nlm.nih.gov
TBASE—database for transgenic and knockout mice	http://tbase.jax.org/
The GeneNetwork—bioinformatic resources for systems genetics and complex trait analysis in mouse, rat, and *Arabidopsis*[a]	http://www.genenetwork.org/
US Mouse Phenome Project[a]	http://www.jax.org/phenome
Whole Mouse Catalog—information website	http://www.rodentia.com/wmc

[a] Internet resources highlighted in text.

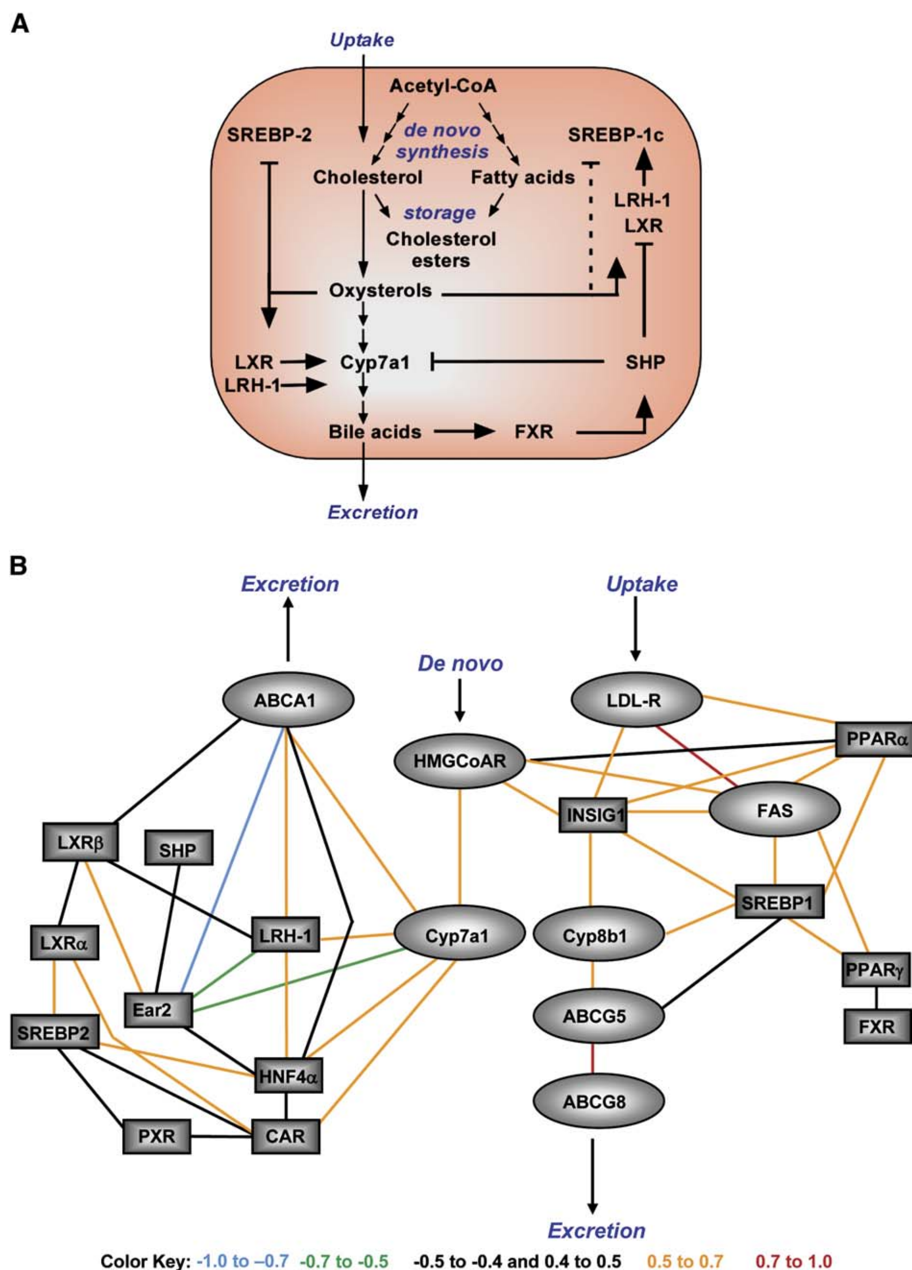


Figure 3. Exploitation of complex information for systems biology approaches

A) The metabolic reconstruction of lipid homeostasis according to the combined analysis of QTL mapping, GEMMs, and pharmacological studies.

B) The "computed" lipid network using the GeneNetwork function association network, which transforms gene-expression correlations of a user-defined list of genes into a network of statistical associations. The list of genes was derived from those identified by wet biology (see [**A**]), and the genetic correlations were based on the liver-gene expression database of the B6BTBRF2 mouse cross. The lines between the nodes represent Pearson correlation coefficients, with higher absolute values being of greater significance as reflected in the color of the line. For detailed information, see text and http://www.genenetwork.org). Abbreviations are given in the text except for the following: ABC, ATP binding cassette protein; CAR, constitutive androstane receptor; PXR, pregnane X receptor; HMGCoAR, 3-hydroxy-3-methylglutaryl-coenzyme A reductase; LDL-R, low-density lipoprotein receptor; HNF, hepatic nuclear factor; and FAS, fatty acid synthase.

series of knockout mice, the importance of FXR/SHP/LXR to control SREBP-1c and triglyceride metabolism was also demonstrated (Watanabe et al., 2004; Figure 3A). Mice deficient in the transcription intermediary factor 2 (TIF2) also presented a silent metabolic phenotype (Picard et al., 2002). $TIF2^{-/-}$ are normal under chow-fed conditions but are remarkably protected against diet-induced obesity due to enhanced adaptive thermogenesis, identifying this p160 coregulator as a modulator of energy homeostasis. Similarly, the phenotype of mice that express an activated form of PPARβ/δ in skeletal muscle was only revealed subsequent to endurance exercise training (Wang et al., 2004). In fact, PPARβ/δ promoted oxidative muscle fiber transformation and mitochondrial biogenesis, enabling transgenic mice to run longer. Since all these phenotypes were only unveiled after challenges (diet in $LXR^{-/-}$ and $TIF2^{-/-}$, pharmacological in $SHP^{-/-}$, and exercise in $PPAR\beta/\delta^{-/-}$ mice), it is inherent that many genes involved in metabolic control are environmentally sensitive.

Systems metabolism: Putting it all together

Systems biology, as applied to metabolism, aims to convert metabolic parts to metabolic systems. The success of reconstructing metabolic networks will rely on at least two things, how comprehensively we define the metabolic parts in our databases and how well we model these into systems. High throughput strategies have quickly advanced our capacity to catalog various metabolic parts based on interactions between genes and their gene products in terms of genetic control, transcriptional regulation, expression correlations, sequence homology, and protein:protein interactions. Incorporation of data sets derived from the analysis of GRPs and traditional GEMMs will also advance the reconstruction of metabolic networks. The GEMMs

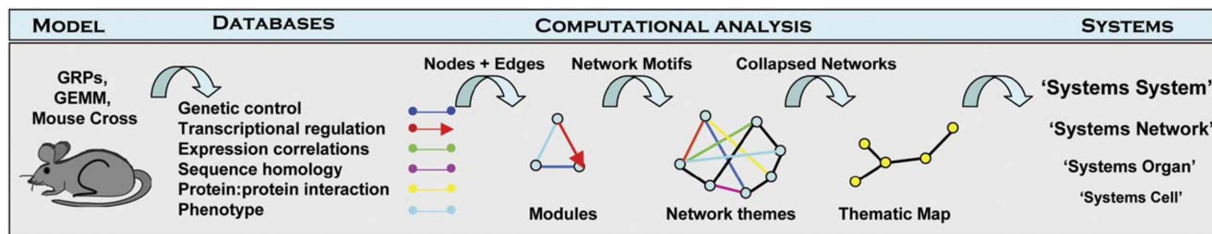


Figure 4. Exploitation of complex information for systems biology approaches

Computational architecture can be based on a modular structure consisting of nodes, or genes, their gene products or phenotype, and edges, or some type of interaction. The edges, depicted simplistically as colored lines, represent multiple and diverse interactions. A three node interconnection pattern is shown as an example, which is then synthesized into enriched multicolor network motifs. These motifs can be combined to form network themes or classes of higher-order recurring interconnection patterns that encompass multiple occurrences of network motifs. These network themes can then be collapsed and assembled into thematic maps to provide a bird's-eye view of the functional relationships between diverse subsystems of the overall biochemical networks. These maps are closely tied to specific biological phenomena. Applying these strategies at multiple levels allows for the building of systems cell, systems tissue, systems network, and finally systems system, inherent structures of higher organisms.

cover the "simplest" basis of disease because they represent an isolated genetic event causing a phenotype, whereas the complex models of GRPs incorporate the natural range of genetic variation to precipitate disease. However, in order to truly convert parts to networks and systems, most of our datasets have to become "full of life" and incorporate the reality that only subsets of these interactions are active at any given point in time under any one particular biological condition. Thus, in order to encapsulate condition-dependent network structures, dynamic enviro-, chemico- and pharmacogenomics data sets derived from exposing mice to environmental (e.g., diet) and chemical (e.g., drugs) challenges, are needed. However, as this requires gathering, storing, and analyzing information of a diverse nature, crossdisciplinary communication and collaboration needs to be improved concurrently with the evolution of the cyberinfrastructure that supports the subsequent integration of these diverse and large-scale metabolic databases into metabolic systems. Existing bioinformatic tools readily support access and storage of information, but the ability and ease to abstract, integrate, graph, and translate this information into knowledge is still developing. It is hoped that in the end the derivation of static networks, their rewiring to fit dynamic conditions, and the ability to extract empirical rules underpinning metabolic organizational principles will be accessible for all biologists (Hey and Trefethen, 2005).

One practical scheme of how complex information can be exploited for systems biology approaches is summarized in Figure 4. The trait databases, supported from any population, from any tissue, from any species, and of any type, serve as the building blocks for the reconstruction of metabolic networks. Current computational architecture is based on modular structures, defined as nodes, which can represent genes, their gene products and phenotypes, and edges, which represent some type of interaction. The edges, depicted simplistically as colored lines, can represent multiple and diverse interactions all the way from genetic control, transcriptional regulation, expression correlation, and sequence homology to protein:protein interactions. In its "simplest" form, a module can be based on single biological interactions such as those generated by association-networking graphing functions, which graphically transform differentially expressed or correlated genes, the nodes, into a network of statistical associations or interactions, the edges (see Figure S1 in the Supplemental Data available with this article online). For example, we can input a partial list of genes identified as regulators of cholesterol homeostasis by wet biology (Figure 3A) and then compute a lipid-gene network using statistically significant gene expression correlations found in the liver of a genetical genomic data set from a B6BTBRF2 mouse cross (Figure 3B; www.genenetwork.org). In fitting with prior literature, in this network there are strong genetic interactions between LRH-1 and Cyp7a1, ABCA1 and LXR, and SREBP-2 and LXR (Davis et al. [2002] and references within). There are, however, also obvious differences in the placement of these nodes as compared to that derived by wet biology (compare Figures 3A and 3B). This differential rewiring likely reflects the modular nature of biological systems and condition-dependent interactions, which, for example, allows for different metabolic compensation in GEMMs and GRPs.

Beyond the modular type analysis is set-type enrichment analysis (SEA) that can be used to evaluate the representation of any a priori-defined module in a given set. Statistical overrepresentation of a set can reveal motifs (Aderem, 2005). For example, a gene-set enrichment analysis (GSEA), determines whether any modular gene set is statistically enriched or overrepresented in a list of genes (Mootha et al., 2003). In the case of GEMMs, the list of genes represents differentially expressed genes between wild-type and mutant mice, whereas in the case of GRPs it could correspond to genes whose expression levels correlate with the gene of interest. The theory is that quantitative gene expression patterns may correspond to functional gene categories within one tissue or even more so, across many different but functionally related tissues (Hughes et al., 2000). GSEA is commonly performed according to gene ontology (GO)-, Biocarta-, and/or KEGG-a priori-defined modules which statistically place genes of interest in structured networks according to molecular function, biological processes, cellular components, and higher-order functional meanings (Figure S2). These types of analyses serve as global-positioning tools that orient the researcher in some meaningful direction in the extensive metabolic atlas inherent of higher organisms. SEA can also be applied to the systematic detection of *cis*-regulatory, *trans*-regulatory, and protein:protein-interacting domains using the concept MotifADE, or motifs associated with differential expression (Giallourakis et al., 2005; Mootha et al., 2004). Such a strategy predicts that modular genes share both a similar expression profile and a similar profile of motifs that either controls their transcription (e.g., regulatory elements in genomic DNA) or their interaction with other proteins (e.g., protein-interaction domains) or DNA. These motifs, either predefined or determined de novo, can be computationally detected in the promoter or protein sequences of a list of genes using enrichment-type

algorithms. Ideally, the culmination of multiple SEAs can help to build and later to navigate the metabolic networks.

Another type of analysis aimed at defining genetic modules and networks is genetical genomics or expression genetics, which involves carrying out genome wide analyses of gene-expression data and then treating individual transcript abundances as quantitative traits to identify QTL(s) that influence their expression levels. The genetic regions or loci that are revealed are termed expression QTL (eQTL; Damerval et al., 1994; Jansen and Nap, 2001; Klose et al., 2002; Brem et al., 2002; Chesler et al., 2003; Schadt et al., 2003). eQTL mapping has established that the basis for the variation in transcript abundance is highly heritable and can be influenced by a locus at or near its own locus (*cis*-acting QTL) or by a completely different genomic loci (*trans*-acting QTL). Once an eQTL is found, these chromosomal regions can be translated into lists of causal and reactive genes for the trait of interest. One potential "filtering" step is to focus on those genes whose expression level correlates with the trait of interest and that are located in the QTL region. Another approach is to cluster groups of correlated gene-expression traits to reveal shared eQTLs or hot spots, an idea that stems from the fact that modular genes share both a similar expression profile and by extension QTLs. These chromosomal regions in turn point toward regulatory elements that affect the expression levels of groups of genes and could provide explanations as to how expression levels of different genes in the same pathway are correlated/regulated. In such an analysis, traits are clustered along one axis by phenotypic similarity, whereas the genomic localization of the QTLs maps, chromosome by chromosome, are represented on the other axis by heat maps that indicate the probability of linkage. The list of traits that is clustered can be of any nature. Chromosomal locations where these genes share common eQTLs or hot spots generally appear as colored lines along the cluster tree and can be readily identified (Figure S3).

Particularly powerful is the combination of eQTL data analyses with complex-trait QTL data since cQTL mapping on its own often includes gross clinical measurements that are far removed from their causative biological processes. The identification of a common chromosomal location for *cis*-acting eQTLs and disease-trait cQTLs can be used to nominate genes in the disease-susceptibility locus, which can bypass the requirement for fine mapping of the region (Schadt et al., 2003). For example, eQTL analysis was performed on a cohort of high-fat-fed C57BL/6XDBA F2 mice to define the obesity trait subcutaneous fat pad mass (FPM; Schadt et al., 2003). The most differentially expressed sets of genes in these mice comprising the upper (FPM-high) and lower (FPM-low) percentiles were compiled, and five hot spots or regions significantly enriched with eQTL were identified. In these, more than 50% of the genes in the FPM set were genetically linked. Furthermore, in the FPM-high group, two distinct expression patterns were identified, indicating "causative" heterogeneity. Thus, defining metabolic phenotypes according to patterns of gene expression or molecular phenotype allows for the classification of subpopulations that are homogenous with respect to the underlying causes of metabolic disease. Computational analysis aimed at providing likelihood-based tests for causality to extract information of possible upstream modulators of gene transcripts from *cis*- and *trans*-acting eQTL and cQTL data is continuing to evolve and holds great promise for unraveling complex traits, including obesity (Li et al., 2005; Schadt et al., 2005).

Several insights on how to subsequently multidimensionally model the biological interactions beyond statistical terms into a unifying network obedient to fundamental mechanical and physicochemical principles have been put forth at the level of unicellular organisms such as *S. cerevisiae*. In *S. cerevisiae*, five color-coded data sets were recently decomposed into three or four node-interconnection patterns representing multiple interaction types and were then synthesized into enriched multicolor network motifs (Herrgård and Palsson, 2005; Zhang et al., 2005; Figure 4). These motifs were then combined to form network themes or classes of higher-order recurring interconnection patterns that encompass multiple occurrences of network motifs. Similarly, genetic interaction networks from two primary data sources, genetic and physical interactions, have been assembled into within-pathway models defined by dense interactions within a single group of proteins (Kelley and Ideker, 2005). Between-pathway models were then built from within-pathways, which were defined by enriched or dense genetic interactions connecting two separate nonoverlapping groups of proteins, where each group was densely connected by physical interactions. By collapsing these network themes and/or between-pathway models and then assembling them into thematic maps, a bird's-eye view of functional relationships between diverse subsystems of the overall biochemical networks can be revealed. These thematic maps can be interpreted as closely tied to specific biological phenomena and representative of more fundamental network design principles. For example, many of the transcription factors and protein complexes are often linked to the same biological process, and conversely complexes of related function are often linked to the same transcription factor (Herrgård and Palsson, 2005; Zhang et al., 2005).

These types of computational strategies are well on their way to being successful at a "systems cell" level, in that they will be able to synthesize and decompose the over 150,000 functions and physical interactions in yeasts into integrated biological networks within a single yeast cell. In higher organisms, there are additional challenges to consider, in that systems cells must be built into the context of systems organs, systems networks, and finally into systems systems. For instance, a hepatocyte is part of the liver, which coordinates with other organs and their networks the fate of metabolism on a whole-organism basis. Although we can envision a similar computational development to model metabolism in complex organisms, computational metabolic navigation in mammals is still several mouse clicks away. Although systems biology approaches are still in their development stages, several user-friendly internet resources already exist that support these types of efforts (Table 1 summarizes several useful mouse resources). To demonstrate their potential, we invite the reader to navigate an example found in the Supplemental Data.

Future prospects

Techniques in mouse genetics and phenomics have rapidly evolved over the past decade, which with the renaissance of integrative physiology positioned scientists well for the emergence of the discipline systems biology. Such systems biology approaches hold great promise in the area of metabolism. The efficient conversion of metabolic parts to metabolic systems will depend on the integration of discovery-based and hypothesis-driven phenogenomic approaches, which rely on the study of mouse models derived from both reverse- and

forward-genetic strategies and complex models of GRPs. A stable marriage of in silico with wet biology that allows for reiterative testing of the computationally identified networks will be crucial for the success of systems biology approaches. Such integrated approaches will become indispensable future resources to understand complex metabolic networks. It is expected that such insight will ultimately translate into better diagnostic and therapeutic strategies for metabolic diseases, which have acquired epidemic proportions.

Supplemental data

Supplemental Data include three figures and can be found with this article online at http://www.cellmetabolism.org/cgi/content/full/2/6/349/DC1/.

Acknowledgments

We thank Robert Williams from GeneNetwork for his invaluable input and assistance and Alan Attie, Christina Kendziorski, Brian Yandell, and Hong Lan at the University of Wisconsin-Madison for the B6XBTBRF2-ob/ob liver mRNA M430AB array data. We also thank members of the Auwerx lab for their helpful discussions and CNRS, Inserm, Hopitaux Universitaires de Strasbourg, European Union (QLRT-2001-00930 and LSHM-CT-2004-512013), and the NIH (POI-DK59820-01 and ROI-DK067320-01) for grant support. C.A.A. is supported by a HSFC fellowship.

References

Accili, D., Drago, J., Lee, E.J., Johnson, M.D., Cool, M.H., Salvatore, P., Asico, L.D., Jose, P.A., Taylor, S.I., and Westphal, H. (1996). Early neonatal death in mice homozygous for a null allele of the insulin receptor gene. Nat. Genet. *12*, 106–109.

Aderem, A. (2005). Systems biology: its practice and challenges. Cell *121*, 511–513.

Austin, C.P., Battey, J.F., Bradley, A., Bucan, M., Capecchi, M., Collins, F.S., Dove, W.F., Duyk, G., Dymecki, S., Eppig, J.T., et al. (2004). The knockout mouse project. Nat. Genet. *36*, 921–924.

Auwerx, J., Avner, P., Baldock, R., Ballabio, A., Balling, R., Barbacid, M., Berns, A., Bradley, A., Brown, S., Carmeliet, P., et al. (2004). The European dimension for the mouse genome mutagenesis program. Nat. Genet. *36*, 925–927.

Bickel, P.E. (2004). Metabolic fuel selection: the importance of being flexible. J. Clin. Invest. *114*, 1547–1549.

Branda, C., and Dymecki, S. (2004). Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice. Dev. Cell *6*, 7–28.

Brem, R.B., Yvert, G., Clinton, R., and Kruglyak, L. (2002). Genetic dissection of transcriptional regulation in budding yeast. Science *296*, 752–755.

Brendel, C., Schoonjans, K., Botrugno, O.A., Treuter, E., and Auwerx, J. (2002). The small heterodimer partner interacts with liver X receptor alpha and represses its transcriptional activity. Mol. Endocrinol. *16*, 2065–2076.

Bruning, J.C., Michael, M.D., Winnay, J.N., Hayashi, T., Horsch, D., Accili, D., Goodyear, L.J., and Kahn, C.R. (1998). A muscle-specific insulin receptor knockout exhibits features of the metabolic syndrome of NIDDM without altering glucose tolerance. Mol. Cell *2*, 559–569.

Cederberg, A., Grönning, L.M., and Ahrén, B. (2001). FOXC2 is a winged helix gene that counteracts obesity, hypertriglyceridemia, and diet-induced insulin resistance. Cell *106*, 563–573.

Champy, M., Selloum, M., Piard, L., Zietler, V., Caradec, C., Chambon, P., and Auwerx, J. (2004). Mouse functional genomics requires standardization of mouse handling and housing conditions. Mamm. Genome *15*, 768–783.

Chen, H., Charlat, O., Tartaglia, L.A., Woolf, E.A., Weng, X., Ellis, S.J., Lakey, N.D., Culpepper, J., Moore, K.J., Breitbart, R.E., et al. (1996). Evidence that the diabetes gene encodes the leptin receptor: identification of a mutation in the leptin receptor gene in db/db mice. Cell *84*, 491–495.

Chesler, E.J., Wang, J., Lu, L., Qu, Y., Manly, K.F., and Williams, R.W. (2003). Genetic correlates of gene expression in recombinant inbred strains: a relational model system to explore neurobehavioral phenotypes. Neuroinformatics *1*, 343–357.

Churchill, G.A., Airey, D.C., Allayee, H., Angel, J.M., Attie, A.D., Beatty, J., Beavis, W.D., Belknap, J.K., Bennett, B., Berrettini, W., et al. (2004). The collaborative cross, a community resource for the genetic analysis of complex traits. Nat. Genet. *36*, 1133–1137.

Cock, T.A., Back, J., Elefteriou, F., Karsenty, G., Kastner, P., Chan, S., and Auwerx, J. (2004). Enhanced bone formation in lipodystrophic PPARgamma (hyp/hyp) mice relocates haematopoiesis to the spleen. EMBO Rep. *5*, 1007–1012.

Colinayo, V.V., Qiao, J., Wang, X., Krass, K.L., Schadt, E., Lusis, A.J., and Drake, T.A. (2003). Genetic loci for diet-induced atherosclerotic lesions and plasma lipids in mice. Mamm. Genome *14*, 464–471.

Collins, S.C., Wallis, R.H., Wallace, K., Bihoreau, M.T., and Gauguier, D. (2003). Marker-assisted congenic screening (MACS): a database tool for the efficient production and characterization of congenic lines. Mamm. Genome *14*, 350–356.

Damerval, C., Maurice, A., Josse, J.M., and de Vienne, D. (1994). Quantitative trait loci underlying gene product variation: a novel perspective for analyzing regulation of genome expression. Genetics *137*, 289–301.

Davis, R.A., Miyake, J.H., Hui, T.Y., and Spann, N.J. (2002). Regulation of cholesterol-7alpha-hydroxylase: BAREly missing a SHP. J. Lipid Res. *43*, 533–543.

Dykxhoorn, D.M., Novina, C.D., and Sharp, P.A. (2003). Killing the messenger: short RNAs that silence gene expression. Nat. Rev. Mol. Cell Biol. *4*, 457–467.

Giallourakis, C., Henson, C., Reich, M., Xie, X., and Mootha, V.K. (2005). Disease gene discovery through integrative genomics. Annu. Rev. Genomics Hum. Genet. *6*, 381–406.

Goodwin, B., Jones, S.A., Price, R.R., Watson, M.A., McKee, D.D., Moore, L.B., Galardi, C., Wilson, J.G., Lewis, M.C., Roth, M.E., et al. (2000). A regulatory cascade of the nuclear receptors FXR, SHP-1, and LRH-1 represses bile acid biosynthesis. Mol. Cell *6*, 517–526.

Green, E.C., Gkoutos, G.V., Lad, H.V., Blake, A., Weekes, J., and Hancock, J.M. (2005). EMPReSS: European mouse phenotyping resource for standardized screens. Bioinformatics *21*, 2930–2931.

Guo, G., Wang, W., and Bradley, A. (2004). Mismatch repair genes identified using genetic screens in Blm-deficient embryonic stem cells. Nature *429*, 891–895.

Halaas, J.L., Gajiwala, K.S., Maffei, M., Cohen, S.L., Chait, B.T., Rabinowitz, D., Lallone, R.L., Burley, S.K., and Friedman, J.M. (1995). Weight-reducing effects of the plasma protein encoded by the obese gene. Science *269*, 543–546.

He, W., Barak, Y., Hevener, A., Olson, P., Liao, D., Le, J., Nelson, M., Ong, E., Olefsky, J.M., and Evans, R.M. (2003). Adipose-specific peroxisome proliferator-activated receptor gamma knockout causes insulin resistance in fat and liver but not in muscle. Proc. Natl. Acad. Sci. USA *100*, 15712–15717.

Herrgård, M.J., and Palsson, B.O. (2005). Untangling the web of functional and physical interactions in yeast. J. Biol. *4*, 5.

Hey, T., and Trefethen, A.E. (2005). Cyberinfrastructure for e-science. Science *308*, 817–821.

Hughes, T.R., Marton, M.J., Jones, A.R., Roberts, C.J., Stoughton, R., Armour, C.D., Bennett, H.A., Coffey, E., Dai, H., He, Y.D., et al. (2000). Functional discovery via a compendium of expression profiles. Cell *102*, 109–126.

Imai, T., Takakuwa, R., Marchand, S., Dentz, E., Bornert, J.M., Messaddeq, N., Wendling, O., Mark, M., Desvergne, B., Wahli, W., et al. (2004). Peroxisome proliferator-activated receptor gamma is required in mature white and brown adipocytes for their survival in the mouse. Proc. Natl. Acad. Sci. USA *101*, 4543–4547.

Inoue, M., Sakuraba, Y., Motegi, H., Kubota, N., Toki, H., Matsui, J., Toyoda, Y., Miwa, I., Terauchi, Y., Kadowaki, T., et al. (2004). A series of maturity onset diabetes of the young, type 2 (MODY2) mouse models generated by a large-scale ENU mutagenesis program. Hum. Mol. Genet. *13*, 1147–1157.

Jansen, R.C., and Nap, J.P. (2001). Genetical genomics: the added value from segregation. Trends Genet. *17*, 388–391.

Katz, E.B., Stenbit, A.E., Hatton, K., DePinho, R., and Charron, M.J. (1995). Cardiac and adipose tissue abnormalities but not diabetes in mice deficient in GLUT4. Nature *377*, 151–155.

Kelley, R., and Ideker, T. (2005). Systematic interpretation of genetic interactions using protein networks. Nat. Biotechnol. *23*, 561–566.

Kerr, T.A., Saeki, S., Schneider, M., Schaefer, K., Berdy, S., Redder, T., Shan, B., Russell, D.W., and Schwarz, M. (2002). Loss of nuclear receptor SHP impairs but does not eliminate negative feedback regulation of bile acid synthesis. Dev. Cell *2*, 713–720.

Kido, Y., Burks, D.J., Withers, D., Bruning, J.C., Kahn, C.R., White, M.F., and Accili, D. (2000). Tissue-specific insulin resistance in mice with mutations in the insulin receptor, IRS-1, and IRS-2. J. Clin. Invest. *105*, 199–205.

Klose, J., Nock, C., Herrmann, M., Stühler, K., Marcus, K., Bluggel, M., Krause, E., Schalkwyk, L.C., Rastan, S., Brown, S.D., et al. (2002). Genetic analysis of the mouse brain proteome. Nat. Genet. *30*, 385–393.

Koutnikova, H., Cock, T.A., Watanabe, M., Houten, S.M., Champy, M.F., Dierich, A., and Auwerx, J. (2003). Compensation by the muscle limits the metabolic consequences of lipodystrophy in PPAR gamma hypomorphic mice. Proc. Natl. Acad. Sci. USA *100*, 14457–14462.

Kuriyama, H., Liang, G., Engelking, L.J., Horton, J.D., Goldstein, J.L., and Brown, M.S. (2005). Compensatory increase in fatty acid synthesis in adipose tissue of mice with conditional deficiency of SCAP in liver. Cell Metab. *1*, 41–51.

Li, R., Lyons, M.A., Wittenburg, H., Paigen, B., and Churchill, G.A. (2005). Combining data from multiple inbred line crosses improves the power and resolution of QTL mapping. Genetics *169*, 1699–1709.

Lu, T.T., Makishima, M., Repa, J.J., Schoonjans, K., Kerr, T.A., Auwerx, J., and Mangelsdorf, D.J. (2000). Molecular basis for feedback regulation of bile acid synthesis by nuclear receptors. Mol. Cell *6*, 507–515.

Machleder, D., Ivandic, B., Welch, C., Castellani, L., Reue, K., and Lusis, A.J. (1997). Complex genetic control of HDL levels in mice in response to an atherogenic diet. Coordinate regulation of HDL levels and bile acid metabolism. J. Clin. Invest. *99*, 1406–1419.

Metzger, D., and Chambon, P. (2001). Site- and time-specific gene targeting in the mouse. Methods *24*, 71–80.

Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. *34*, 267–273.

Mootha, V.K., Handschin, C., Arlow, D., Xie, X., St Pierre, J., Yang, W., Altshuler, D., Puigserver, P., Patterson, N., Willy, P.J., et al. (2004). Erralpha and Gabpa/b specify PGC-1alpha-dependent oxidative phosphorylation gene expression that is altered in diabetic muscle. Proc. Natl. Acad. Sci. USA *101*, 6570–6575.

Nandi, A., Kitamura, Y., Kahn, C.R., and Accili, D. (2004). Mouse models of insulin resistance. Physiol. Rev. *84*, 623–647.

O'Brien, T.P., and Frankel, W.N. (2003). Moving forward with chemical mutagenesis in the mouse. J. Physiol. *554*, 13–21.

O'Doherty, A., Ruf, S., Mulligan, C., Hildreth, V., Errington, M.L., Cooke, S., Sesay, A., Modino, S., Vanes, L., Hernandez, D., et al. (2005). An aneuploid mouse strain carrying human chromosome 21 with Down syndrome phenotypes. Science *309*, 2033–2037.

Paigen, K. (2003). Anecdotal, historical and cricital commentaries on genetics. Genetics *163*, 1227–1235.

Paigen, B., Morrow, A., Brandon, C., Mitchell, D., and Holmes, P. (1985). Variation in susceptibility to atherosclerosis among inbred strains of mice. Atherosclerosis *7*, 65–73.

Peet, D.J., Turley, S.D., Ma, W., Janowski, B.A., Lobaccaro, J.M., Hammer, R.E., and Mangelsdorf, D.J. (1998). Cholesterol and bile acid metabolism are impaired in mice lacking the nuclear oxysterol receptor LXRα. Cell *93*, 693–704.

Picard, F., Géhin, M., Annicotte, J., Rocchi, S., Champy, M.F., O'Malley, B.W., Chambon, P., and Auwerx, J. (2002). SRC-1 and TIF2 control energy balance between white and brown adipose tissue. Cell *111*, 931–941.

Rangwala, S.M., Rhoades, B., Shapiro, J.S., Rich, A.S., Kim, J.K., Schulman, G.I., Kaestner, K.H., and Lazzar, M.A. (2003). Genetic modulation of PPARg phosphorylation regulates insulin sensitivity. Dev. Cell *5*, 657–663.

Raser, J.M., and O'Shea, E.K. (2005). Noise in gene expression: origins, consequences, and control. Science *309*, 2010–2013.

Reifsnyder, P.C., and Leiter, E.H. (2002). Deconstructing and reconstructing obesity-induced diabetes (diabesity) in mice. Diabetes *5*, 825–832.

Schadt, E.E., Monks, S.A., Drake, T.A., Lusis, A.J., Che, N., Colinayo, V., Ruff, T.G., Milligan, S.B., Lamb, J.R., Cavet, G., et al. (2003). Genetics of gene expression surveyed in maize, mouse and man. Nature *422*, 297–302.

Schadt, E.E., Lamb, J., Yang, X., Zhu, J., Edwards, S., Guhathakurta, D., Sieberts, S.K., Monks, S., Reitman, M., Zhang, C., et al. (2005). An integrative genomics approach to infer causal associations between gene expression and disease. Nat. Genet. *37*, 710–717.

Schwarz, M., Davis, D.L., Vick, B.R., and Russell, D.W. (2001). Genetic analysis of intestinal cholesterol absorption in inbred mice. J. Lipid Res. *42*, 1801–1811.

Seong, E., Saunders, T.L., Stewart, C.L., and Burmeister, M. (2004). To knockout in 129 or in C57BL/6: that is the question. Trends Genet. *20*, 59–62.

Sinal, C.J., Tohkin, M., Miyata, M., Ward, J.M., Lambert, G., and Gonzalez, F.J. (2000). Targeted disruption of the nuclear receptor FXR/BAR impairs bile acid and lipid homeostasis. Cell *102*, 731–744.

Singer, J.B., Hill, A.E., Burrage, L.C., Olszens, K.R., Song, J., Justice, M., O'Brien, W.E., Conti, D.V., Witte, J.S., Lander, E.S., and Nadeau, J.H. (2004). Genetic dissection of complex traits with chromosome substitution strains of mice. Science *304*, 445–448.

Skarnes, W.C. (2005). Two ways to trap a gene in mice. Proc. Natl. Acad. Sci. USA *102*, 13001–13002.

Soutschek, J., Akinc, A., Bramlage, B., Charisse, K., Constien, R., Donoghue, M., Elbashir, S., Geick, A., Hadwiger, P., Harborth, J., et al. (2004). Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. Nature *432*, 173–178.

Threadgill, D.W., Hunter, K.W., and Williams, R.W. (2002). Genetic dissection of complex and quantitative traits: from fantasy to reality via a community effort. Mamm. Genome *13*, 175–178.

Toye, A.A., Moir, L., Hugill, A., Bentley, L., Quarterman, J., Mijat, V., Hough, T., Goldsworthy, M., Haynes, A., Hunter, A.J., et al. (2004). A new mouse model of type 2 diabetes, produced by N-ethyl-nitrosourea mutagenesis, is the result of a missense mutation in the glucokinase gene. Diabetes *53*, 1577–1583.

Tsai, Y.S., Kim, H.J., Takahashi, N., Kim, H.S., Hagaman, J.R., Kim, J.K., and Maeda, N. (2004). Hypertension and abnormal fat distribution but not insulin resistance in mice with P465L PPARgamma. J. Clin. Invest. *114*, 240–249.

Wang, L., Lee, Y.K., Bundman, D., Han, Y., Thevananther, S., Kim, C.S., Chua, S.S., Wei, P., Heyman, R.A., Karin, M., and Moore, D.D. (2002). Redundant pathways for negative feedback regulation of bile acid production. Dev. Cell *2*, 721–731.

Wang, Y.X., Zhang, C.L., Yu, R.T., Cho, H.K., Nelson, M.C., Bayuga-Ocampo, C.R., Ham, J., Kang, H., and Evans, R.M. (2004). Regulation of muscle fiber type and running endurance by PPARdelta. PLoS Biol. *2*, e294 10.1371/journal.pbio.0020294.

Watanabe, M., Houten, S.M., Wang, L., Moschetta, A., Mangelsdorf, D.J., Heyman, R.A., Moore, D.D., and Auwerx, J. (2004). Bile acids lower triglyceride levels via a pathway involving FXR, SHP, and SREBP-1c. J. Clin. Invest. *113*, 1408–1418.

Williams, R.W., Gu, J., Qi, S., and Lu, L. (2001). The genetic structure of recombinant inbred mice: high-resolution consensus maps for complex trait analysis. Genome Biol. *2*, RESEARCH0046.

Yusa, K., Horie, K., Kondoh, G., Kouno, M., Maeda, Y., Kinoshita, T., and Takeda, J. (2004). Genome-wide phenotype analysis in ES cells by regulated disruption of Bloom's syndrome gene. Nature *429*, 896–899.

Zhang, L.V., King, O.D., Wong, S.L., Goldberg, D.S., Tong, A.H., Lesage, G., Andrews, B., Bussey, H., Boone, C., and Roth, F.P. (2005). Motifs, themes and thematic maps of an integrated Saccharomyces cerevisiae interaction network. J. Biol. *4*, 6.