Genes Nutr (2013) 8:19–27
DOI 10.1007/s12263-012-0313-7

REVIEW

Current status on genome–metabolome-wide associations: an opportunity in nutrition research

Ivan Montoliu · Ulrich Genick · Mirko Ledda · Sebastiano Collino · François-Pierre Martin · Johannes le Coutre · Serge Rezzi

Received: 15 February 2012 / Accepted: 2 August 2012 / Published online: 16 October 2012

Abstract Genome-wide association studies (GWASs) have become a very important tool to address the genetic origin of phenotypic variability, in particular associated with diseases. Nevertheless, these types of studies provide limited information about disease etiology and the molecular mechanisms involved. Recently, the incorporation of metabolomics into the analysis has offered novel opportunities for a better understanding of disease-related metabolic deregulation. The pattern emerging from this work is that gene-driven changes in metabolism are prevalent and that common genetic variations can have a profound impact on the homeostatic concentrations of specific metabolites. A particularly interesting aspect of this work takes into account interactions of environment and lifestyle with the genome and how this interaction translates into changes in the metabolome. For instance, the role of PYROXD2 in trimethylamine metabolism points to an interaction between host and microbiome genomes (host/microbiota). Often, these findings reveal metabolic deregulations, which could eventually be tuned with a nutritional intervention. Here we review the development of gene–metabolism association studies from a single-gene/single-metabolite to a genome-wide/metabolome-wide approach and highlight the conceptual changes associated with this ongoing transition. Moreover, we report some of our recent GWAS results on a cohort of 265 individuals from an ethnically diverse population that validate and refine previous findings on gene–urine metabolism interactions. Specifically, our results confirm the effect of PYROXD2 polymorphisms on trimethylamine metabolism and suggest that a previously reported association of *N*-acetylated compounds with the ALMS1/NAT8 locus is driven by SNPs in the ALMS1 gene.

Keywords Metabolome wide associations · Genome wide associations · Metabolomics · Nutrition

Ivan Montoliu and Ulrich Genick contributed equally to this work.

Special section: "Foodomics"; Guest Editors Dr. A. Bordoni and F. Capozzi.

Electronic supplementary material The online version of this article (doi:10.1007/s12263-012-0313-7) contains supplementary material, which is available to authorized users.

I. Montoliu · S. Collino · F.-P. Martin · S. Rezzi (✉)
Nestlé Research Center, Bioanalytical Science, Nestec Ltd., 1000 Lausanne 26, Switzerland
e-mail: serge.rezzi@rdls.nestle.com; Serge.Rezzi@rd.nestle.com

U. Genick · M. Ledda · J. le Coutre
Nestlé Research Center, Perception Physiology, Nestec Ltd., 1000 Lausanne 26, Switzerland

Present Address:
S. Collino · F.-P. Martin · S. Rezzi
Nestlé Institute of Health Sciences SA, Campus EPFL, Quartier de l'innovation, bâtiment G, 1015 Lausanne, Switzerland

J. le Coutre
Organization for Interdisciplinary Research Projects, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

Introduction

One of the important aspects of nutrition and health today is to determine the biochemical effects of diets on individuals' metabolism and to unravel the underlying mechanism of action. However, this task is very challenging because of the web of interactions among food bioactives, but also the complex mosaic of both genomic and

metagenomic (i.e., gut microbiota) effects and environmental factors. Indeed, both system-wide (i.e., whole organism) and organ-specific changes in biochemical processes have components driven by these factors (Martin et al. 2009a, b; Nicholson et al. 2005; Claus et al. 2011; Mestdagh et al. 2011; Merrifield et al. 2011; Wikoff et al. 2009).

It is widely considered that the identification of metabolic signatures associated with specific genotypes in free-living populations remains challenging due to the relatively low amplitude of these associations when compared to inherent intra- and inter-individual variability that results from dominant factors (i.e., lifestyle, food habits and aging). In spite of this variability, studies that match metabolic profiles and genotype have been reported. These studies are usually based on large cohorts recruited for other purposes, such as epidemiological studies in general, but also on areas such as aging, environment and disease. Moreover, GWAS with metabolic traits is always dependent on the set of metabolites analyzed. This is critical and is usually driven by previous knowledge in the goals of the main study. In general most of the studies include families of compounds such as fatty acids, carnitines, sphingomyelins and amino acids. The amounts of these circulating compounds and their regulatory pathways are likely to be affected by specific nutritional interventions.

Genome-wide association studies (GWASs)

Over the last 10 years, miniaturization, automation and massive fabrication have dramatically decreased the cost of microarray-based whole-genome genotyping (Hoheisel 2006) with so-called genotyping chips. With more than one million of carefully selected genetic probes, these modern genotyping chips are able to explore the vast majority of common genetic variations in humans (Li et al. 2008). With these rapid advances, it has been possible to extend genotype–phenotype association studies from specific target genes to the entire genome leading to the development of so-called genome-wide association studies (GWASs) (McCarthy and Hirschhorn 2008; McCarthy et al. 2008). For a GWAS all study participants are genotyped with a genotyping chip and parallel statistical association tests between the genotype at each of the $\sim$1 million probed genetic sites, and the phenotype of interest are performed. By analyzing the strength of the association along the entire genome, it may then be possible to identify a gene, or even a specific region within a gene, that drives variation in the phenotypic trait. An important consequence of performing so many independent association tests (i.e., one test for each of the one million genotyped sites) is that associations with very low p-values are expected to occur by chance. Multiple-testing correction procedures, such as Bonferroni and false discovery rate corrections, are therefore indispensible. Depending on the correction applied, p-values lower than $10^{-7.5}$ are necessary to achieve statistical significance in a typical GWAS. To achieve such strong association signals, GWASs usually require substantially larger cohorts than single-gene association studies.

The development of gene–metabolism studies from single-compound/single-gene studies to a holistic approach

The existence and heritable nature of inter-individual differences in human metabolism and their potential impact on human nutritional requirements has long been recognized. But in the last 5–10 years the study of gene–metabolism interactions has undergone a profound transition in which classical phenotype-to-genotype work on single-metabolite/single-gene associations is increasingly replaced by more holistic approaches that aim to survey the entire genome for drivers of overall metabolic patterns. To illustrate this transition and its implications for the future of gene–metabolism research, we here discuss selected examples of successful studies that represent various stages of this transition.

The classic approach to gene–metabolism research is well represented by the work on two metabolic phenotypes, favism and lactose intolerance. Individuals suffering from favism may display symptoms of acute non-immune hemolytic anemia after eating broad beans. Through laborious biochemical work conducted in the first half of the last century (see the review by Beutler (2008) for a historical perspective), it was finally discovered that sufferers of favism have a deficient glucose-6-phosphate dehydrogenase (G6PD) enzyme that results in a malfunction of the pentose phosphate pathway. This pathway is essential for the red blood cells' protection mechanism against oxidative stress. In people who lack G6PD, oxidative stress triggered by the compounds vicine and divicine contained in fava beans can therefore cause irreversible damage to red blood cells, resulting in hemolysis and, potentially, kidney failure.

In the case of lactose intolerance, some individuals carry a genetic variation (Enattah et al. 2002) in the regulatory region of the LCT gene that leads to the continued expression of the lactase enzyme beyond weaning, at which stage the production of this enzyme stops in most mammals. Individuals carrying this variation (i.e., the lactase persistence form of the LCT gene) continue to produce the lactase enzyme and secrete it into the duodenum. There this enzyme breaks down the disaccharide lactose into monosaccharides glucose and galactose monosaccharides, which

can then be absorbed into the body. Individuals who carry only non-persistence variants of the gene stop to make the lactase enzyme, so that the lactose passes through the duodenum undigested and enters into the colon where it is metabolized by resident bacteria. The bacterial fermentation process generates large quantities of gases causing bloating and discomfort.

For both favism and lactose intolerance, the metabolic phenotypes and heritable nature of the metabolic phenotypes are known since antiquity. In this sense the gene–metabolism research on this topic started from a phenotypic observation, and the discovery of the underlying genetic variation represented the endpoint of extensive studies into the biochemistry and physiology of these phenotypes.

A next stage in the evolution of gene–metabolism research is represented by single-gene/single-metabolite studies. In these studies the gene of interest, often an enzyme, as well as the affected metabolite is known prior to the start of the main studies, and the principal goal of the studies is to understand how a specific genetic variation impacts medical outcomes via a specific metabolic mechanism. A typical example is the investigation of inter-individual differences in caffeine metabolism. Caffeine is a widely consumed stimulant, the breakdown of which proceeds primarily via a P450 enzyme (P4501A2) encoded by the CYP1A2 gene. A single nucleotide polymorphism in intron 1 of CYP1A2 appears to influence the inducibility of this gene and the efficiency to metabolize caffeine (Rasmussen et al. 2002; Sachse et al. 1999). These studies aim to understand how genetic differences ultimately cause differences in behavioral and medical phenotypes as diverse as coffee consumption (Cornelis et al. 2007) and reproductive health (Sata et al. 2005). The advantage of this type of study is that statistical analysis is relatively simple to perform and that due to the limited number of individual statistical tests, multiple-testing correction does not dilute a potential association signal, thus allowing the discovery of significant association with comparatively small study panels. The obvious and inherent drawback of the single-gene studies—often conducted on just a single SNP within the gene—is that the effect of other genetic polymorphism that might exist in the genome with eventual bigger functional effects will go unnoticed.

The arrival of the GWAS triggered a radical transformation of gene–metabolism research. With this technology, it becomes possible to perform an unbiased search of the entire human genome to identify all common genetic factors that affect a specific metabolic phenotype. Early GWASs were often directed toward common medical conditions such as cardiovascular disease (CVD) or diabetes. Perhaps as a consequence of this, initial gene–metabolism GWASs focused on metabolic phenotypes such as blood cholesterol or triglyceride levels, for which links to health outcomes were already well established (Kathiresan et al. 2008a, b).

These early GWASs both replicated gene–metabotype association that was previously identified in single-gene/single-metabotype studies and identified new genes that were not previously implicated in the metabolism of the targeted compounds. With the usefulness of the GWAS approach demonstrated, whole-genome genotyping of the participants in medical study cohorts has become common practice. As a result, new GWASs on blood lipid phenotypes are now conducted on cohorts of more than 100,000 subjects. This increase in panel sizes leads to a parallel increase in statistical power, so that more genetic factors can be identified. As a result, 95 genetic loci distributed throughout the genome are now known to be associated with blood lipid phenotypes (Teslovich et al. 2010). Roughly two-thirds of these loci were previously not implicated in blood lipid metabolism. These discoveries have opened up new avenues for mechanistic studies that have since demonstrated the functional relevance of several of these loci and have expanded knowledge on the metabolic and regulatory lipid networks. (For a detailed review of relevant examples see (Kathiresan and Srivastava 2012). GWASs with blood lipid phenotypes had thus a significant impact on cardiovascular health research. Despite the large number of identified genetic loci, the overall variance in blood lipid phenotypes that is explained by the combined genotypes at these loci remains relatively small ($\sim$10 %) (Teslovich et al. 2010). Therefore, prediction of individual blood lipid levels or the design of medical intervention schemes based on an individual's genotype at these loci is not useful for the general population.

Targeted metabolomic phenotypes in GWAS

Up to this point, GWASs have provided a lot of information on the genetic contribution to one specific metabolic phenotype (e.g., the blood concentration of cholesterol), but few about related metabolic phenotypes. To address this point, the number of targeted metabolites from various metabolic pathways was increased. The investigation of gene's impact on broader metabolic patterns rather than just a single compound became possible.

A good example of this update in GWAS strategy can be found in the recent work of Gieger et al. (2008) that investigated broader lipid metabolic patterns with implications for CVDs. The work was conducted on Kooperative Gesundheitsforschung in der Region Augsburg (KORA) F3 GWA study cohort that had undergone whole-genome genotyping with SNP chips.

Targeted metabolic profiles of >300 metabolites were generated from serum samples collected for each individual.

Parallel metabolite concentration measurements were performed with the help of isotope-labeled reference compounds. The study resulted in a list of five strongly associated genes (FADS1, PLEK, PARK2, ANKRD30A and LIPC), which could be partially linked to the expected role of each gene in metabolism. For instance, the FADS1 gene codes for fatty acid delta-5 desaturase, a key enzyme in the metabolism of long-chain polyunsaturated and omega 3/6 fatty acids. The minor allele variant of this gene (rs174548) is associated with a decreased efficiency for the delta-5 desaturase reaction, which in turn results in decreased serum levels of phosphatidylcholines, plasmalogen/plasminogen and phosphatidylinositol. In addition, glycerophospholipids carrying less than 4 unsaturated bonds in their fatty acid chains are increased in individuals who carry the less efficient variant of FADS1, while sphingomyelins are decreased, suggesting a modification in phosphatidylcholines' homeostasis. These observations not only constituted a full replication of the association of this locus with arachidonic acid metabolism as previously reported (Schaeffer et al. 2006; Malerba et al. 2008), but also underpinned changes in the efficiency of the delta-5 desaturase reaction. This study further demonstrated that metabolic phenotype characterization can be improved by the use of ratios of metabolite concentrations. To some extent metabolic ratios reflect the rate of biochemical reactions that link these metabolites. Genetic variants that affect the efficiency of a specific enzyme should be best reflected in the ratio of this enzyme's relative substrates and products rather than in the absolute concentration of either of them. This metabolite ratio approach was introduced by Altmaier et al. (2008), and its utility was demonstrated by a pronounced increase in the strength of the association that was observed for FADS1. Here, changes in the catalytic activity of FADS1 due to its polymorphism may alter the levels of eicosatrienoyl-CoA (C20:3) and arachidonyl-CoA (C20:4), which in turn translates into changes on the PC 36:3 and PC 36:4 concentrations. This use of metabolite concentration ratios also proved to be very efficient in the identification of two new loci: the short-chain acyl-coenzyme A dehydrogenase (SCAD) and medium-chain acyl-coenzyme A dehydrogenase (MCAD). Both genes code for enzymes of fatty acids β-oxidation, but affect substrates of different chain length. The strongest association with the SNP rs2014355 in SCAD is the ratio between two short-chain acyl-carnitines (C3/C4), while the strongest with SNP rs11161510 in MCAD is found for C12/C8 ratios (medium-length chain). The directional effect observed for these polymorphisms (e.g., high concentration of long fatty acids versus reduced concentration of short-chain fatty acids) suggests a reduction in the dehydrogenase activity.

In summary, the use of metabolic ratios in GWAS presents several advantages, such as an increase in gene–phenotype association strength and a better understanding of the metabolic variation associated with the phenotype. However, the use of targeted metabolomics approaches (indeed, traditional in this kind of applications) becomes a limitation as it narrows down the possible phenotype variations to a limited set of compounds.

Holistic metabolomics phenotypes in GWAS

Gene–metabolism research evolved with holistic studies on the genetic and on the phenotypic side. This type of studies is currently ongoing in multiple research groups, and a first report has been recently published (Nicholson et al. 2011). For their study, Nicholson and coworkers collected 1D-NMR data of urine and plasma and divided the resulting NMR data into 526 spectral peaks with the intensity of those peaks serving as input values for 526 parallel GWASs. The division of the spectra into these peaks was performed without prior assignment of the peaks to specific metabolites. Notably, this first study, which was conducted on a rather small panel of less than just 150 subjects, has already resulted in the discovery of two very strong gene–metabolism interactions. The first association is between an *N*-acetylated compound and a genetic locus containing the ALMS1, NAT8, TPRKB and DUSP11 genes. This single locus explains >50 % of the observed variance in the urine concentration of this compound, and rare mutations in the ALMS1 gene are known to generate severe kidney dysfunctions. The second association between urine trimethylamine (TMA) concentration and variants in the PYROXD2 gene explains a somewhat smaller, yet still significant, proportion of the observed variance in the concentration of this metabolite. Special poignancy is lent to this second association by the recent discovery that trimethylamine oxide (TMAO), a metabolic product of TMA, is a major risk factor for the development of CVD (Wang et al. 2011). Briefly, the dietary precursor of TMA, phosphatidylcholine (PC), is found in meat, milk, fish and eggs. Gastric enzymes release the choline moiety from PC, and gut bacteria convert choline into TMA, which is then absorbed into the human body. Part of this TMA is then converted to TMAO by flavin-containing monooxygenases (FMOs) particularly by FMO3 (Holmes et al. 2008) in the human body. While it is not immediately clear how exactly the genetic variation in PYROXD2 influences the amount of urine-secreted TMA, it seems that the metabolism of a precursor to a major metabolic marker for cardiovascular risk deserves further research.

Clearly, the study by Nicholson et al. is just the first study of its kind, but it already suggests that genetic factors with very strong impact on certain human metabolites have remained unnoticed and that holistic genome-wide/

Table 1 Selection of reported gene-metabolism associations

Gene	Association	Metabolic trait	Metabolic process
ACADM	Acyl-Coenzyme A dehydrogenase	Hexanoylcarnitine/oleate	Mitochondrial fatty acids β-oxidation
AGXT2	Alanine-glyoxilate aminotransferase-2	β-Aminoisobutyrate	β-Aminoisobutyrate/pyruvate reaction (Alanine synthesis)
ALMS1	Alstrom syndrome	*N*-Acetylated compounds	–
APOA1 APOC3 APOA4 APOA5	Apolipoprotein cluster	Triglyceride levels and phosphatidylcholine ratios (PC 36:2/PC 38:1)	Protein composition of HDL in plasma
ATP10D	Phospholipid-transporting ATPase	Glucosylceramides 16:0 and 24:1	Ceramide transport
CPS1	Carbamoyl-phosphate synthase	Glycine	Protein/nitrogen metabolism
ELOVL2	Elongase	Eicosapentaenoic acid (EPA), docosapentaenoic acid (DPA), docosahexaenoic acid (DHA)	Elongation of polyunsaturated (n−3) fatty acids
FADS1	Fatty acid delta-5 desaturase affecting polyunsaturated and omega 3/6 long-chain fatty acids (PUFA)	Eicosatrienoyl-CoA (C20:3) arachidonyl-CoA (C20:4)	Regulation of LDL, HDL cholesterol and triglyceride levels
GCKR	Glucose kinase regulator protein	Glucose/mannose	Glucose homeostasis
LASS4	Ceramide synthase	Sphingomyelins 18:0, 20:0, 20:1, Ceramides 20:0	Ceramide synthesis
LIPC	Breaking down triglycerides to mono- and diacylglycerols and fatty acids	Phosphatidylethanolamine PE aa C38:6	Regulation of HDL cholesterol and triglycerides
MCAD	Medium-chain acyl-coenzyme A dehydrogenase	Acyl-carnitines (C12/C8)	Fatty acids β-oxidation
MTNR1B	High-affinity receptor for melatonin	Glucose (fasting)	Inhibitory effect of melatonin on insulin secretion
NAT2	Risk locus for coronary artery disease. Response to drug toxicity	1-Methylxanthine/ 4-acetamidobutanoate; formate/ succinate ratio	Regulation of triglyceride levels
NAT8	Cysteinyl-conjugate *N*-acetyltransferase	*N*-Acetyl compounds	Acetylation of cysteine S-conjugates to mercapturic acids.
PANK1	Pantothenate kinase Coenzyme A synthesis	Insulin	Glucose metabolism
PARK2	Parkin (ligase)	Lysine	Glutamate/aminoacid metabolism
PLEK	Pleckstrin protein	Sphingomyelin C14:0	Protein/lipid interactions
PYROXD2	Pyridine nucleotide-disulfide oxidoreductase	Trimethylamine (urine), dimetylamine (plasma)	Oxidoreductase enzyme
SCAD	Short-chain acyl-coenzyme A dehydrogenase	Acyl-carnitines (C3/C4)	Fatty acids β-oxidation
SGPP1	Sphingosine-1-phosphatase 1	Sphingomyelins 14:0, 15:0, 23:0, 24:0, 22:1, 24:1dihydrosphingomyelin 16:0 and 18:0	Recycling of sphingosine into long-chain ceramides
SLC7A9	Glycoprotein-associated amino acid transporter	Glutarylcarnitine/lysine	Involved in the high-affinity, sodium-independent transport of cystine and neutral and dibasic amino acids
SPTLC3	Serine palmitoyltransferase	Cer16:0, 22:0, 23:0, 24:0, 24:1, saturated and unsaturated ceramides, sphingomyelins 17:0, 16:1	Ceramide synthesis and transport

Gieger et al. 2008; Illig et al. 2010; Nicholson et al. 2011; Suhre et al. 2011

metabolome-wide association studies provide a powerful tool to identify them. Notably, the strength of the observed associations is so great that they easily overcome the multiple-testing corrections necessitated by the large number of parallel GWASs inherent in metabolome-wide studies.

A comparison of these recent studies to those on favism and lactose intolerance discussed above indicates how dramatically gene–metabolism research has changed over the past decade and a half. In the past, linking variations in metabolism to a specific gene was often the endpoint of extensive mechanistic and epidemiological studies. By contrast, the newest generation of whole-genome/whole-metabolome studies is typically the starting point that now guides these studies. In parallel, the nature of the metabolic phenotypes addressed by these studies has changed as well: from acute metabolic phenotypes that can be readily perceived by the affected individual, as is the case for lactose intolerance and favism, to subtle, long-term health-risk phenotypes that are not easily perceived by the affected individual and that might even have evaded detection by routine medical testing (e.g., TMAO levels and CVD). For this latter type of metabolic phenotype, early gene- or metabolome-based identification of individual risk factors will be essential to address these risks before any, potentially irreversible, symptoms manifest themselves.

In summary, the literature illustrates that GWAS is a powerful approach to study the genetic predisposition for metabolic disorders (Table 1 collects key examples of the gene–metabolism interactions that have been discovered or

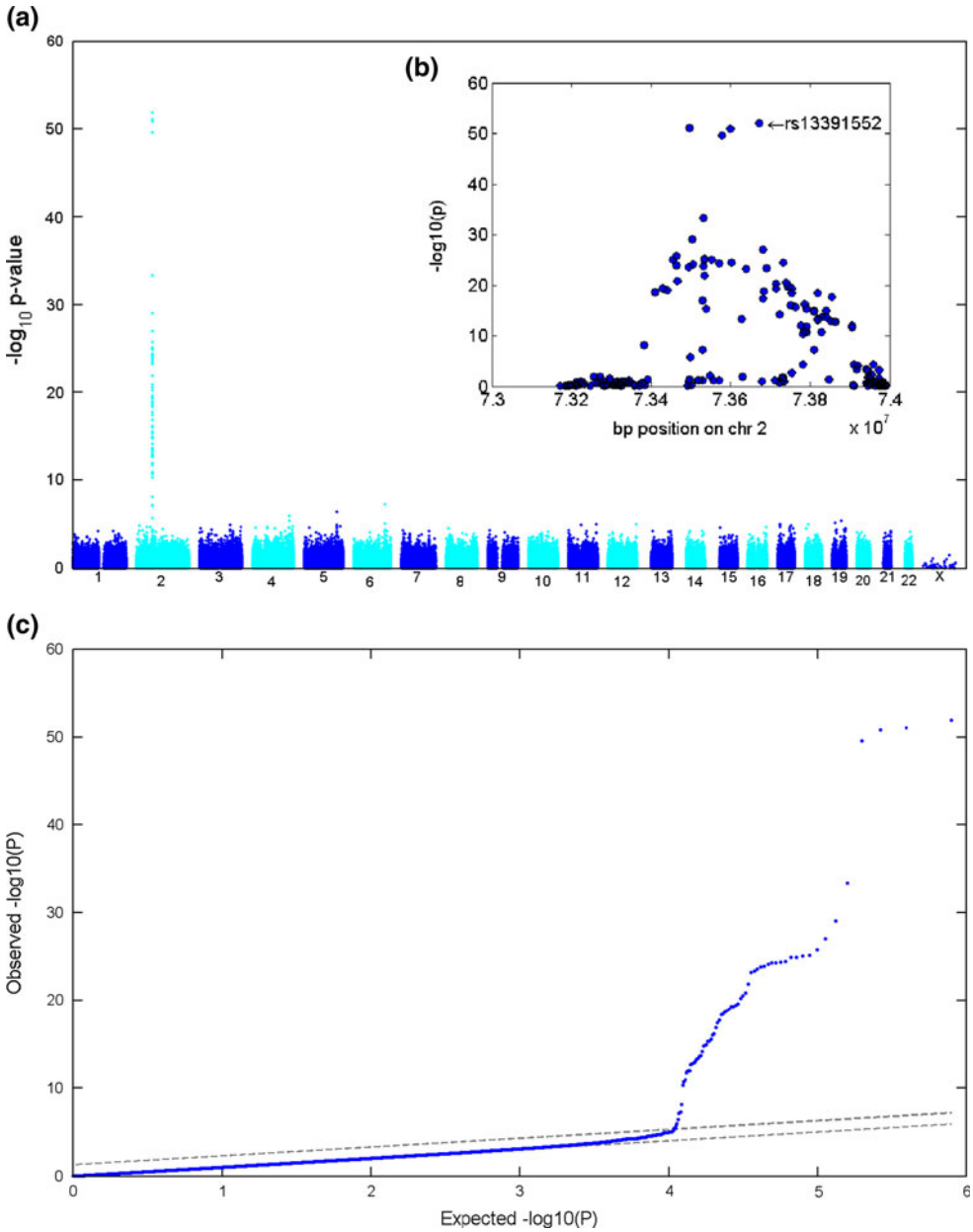

Fig. 1 **a** Manhattan plot resulting from a GWAS that uses as input phenotype the intensity of a ^{1}H-NMR chemical shift centered at 2.032 ppm. This chemical shift peak has been assigned to an *N*-acetylated compound, possibly *N*-acetylated proteins. The plot shows a strong association with a large number of SNPs on chromosome 2. A close-up (**b**) of the genomic regions shows that the center of the association peak lies clearly on the ALMS1 gene and not on the neighboring NAT8 gene. The QQ-plot (**c**) shows that a large number of SNPs are strongly associated with this compound. The fact that all the significant associations come from the same locus (*panel* **a** and **b**) underlines the existence of strong LD in this genome region

validated using the GWAS approach). The use of this genetic–metabolomic information may be useful in a preventive care context, enabling the proposal of appropriate dietary interventions to modulate these long-term metabolic imbalances and preventing the onset of disease.

Replication of gene–metabolism associations

In one of our recent studies, we carried out a GWAS with urine metabolic traits in a cohort of 265 subjects from the general population of the São Paulo metropolitan area of Brazil (Galindo-Cuspinera et al. 2009; Genick et al. 2011). This population is particularly interesting because of its high degree of ethnic and environmental diversity. Subjects were aged 18–47 (mean, 32.8), and the panel was balanced according to gender (49 % of subjects were male and 51 % female) (Supplement I). We deployed an analytical strategy similar to the one previously selected by Nicholson et al. (2011). The NMR data complexity was reduced by binning the spectra into evenly spaced chemical shift regions of 0.004 ppm width. Parallel GWAS on all chemical shift bins resulted in a very pronounced association between the NMR signal at 2.032 ppm (indicative of an *N*-acetyl group) and SNPs located in the ALMS1 and NAT8 genes on chromosome 2 (Fig. 1). This observation

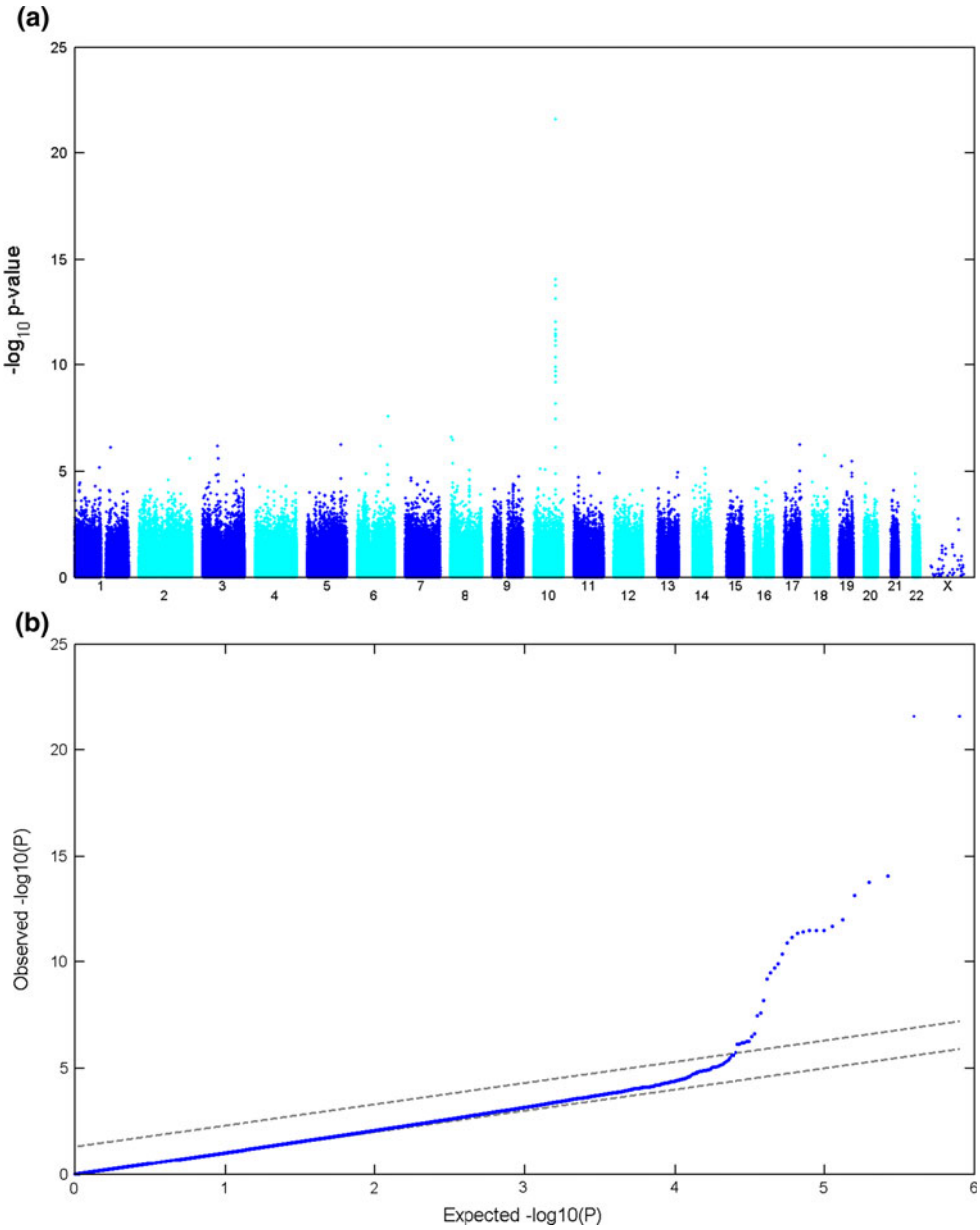

Fig. 2 **a** Manhattan plot resulting from a GWAS that uses as input phenotype the intensity in ^{1}H-NMR chemical shift centered at 2.856 ppm. This chemical shift peak has been assigned to Trimethylamine. The plot shows a strong association with a series of SNPs on chromosome 10. These SNPs are located in or around the gene PYROXD2, which encodes a probable pyridine nucleotide-disulfide oxidoreductase. **b** QQ-plot corresponding to the Manhattan plot shown in subfigure (**a**). The plot underlines the significance of the association and indicates that the strength of the association is not driven by lambda inflation

replicates the association observed by Nicholson et al. (2011). Moreover, our results show that the associations with SNPs located within the ALMS1 gene are much stronger than with those within the neighboring NAT8 gene. We suspect that the greater genetic diversity of our study population has resulted in a breakdown of the linkage disequilibrium in the ALMS1/NAT8 region of the genome, which resulted in a sharper association peak. While the molecular function of ALMS1 is not known, genetic variations within ALMS1 have been implicated in a number of kidney health disorder phenotypes (Chambers et al. 2010) including a rare genetic disease called Alström syndrome (Li et al. 2007). We further confirmed the association (Nicholson et al. 2011) observed for trimethylamine (TMA), thus confirming the strong impact of genetics on the metabolism of a precursor to the major cardiovascular risk factor TMAO (Wang et al. 2011).

Our results support that TMA concentration in urine is strongly associated with natural genetic variations in and around the gene for the putative pyridine nucleotide-disulfide oxidoreductase gene PYROXD2 (Fig. 2). Yet, the question about how PYROXD2 affects the amount of TMA in urine remains unsolved. Both ALMS1/NAT8 and PYROXD2 regions of the human genome show signs of strong and recent evolutionary pressure (Tang et al. 2007). The extent of this pressure in combination with the geographic distribution of the genetic variations indicates that carrying a particular variant of these genomic regions must have provided significant advantages in adapting to the different climatic, dietary and infectious environments. These novel findings demonstrate that genetic variations in both the ALMS1/NAT8 and PYROXD2 regions lead to marked differences in the concentration of relatively abundant human metabolites. We can also hypothesize that changes in the gut microflora composition may play an important role along this process of adaptation. Ultimately, these differences in metabolic/genetic heritage may translate into different health status and disease predisposition for different groups of individuals. GWAS has been able to highlight associations between the genetic polymorphism of the host and a metabolic pathway modulated by gut microbiota. Moreover, this pathway has been demonstrated to have important health implications, in particular for CVD.

Perspectives for genome-wide association studies to nutritional research

In summary, GWAS applications with metabolic traits offer potential for nutritional research. In typical case–control studies, frequent in Nutrimetabonomics (Rezzi et al. 2007; Martin et al. 2009a; Heinzmann et al. 2011; van Velzen et al. 2008; 2009), the genetic effects on homeostasis regulation may introduce undesired variance in metabolic profiles. These effects may obscure diet-induced metabolic changes, which may be of relatively low magnitude.

The incorporation of metabolite ratios has proven to be effective to strengthen genome–metabolome associations, which are of particular interest in short cohort studies. As it has been described, genome–metabolome associations, in particular those based in urine metabolic profiles, are characterized by very large effect sizes (Nicholson et al. 2011). This gives an opportunity to develop genome-wide/metabolome-wide studies with sample sizes typical of standard metabolomic studies. In this context, the use of metabolite ratios appears as an interesting approach to understand the metabolic information at the pathway level. The analysis of these ratios, showing higher association with a particular gene, may provide information about which compounds belong to the same pathway or, eventually, are being regulated by the same genetic mechanism. This can provide more insight into the functionalities and roles of different metabolites. Moreover, this understanding is not limited to the host. Cases such as the relationship between TMA synthesis and PYROXD2 clearly show the possible interaction between two genomes (host/gut microbiome) in more detail, and its expression in the host metabolism.

In the future, GWAS with metabolic traits might offer good potential for personalized nutrition. Through this methodology it is now possible to link metabolic phenotype with genetic background while providing new metabolic pathway target for tailor-made nutritional solutions. With this knowledge at hand, the time for proposing customized nutritional interventions to recover unbalanced pathways to a healthy status is coming closer.

References

Altmaier E, Ramsay S, Graber A, Mewes H-W, Weinberger K, Suhre K (2008) Bioinformatics analysis of targeted metabolomics—Uncovering old and new tales of diabetic mice under medication. Endocrinology 149(7):3478–3489

Beutler E (2008) Glucose-6-phosphate dehydrogenase deficiency: a historical perspective. Blood 111(1):16–24

Chambers JC, Zhang W, Lord GM, van der Harst P, Lawlor DA, Sehmi JS et al (2010) Genetic loci influencing kidney function and chronic kidney disease. Nat Genet 42:373–375

Claus SP, Ellero SL, Berger B, Krause L, Bruttin A, Molina J et al (2011) Colonization-induced host-gut microbial metabolic interaction. mBio 2(2):e00271-10

Cornelis MC, El-Sohemy A, Campos H (2007) Genetic polymorphism of the adenosine A2A receptor is associated with habitual caffeine consumption1,2,3. Am J Clin Nutr 86(1):240–244

Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, Järvelä I (2002) Identification of a variant associated with adult-type hypolactasia. Nat Genet 30:233–237

Galindo-Cuspinera V, Waeber T, Antille N, Hartmann C, Stead N, Martin N (2009) Reliability of threshold and suprathreshold methods for taste phenotyping: characterization with PROP and sodium chloride. Chemosens Percept 2(4):214–228

Genick UK, Kutalik Z, Ledda M, Souza Destito MC, Souza MM, Cirillo CA, et al (2011) Sensitivity of genome-wide-association signals to phenotyping strategy: the PROP-TAS2R38 taste association as a benchmark. PLoS ONE e27745

Gieger C, Geistlinger L, Altmaier E, Hrabé de Angelis M, Kronenberg F, Meitinger T et al (2008) Genetics meet metabolomics: a genome-wide association study of metabolite profiles in human serum. PLoS Genet 4:e1000282

Heinzmann SE, Merrifield CA, Rezzi S, Kochhar S, Lindon JC, Holmes E et al (2011) Stability and robustness of human metabolic phenotypes in response to sequential food challenges. J Proteome Res 11(2):643–655

Hoheisel JD (2006) Microarray technology: beyond transcript profiling and genotype analysis. Nat Rev Genet 7:200–210

Holmes E, Wilson ID, Nicholson JK (2008) Metabolic phenotyping in health and disease. Cell 134(5):714–717

Illig T, Gieger C, Zhai G, Römisch-Margl W, Wang-Sattler R, Prehn C et al (2010) A genome-wide perspective of genetic variation in human metabolism. Nat Genet 42(2):137–141

Kathiresan S, Srivastava D (2012) Genetics of human cardiovascular disease. Cell 148(6):1242–1257

Kathiresan S, Melander O, Giuducci C, Surti A, Burtt NP, Rieder MJ et al (2008a) Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. Nat Genet 40:189–197

Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE et al (2008b) Common variants at 30 loci contribute to polygenic dyslipidemia. Nat Genet 41:56–65

Li G, Vega R, Nelms K, Gekakis N, Goodnow C, McNamara P et al (2007) A role for Alström syndrome protein, Alms1, in kidney ciliogenesis and cellular quiescence. PLoS Genet 3(1):e8

Li M, Li C, Guan W (2008) Evaluation of coverage variation of SNP chips for genome-wide association studies. Eur J Hum Genet 16:635–643

Malerba G, Schaeffer L, Xumerle L, Klopp N, Trabetti E, Biscuola M et al (2008) SNPs of the FADS gene cluster are associated with polyunsaturated fatty acids in a cohort of patients with cardiovascular disease. Lipids 43(4):289–299

Martin F-PJ, Rezzi S, Peré-Trepat E, Kamlage B, Collino S, Leibold E et al (2009a) Metabolic effects of dark chocolate consumption on energy, gut microbiota, and stress-related metabolism in free-living subjects. J Proteome Res 8(12):5568–5579

Martin F-P, Sprenger N, Yap IK, Wang Y, Bibiloni R, Rochat F et al (2009b) Panorganismal gut microbiome–host metabolic crosstalk. J Proteome Res 8(4):2090–2105

McCarthy M, Hirschhorn J (2008) Genome-wide association studies: past, present and future. Hum Mol Genet 17(2):R100–R101

McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP et al (2008) Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 9:356–369

Merrifield CA, Lewis M, Claus SP, Beckonert OP, Dumas M-E, Duncker S et al (2011) A metabolic system-wide characterisation of the pig: a model for human physiology. Mol BioSyst 7:2577–2588

Mestdagh R, Dumas M-E, Rezzi S, Kochhar S, Holmes E, Claus SP et al (2011) Gut microbiota modulate the metabolism of brown adipose tissue in mice. J Proteome Res 11(2):620–630

Nicholson JK, Holmes E, Wilson ID (2005) Gut microorganisms, mammalian metabolism and personalized health care. Nat Rev Microbiol 3:431–438

Nicholson G, Rantalainen M, Li JV, Maher AD, Malmodin D, Ahmadi KR et al (2011) A genome-wide metabolic QTL analysis in Europeans implicates two loci shaped by recent positive selection. PLoS Genet 7(9):e1002270

Rasmussen BB, Brix TH, Kirsten KO, Brøsen K (2002) The interindividual differences in the 3-demthylation of caffeine alias CYP1A2 is determined by both genetic and environmental factors. Pharmacogenetics 12(6):473–478

Rezzi S, Ramadan Z, Fay LB, Kochhar S (2007) Nutritional metabonomics: applications and perspectives. J Proteome Res 6(2):513–525

Sachse C, Brockmöller J, Bauer S, Roots I (1999) Functional significance of a C→A polymorphism in intron 1 of the cytochrome P450 CYP1A2 gene tested with caffeine. Br J Clin Pharmacol 47:445–449

Sata F, Yamada H, Suzuki K, Saijo Y, Kato EH, Morikawa M et al (2005) Caffeine intake, CYP1A2 polymorphism and the risk of recurrent pregnancy loss. Mol Hum Reprod 11(5):357–360

Schaeffer L, Gohlke H, Müller M, Heid IM, Palmer LJ, Kompauer I et al (2006) Common genetic variants of the FADS1 FADS2 gene cluster and their reconstructed haplotypes are associated with the fatty acid composition in phospholipids. Hum Mol Genet 15(11):1745–1756

Suhre K, Shin S-Y, Petersen A-K, Mohney RP, Meredith D, Wägele B et al (2011) Human metabolic individuality in biomedical and pharmaceutical research. Nature 477:54–60

Tang K, Thornton KR, Stoneking M (2007) A new approach for using genome scans to detect recent positive selection in the human genome. PLoS Biol 5(7):e171

Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M et al (2010) Biological, clinical and population relevance of 95 loci for blood lipids. Nature 466:707–713

van Velzen EJ, Westerhuis JA, van Duynhoven P, van Dorsten FA, Hoefsloot HC, Jacobs DM et al (2008) Multilevel data analysis of a crossover designed human nutritional intervention study. J Proteome Res 7(10):4483–4491

van Velzen EJ, Westerhuis JA, Duynhoven JP, van Dorsten FA, Grun CH, Jacobs DM et al (2009) Phenotyping tea consumers by nutrikinetic analysis of polyphenolic end-metabolites. J Proteome Res 8(7):3317–3330

Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, DuGar B et al (2011) Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature 472:57–63

Wikoff WR, Anfora AT, Liu J, Schultz PG, Lesley SA, Peters EC et al (2009) Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. Proc. Natl. Acad. Sci. USA 106(10):3698–3703