

ORIGINAL ARTICLE

The heritability of metabolic profiles in newborn twins

FY Alul¹, DE Cook¹, OA Shchelochkov¹, LG Fleener¹, SL Berberich², JC Murray¹ and KK Ryckman¹

Identifying genetic and metabolic biomarkers in neonates has the potential to improve diagnosis and treatment of common complex neonatal diseases, and potentially lead to risk assessment and preventative measures for common adulthood illnesses such as diabetes and cardiovascular disease. There is a wealth of information on using fatty acid, amino acid and organic acid metabolite profiles to identify rare inherited congenital diseases through newborn screening, but little is known about these metabolic profiles in the context of the 'healthy' newborn. Recent studies have implicated many of the amino acid and fatty acid metabolites utilized in newborn screening in common complex adult diseases such as cardiovascular disease, insulin resistance and obesity. To determine the heritability of metabolic profiles in newborns, we examined 381 twin pairs obtained from the Iowa Neonatal Metabolic Screening Program. Heritability was estimated using multilevel mixed-effects linear regression adjusting for gestational age, gender, weight and age at time of sample collection. The highest heritability was for short-chain acylcarnitines, specifically C4 ($h^2 = 0.66$, $P = 2 \times 10^{-16}$), C4-DC ($h^2 = 0.83$, $P < 10^{-16}$) and C5 ($h^2 = 0.61$, $P = 1 \times 10^{-9}$). Thyroid stimulating hormone ($h^2 = 0.58$, $P = 2 \times 10^{-5}$) and immunoreactive trypsinogen ($h^2 = 0.52$, $P = 3 \times 10^{-9}$) also have a strong genetic component. This is direct evidence for a strong genetic contribution to the metabolic profile at birth and that newborn screening data can be utilized for studying the genetic regulation of many clinically relevant metabolites.

Heredity (2013) **110**, 253–258; doi:10.1038/hdy.2012.75; published online 14 November 2012

Keywords: heritability; newborn screen; metabolic profile; thyroid stimulating hormone

INTRODUCTION

Metabolomics is the study of metabolites, small molecules that make up the building blocks of cellular processes (Goonewardena *et al.*, 2010). Metabolomics is a rapidly developing field of great importance in understanding the physiology of dynamic cellular processes and for diagnosis of human diseases. Better understanding of the metabolic profiles present in the immediate postnatal period is important for their short-term contribution to diagnosis, as well as for the long-term prediction of outcomes. Metabolic profiles of the neonate, defined by levels of amino acid, organic acid and fatty acid oxidation metabolites may be correlated with metabolic conditions in adulthood, such as obesity, diabetes and cardiovascular disease through the phenomenon of metabolic programming (Srinivasan and Patel, 2008).

Understanding the genetic contribution to metabolic profiles, particularly in the neonate, may provide insight into complex diseases and their risks. The underlying genetic contribution to metabolic traits has been studied in plants (Keurentjes *et al.*, 2006), mice (Ferrara *et al.*, 2008) and more recently, in adult humans (Shah *et al.*, 2009). However, to-date, we are aware of no studies examining the heritability of metabolic traits related to amino acid and fatty acid oxidation in the neonatal period.

Population-based metabolic profiling is most commonly implemented in State-mandated neonatal screening programs used to detect numerous endocrine and metabolic disorders at birth. Levels of amino acids, including branched chain amino acids and acylcarnitines, are measured to detect newborns with rare disorders that are treatable, if detected early (Wilcken and Wiley, 2008). Some

hormones and enzyme activities are also measured including thyroid stimulating hormone (TSH), immunoreactive trypsinogen (IRT) and 17-hydroxyprogesterone (17-OHP) to detect disorders such as congenital hypothyroidism, cystic fibrosis and congenital adrenal hyperplasia, respectively (Centers for Disease Control and Prevention, 2001; Khoury *et al.*, 2003; Votava *et al.*, 2005). In addition to detecting disorders at birth, branched chain amino acids have been associated with adult diseases including type 2 diabetes and cardiovascular disease (Shah *et al.*, 2010; Wang *et al.*, 2011).

There is several genome-wide association studies that have identified genetic loci associated with metabolite measurements from adult human blood and/or urine samples (Gieger *et al.*, 2008; Illig *et al.*, 2010; Suhre *et al.*, 2011a,b; Kettunen *et al.*, 2012). While initial results are promising, less effort has been devoted to identifying the underlying heritability of metabolic traits in humans at different times of development, particularly at birth. The heritability of biomarkers for amino acid and fatty acid oxidation is virtually unexplored in adults or neonates with the exception of a single study examining an extensive panel of metabolic markers from adult plasma samples (Shah *et al.*, 2009). The authors demonstrated high heritability among several amino acids and acylcarnitines in eight multiplex families with a family history of premature cardiovascular disease (Shah *et al.*, 2009). A recent study by Kettunen *et al.* (2012) examined the heritability for amino acids, lipoproteins, lipids and some other metabolites such as glucose and urea in adult serum samples; however, the paper does not investigate acylcarnitines, which are important for fatty acid oxidation. An additional study by Nicholson *et al.* (2011) investigates the genetic and environmental contribution

¹Department of Pediatrics, University of Iowa Hospitals and Clinics, Iowa City, IA, USA and ²State Hygienic Laboratory, University of Iowa, Iowa City, IA, USA
Correspondence: Dr KK Ryckman, Department of Pediatrics, University of Iowa Hospitals and Clinics, 500 Newton Road, 2167ML, Iowa City, IA 52242, USA.
E-mail: kelli-ryckman@uiowa.edu.

Received 28 March 2012; revised 9 August 2012; accepted 21 August 2012; published online 14 November 2012

to human metabolic profiles, and the authors report estimations of familial variation for different metabolite measurements; however, they state that their sample size is insufficient for heritability estimations. A few other studies have shown a high degree of heritability in adult hormone measurements, particularly, adult serum levels of TSH where heritability estimates ranged from 32% (Samollow *et al.*, 2004) to 64% (Hansen *et al.*, 2004) to 65% (Panicker *et al.*, 2008).

In this study, we obtained twin samples from the Iowa Neonatal Metabolic Screening Program to determine the heritability of hormone, enzyme, acylcarnitine and amino acid measurements measured during routine newborn screening. Understanding the heritability of these analytes at birth may increase our understanding of physiologic processes and metabolic programming in various adulthood diseases. Furthermore, identifying the heritable components of the metabolome, particularly in the neonatal period, may pave the way for future genetic association studies that may provide insight into the physiology of different cellular processes, as well as the biology of complex neonatal and adulthood diseases.

MATERIALS AND METHODS

Study population

Metabolic data for 47 analytes from routine newborn screening were obtained from the University of Iowa State Hygienic Laboratory for 243 same sex twin pairs (130 male–male twin pairs and 113 female–female twin pairs) and 165 male–female twin pairs born in 2009 (Supplementary Table 1). Twins were identified by the State Hygienic Laboratory based on both infants having the same birth date, gestational age, mother's first name and facility identification number. Newborn dried blood spots (DBS) were obtained from a heel stick 24–72 h after birth. DBS specimens were collected, dried and handled according to the Clinical Laboratory Standards Institute guideline (Hannon *et al.*, 2007). DBS specimens were transported to the State Hygienic Laboratory within 5 days of collection, and evaluated for quality at the time of arrival by trained technical staff. Blood spot cards were obtained for DNA extraction and zygosity testing for all same sex twin pairs. Approval for use of the de-identified data and blood spot cards was granted by the Iowa Department of Public Health and a waiver of consent was obtained from the Institutional Review Board of the University of Iowa (IRB no.200908793). DNA was extracted from one dried whole blood spot using the AutoGen (Holliston, MA, USA) QuickGene-810 nucleic acid extraction machine with the DNA Tissue Kit (catalog number DT-S) and following the manufacturer's recommendations.

Enzymatic assays

Quantification of 17-OHP, TSH and IRT were determined by solid phase, time-resolved fluoroimmunoassay from dried newborn blood spots using Perkin Elmer's AutoDELFIA platform (Waltham, MA, USA). Galactose-1-phosphate uridylyl transferase (GALT) was determined by a semi-quantitative enzymatic assay by Perkin Elmer (Waltham, MA, USA) based on the Beutler method.

Tandem mass spectrometry

Tandem mass spectrometry is performed in neonatal screening to detect levels of amino acids and acylcarnitines. Screening procedures in Iowa are based on previously established methodology (Chace *et al.*, 2001; Turgeon *et al.*, 2008; Chace *et al.*, 2009). Briefly, a derivatization method is used in which butyl esters of acylcarnitines and amino acids are prepared from the extracts. For succinylacetone, hydrazine derivatives are prepared. Tandem mass spectrometry is performed with Waters Quattro Micro triple quadrupole tandem mass spectrometers, equipped with an electrospray ionization source operated in the positive ion mode. Sensitivity and resolution checks on the instruments are performed daily, and calibration and mass accuracy adjustments are done 2–3 times per year. Multiple reaction monitoring mode is used to scan for specific mass ion intensities. Concentrations are obtained from the ratio of ion

intensity (cps) at the mass that represents a specific analyte compared with its isotopically labeled internal standard and correcting for blood volume in a 1/8 inch DBS punch. Both internal and external spiked control specimens, a normal control specimen and a blank are analyzed with each batch of specimens. The external spiked control specimens are obtained from Newborn Screening Quality Assurance Program at the Centers for Disease Control.

Determination of zygosity

Zygosity was determined based on concordance of genetic markers by genotyping 20 markers with high heterozygosity (~ 0.5) (Supplementary Table 2) that were not in linkage disequilibrium with one another. Genotyping was performed using the TaqMan chemistry genotyping system (Applied Biosystems, Foster City, CA, USA). All SNP genotyping assays were available and ordered using the Assay-on-Demand service from Applied Biosystems. These genotyping assays included primers to amplify the region containing the SNP of interest and two TaqMan Minor Groove Binder probes that are specific to the polymorphic variant alleles at the site labeled with different fluorescent reporter dyes, FAM and VIC. All reactions were performed using standard conditions supplied by Applied Biosystems. Following thermocycling, fluorescence levels of the FAM and VIC dyes were measured and genotypes were scored using the Sequence Detection Systems 2.2 software (Applied Biosystems). Genotypes were uploaded into a Progeny database (Progeny Software, LLC, South Bend, IN, USA) containing the phenotypic data for subsequent statistical analysis.

Using the following equation (Nyholt, 2006), we determined that there was a $>99\%$ power to accurately differentiate between monozygotic (MZ) and dizygotic (DZ) twin pairs using >10 markers with minor allele frequencies between 0.2 and 0.5.

$$\begin{aligned} M(DZ) &= \frac{1}{4} \text{Prob}(z_0) + \frac{1}{2} \text{Prob}(z_1) + \frac{1}{4} \text{Prob}(z_2) \\ &= \frac{1}{4} M_0 + \frac{1}{2} a_2 + \frac{1}{4} \\ &= \frac{1}{4} \sum_{i=1}^n (p_i)^4 + \sum_{i=1}^n \sum_{j=i+1}^n (2p_i p_j)^2 + \frac{1}{2} \sum_{i=1}^n (p_i)^2 + \frac{1}{4} \end{aligned}$$

Of the 243 pairs, we identified 109 as DZ ($r^2 = 0.4–0.8$) and 107 as MZ ($r^2 > 90\%$). All discordant sex twin pairs ($N = 165$) were considered DZ. We excluded 27 pairs because of one or both twins having low genotyping efficiency that is <10 markers ($N = 18$) or questionable zygosity that is $r^2 < 0.4$ ($N = 5$) or $r^2 = 0.81–0.9$ ($N = 4$). The utility of DNA extracted from DBS cards is challenging because of the very-low DNA yield resulting from these small punches as starting material. This lower quality and quantity of DNA is most probably contributing to the low genotyping efficiency that we observed for some of the samples.

Statistical analysis

Demographic characteristics were compared between DZ and MZ twin pairs. Chi-square tests were used to compare categorical variables (gender, total parenteral nutrition and abnormal screen) and Wilcoxon Rank sum tests were used to compare continuous traits (gestational age, weight and age at screening). Heritability was estimated with multilevel mixed-effects linear regression adjusting for confounders known to influence analyte measurements including gender, age at time of sample collection, gestational age and birth weight. P -values from the additive genetic component are presented. Heritability was estimated from the division of the beta coefficient of the additive component over the sum of the additive, shared environment and residual coefficients. Measurements outside the lower limits of detection (LOD) were given a value of LOD. 17-OHP, TSH and GALT had 10.9%, 5.6% and 0.1% of values, respectively, that were outside the LOD. GALT also had 59 (7.7%) measurements that were outside the upper LOD; these values were assigned as the upper limits. One amino acid (ASA) and eight acylcarnitines (C5:1, C6-DC, C14:2, C14-OH, C16:1-OH, C16-OH, C18-OH and C18:1-OH) had little to no variability (s.d. ≤ 0.01) and, therefore, were excluded from heritability analysis. Succinylacetone was also excluded as $>20\%$ of the measurements were missing. For each analyte, outliers were excluded if values were outside of the mean ± 4 s.d.'s. Equal numbers of DZ and MZ twin pairs

were excluded for deviations ± 4 s.d.'s from the mean ($P=0.58$). The outliers may be due to neonatal illness that we could not account for as we did not have access to medical records. With the exception of C3, which had no outliers, 1–7 twin pairs (0.3–1.8%) were excluded for all other analytes. To normalize measurement distributions each analyte was transformed with the natural logarithm. Of the 37 analytes examined, 21 were normally distributed ($P>0.01$) after transformation; however, MET and C18:1 required additional removal of 15 and 5 twin pair outliers, respectively, and results from these analytes should be interpreted accordingly. For ARG, LEU, PHE, VAL, C5, C10:1, C12:1, C14:1, C16:1, TSH and 17-OHP normal distribution was attained for the residuals from linear regression adjusting for relevant covariates including age at the time of collection, birth weight, gestational age and gender. The residuals were used in heritability analyses for these analytes. TSH, PHE and C10:1 required additional removal of 36, 13 and 1 twin pair outliers, respectively, and results from these analytes should be interpreted accordingly. GALT, C12 and C5-DC were analyzed with the natural log transformed measurement and IRT and C10 were analyzed with the residuals; however, all deviated from normality ($P<0.01$). Additionally, we ran heritability measurements on natural log transformed analyte measurements excluding twin pairs <34 weeks gestation, with one or both twins having received an abnormal screen, with one or both twins on total parenteral nutrition and MZ twin pairs where the birth weight was >20% discordant. This was an attempt to control for infants with medical complications that could interfere with analyte measurements. As this was a population-based de-identified retrospective examination of data we were not able to connect our measurements to medical record information to obtain more detailed information on the severity or type of illness. Results were similar (Supplementary Tables 3 and 4); therefore, the full model including all available twin pairs are presented in the results and discussion below.

Analysis of ratios

There were eight amino acid ratios and 13 acylcarnitine ratios included in the measurements reported with the newborn screen (Supplementary Table 1). We estimated heritability for these ratios as described above for the single analytes. C4/C2, C5/C2, C5-DC/C16 and C16-OH/C16 were excluded because of low variability (s.d. ≤ 0.01). Outliers were excluded if values were outside of the mean ± 4 s.d.'s; 1–7 twin pairs (0.3–1.8%) were excluded for all ratios. Equal numbers of DZ and MZ twin pairs were excluded for deviations ± 4 s.d.'s from the mean ($P=0.54$). To normalize measurement distributions each ratio was transformed with the natural logarithm. Of the 17 analytes examined, 8 were normally distributed ($P>0.01$) after transformation; however ASP/HOMOCIT and LEU/PHE required additional removal of 5 and 2 twin pair outliers and results from these analytes should be interpreted accordingly. For C14:1/C12:1, CIT/ARG, TYR/PHE normal distribution was attained for the residuals from linear regression adjusting for relevant covariates including age at the time of collection, birth weight, gestational age and gender. The residuals were used in heritability analyses for these analytes. CIT/ARG required additional removal of 4 twin pair outliers, respectively, and results from these analytes should be interpreted accordingly. C4/C3 was analyzed

with the natural log transformed measurement, C8/C10 was analyzed with the residuals, C5-DC/C8 was analyzed with the natural log transformed measurement after removal of 5 twin pairs and MET/PHE was analyzed with the residuals after removal of 2 twin pairs; however, all still deviated from normality ($P<0.01$) and should be interpreted with caution.

RESULTS

Heritability analyses were performed on a total of 381 twin pairs (274 DZ and 107 MZ) (Table 1). The DZ and MZ twin pairs were not statistically different from each other when comparing their mean gestational age, total parenteral nutrition, abnormal newborn screening result, weight and age at screening (Table 1). The gender of the two groups differed marginally between DZ and MZ twin pairs ($P=0.06$), with slightly more female MZ twin pairs compared with DZ (Table 1).

Heritability measurements were significant after correction for multiple testing ($P<0.001$) for 10 analytes (Figure 1a, Supplementary Table 3). C4-DC of the short chain acylcarnitines had the highest heritability ($h^2=0.83$, $P<10^{-16}$). Other short chain acylcarnitines with high heritability included C2 ($h^2=0.50$, $P=7\times 10^{-9}$), C3 ($h^2=0.44$, $P=2\times 10^{-6}$), C4 ($h^2=0.66$, $P=2\times 10^{-16}$), C4-OH ($h^2=0.31$, $P=4\times 10^{-5}$) and C5 ($h^2=0.61$, $P=1\times 10^{-9}$). Free carnitine C0 also had a significant heritability of ($h^2=0.45$, $P=4\times 10^{-9}$). The environmental component was significant ($P<0.001$) for all analytes except C4-DC and C5. However, in most cases where shared environment was also significant, the proportion of variability explained by the genetic component exceeded the proportion explained by shared environment. TSH ($h^2=0.58$, $P=2\times 10^{-5}$) and IRT ($h^2=0.52$, $P=3\times 10^{-9}$) also had a significant heritability measurement after correction for multiple testing. The shared environment was significant for IRT ($P=5\times 10^{-4}$) but not TSH ($P=0.29$). The heritability of 17-OHP ($h^2=0.31$, $P=1\times 10^{-3}$) was marginally significant with strong shared environment ($P=7\times 10^{-7}$). The only amino acid significant after correction for multiple testing was glutamate ($h^2=0.35$, $P=5\times 10^{-4}$), the shared environment was also significant ($P=1.0\times 10^{-5}$). There was a strong genetic component for the ratios PHE/TYR ($h^2=0.51$, $P=6.4\times 10^{-6}$), TYR/PHE ($h^2=0.59$, $P=6.6\times 10^{-7}$), C3/C2 ($h^2=0.64$, $P=8.7\times 10^{-13}$), C4/C3 ($h^2=0.61$, $P=2.9\times 10^{-12}$) and C5/C3 ($h^2=0.40$, $P=4.2\times 10^{-6}$) (Figure 1b, Supplementary Table 5).

Of special interest, are the results of the environment and genetic contribution to metabolites in the β -oxidation pathway (Figure 2a). We observe a stronger environmental component for the variation in the long-chain and medium-chain acylcarnitine measurements. Whereas for short-chain acylcarnitines the environmental

Table 1 Demographic characteristics of twin pairs

	DZ (n = 274)	MZ (n = 107)	P-value
Gender of twin pair			0.06 ^a
Female–Female	45 (16.4%)	58 (54.2%)	
Male–Female	165 (60.2%)	0 (0%)	
Male–Male	64 (23.4%)	49 (45.8%)	
Gestational age of twin pair (weeks)	35.6 (2.5)	35.7 (2.3)	1.00
TPN (one or both twins)	29 (10.6%)	9 (8.4%)	0.53
Abnormal screen (one or both twins)	42 (15.3%)	19 (17.8%)	0.56
Weight (grams) (averaged for each twin pair)	2397.9 (496.9)	2364.7 (465.8)	0.19
Age at screening (hours) (averaged for each twin pair)	35.2 (11.8)	35.2 (13.0)	0.55

Abbreviations: DZ, dizygotic; MZ, monozygotic; TPN, total parenteral nutrition.
^aP-value determined comparing only Male–Male and Female–Female twin pairs.

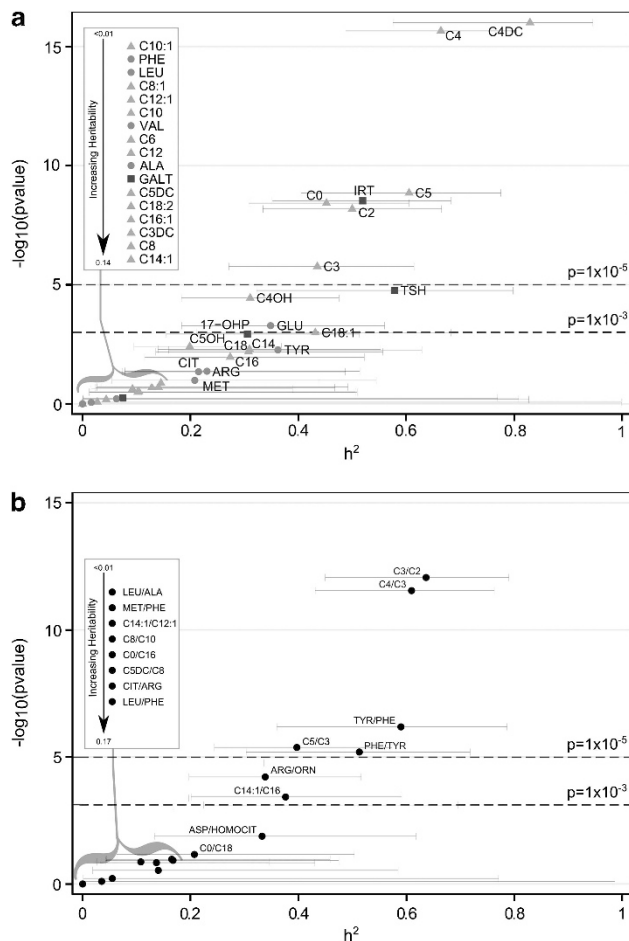


Figure 1 Heritability of analyzed metabolites and analyte ratios. (a) Heritability of analyzed metabolites. The y-axis is the negative \log_{10} of the P -value for additive genetic component and the heritability estimate is on the x-axis. Confidence intervals around the heritability point estimates are in light gray bars. IRT, TSH, 17-OHP and GALT are represented as dark gray squares, amino acids are represented as medium gray circles and acylcarnitines are represented as light gray triangles. (b) Heritability of analyte ratios. The y-axis is the negative \log_{10} of the P -value for additive genetic component and the heritability estimate is on the x-axis. Confidence intervals around the heritability point estimates are in light gray bars.

contribution decreases as the genetic contribution increases. We also observed the same trend of decreasing environmental and increasing genetic contribution for the pathway that represents the catabolism of the branched chain amino acids (valine and total leucine) into C3-DC and C4-DC (Figure 2b).

DISCUSSION

High-throughput technologies that integrate genomics and metabolomics are advantageous for providing insight into the pathophysiology and genetics of complex human diseases (Goonewardena *et al.*, 2010). Hence, it is important to recognize the heritability of a particular metabolic measurement before investing in intensive genetic characterization. To our knowledge, this is the first evaluation of the heritability of a large panel of enzyme, acylcarnitine and amino acid biomarkers in the neonatal period, in addition to being first to use newborn blood spot cards to estimate the genetic heritability. Our findings of high heritability for the acylcarnitines C2, C3, C4-OH and C5, as well as the amino acid glutamate were consistent with a

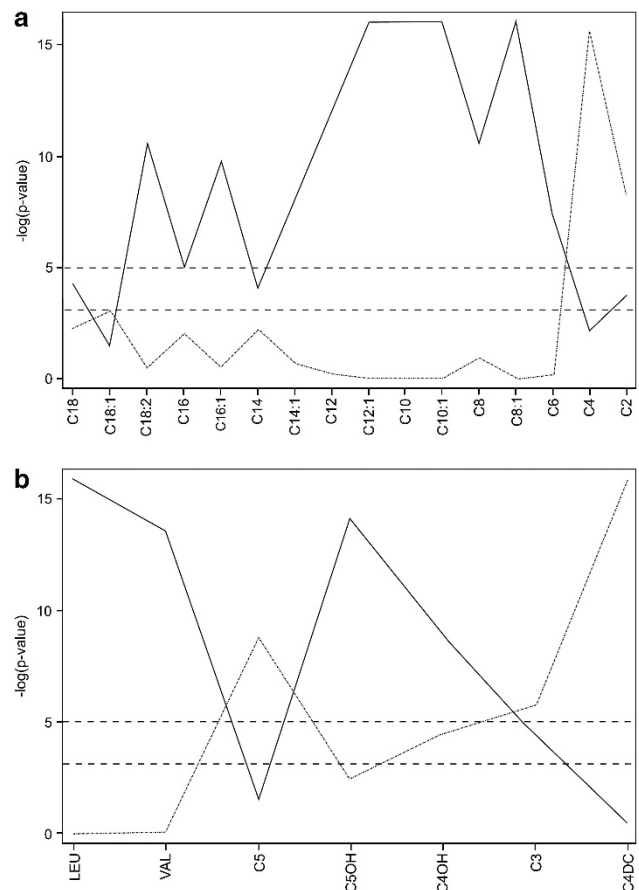


Figure 2 (a) Environment and genetic contribution to metabolites in the β -oxidation pathway. The x-axis is a list of the even-chain acylcarnitines, the y-axis is the negative \log_{10} of the P -value for the correlation coefficients. The dotted line represents the genetic component and the solid line represents the environmental component. (b) Environment and genetic contribution to metabolites along the pathway that represents the catabolism of the branched chain amino acids (valine and total leucine) into C3-DC and C4-DC. The x-axis is a list of the metabolites, the y-axis is the negative \log_{10} of the P -value for the correlation coefficients. The dotted line represents the genetic component and the solid line represents the environmental component.

previous report examining the heritability of metabolites in adults with a family history of premature cardiovascular disease (Shah *et al.*, 2009). These similarities are noteworthy not only because so few studies report heritability estimates for acylcarnitines, but also because there is recent and compelling evidence for the importance of acylcarnitines as biomarkers for cardiovascular disease, obesity and type 2 diabetes in adulthood (Mihalik *et al.*, 2010; Shah *et al.*, 2010).

Interestingly, our highest heritability was found with C4-DC, which in addition to being elevated in diabetics, was also associated with poor glycemic control, implicating this metabolite as a strong biomarker for gluco- and lipotoxicity in type 2 diabetes (Mihalik *et al.*, 2010). Furthermore, the same three metabolite classes (that is, short-chain acylcarnitines, C4-DC, C4-OH) were reported to be significantly associated with adverse outcomes after coronary artery bypass grafting (Shah *et al.*, 2012). Hence, these metabolites can be used as biomarkers of risk to predict performance after coronary artery bypass grafting (Shah *et al.*, 2012). Our study suggests that the

levels of potentially important biomarkers of adult diseases are heritable at birth. Therefore, identifying genetic variants associated with these metabolites at birth may provide important screening tools to identify diseases developed later in life.

Also noteworthy is the significant finding of TSH heritability in neonates. TSH is one of the two analytes used in newborn screening to detect primary congenital hypothyroidism (Wilcken and Wiley, 2008). As a major player in thyroid function, TSH has great influence on both neonatal, as well as adult normal physiology, affecting almost all tissues and maintaining healthy status of all human systems including cognition, cardiovascular, skeletal and metabolic function (Panicker, 2011). Our finding of significant heritability for TSH in newborn twins ($h^2 = 0.58$, $P = 2 \times 10^{-5}$) agrees with previous reports of TSH heritability estimates in adult serum measurements, ranging from 0.32 (Samollow *et al.*, 2004) to 0.64 (Hansen *et al.*, 2004) to 0.65 (Panicker *et al.*, 2008) in different studies. This finding of high TSH heritability, both in neonates and adults, suggests that thyroid function is mainly controlled by genes throughout life, and hence, neonatal screening results of TSH might help predict adult thyroid diseases later in life.

We also find relevance in these results with respect to neonatal metabolic screening programs. The State of Iowa Neonatal Metabolic Screening Program is one of the few that uses the C3/C2 ratio, for which we found strong heritability ($h^2 = 0.64$, $P = 8.7 \times 10^{-13}$), to screen for methylmalonic and propionic acidemias. We have identified patients with confirmed metabolic disorders based on elevated C3/C2 ratios in the presence of C3 below the threshold defined by the newborn screening program. This is consistent with the observation that the C3/C2 ratio has a strong genetic component with little environmental input. This suggests that the C3/C2 ratio may be a useful primary screening marker for methylmalonic and propionic acidemias as it is primarily defined by genetic influence.

We are aware of the limitations in our study. Some of the studied metabolites had large confidence intervals; that may be due to our modest sample size. Furthermore, we were not able to connect our measurements to medical record information to obtain more detailed information on neonatal illness, we attempted to control for this by estimating the heritability excluding twin pairs <34 weeks gestation and twin pairs where one or both twins were on total parenteral nutrition and observed little difference compared with the model where all twin pairs were included (Supplementary Tables 3 and 4). Another limitation to our study was that we had no information on self-identified ancestry. However, in 2009, 86.9% of the births in Iowa were Caucasian. However, it would be interesting for future studies to replicate these results in other races and age groups. We identified high heritability for TSH, and short-chain acylcarnitines all of which are important biomarkers of adult diseases. Our study illustrates the utility of having IRB reviewed access to stored, de-identified newborn dried blood spot samples as these results will guide further studies addressing the predictive ability of metabolic biomarkers in common complex adult diseases, as well as characterizing the genes underlying these observed heritabilities.

DATA ARCHIVING

Data has been deposited at the Jeff Murray Laboratory website, publication section: <http://genetics.uiowa.edu/publications.php>. Publication number 366. File password 'Twins2012'.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We express our thanks to Kim Piper at the Iowa Department of Public Health and the members of the Congenital and Inherited Disorders Advisory Committee for their enthusiastic support and management of this project. We thank Michael Ramirez, Teresa Snell, Franklin Delin and Dariush Shirazi from the State Hygienic Laboratory for their assistance in the acquisition of the newborn screening data and samples. We would like to thank Susan Berends, Tamara Busch and Osayame Ekhaguere at the University of Iowa for their technical assistance with sample management, DNA extraction and genotyping. We thank Sara Copeland at the Health Resources Services Administration for her guidance on this project. This work was supported in part by the March of Dimes (1-FY05-126 and 6-FY08-260) and NIH (R01 HD-52953, R01 HD-57192 and K99 HD-65786).

- Centers for Disease Control and Prevention (2001). Using tandem mass spectrometry for metabolic disease screening among newborns: a report of a work group. *MMWR* **50**: RR-3 1–34.
- Chace DH, DiPerna JC, Mitchell BL, Sgroi B, Hofman LF, Naylor EW (2001). Electrospray tandem mass spectrometry for analysis of acylcarnitines in dried postmortem blood specimens collected at autopsy from infants with unexplained cause of death. *Clin Chem* **47**: 1166–1182.
- Chace DH, Lim T, Hansen CR, De Jesus VR, Hannon WH (2009). Improved MS/MS analysis of succinylacetone extracted from dried blood spots when combined with amino acids and acylcarnitine butyl esters. *Clin Chim Acta* **407**: 6–9.
- Ferrara CT, Wang P, Neto EC, Stevens RD, Bain JR, Wenner BR *et al.* (2008). Genetic networks of liver metabolism revealed by integration of metabolic and transcriptional profiling. *PLoS Genet* **4**: e1000034.
- Gieger C, Geistlinger L, Altmaier E, Hrabce de Angelis M, Kronenberg F, Meitinger T *et al.* (2008). Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. *PLoS Genet* **4**: e1000282.
- Goonewardena SN, Prevette LE, Desai AA (2010). Metabolomics and atherosclerosis. *Curr Atheroscler Rep* **12**: 267–272.
- Hannon WH, Whitley RJ, Davin B, Fernhoff P, Halonen T, Lavochkin M *et al.* (2007). *Blood Collection on Filter Paper for Newborn Screening Programs; Approved Standard*, 5th edn, vol. 27. Clinical and Laboratory Standards Institute (document LA04-A5): Wayne, Pennsylvania, USA, pp 1–13.
- Hansen PS, Brix TH, Sorensen TI, Kyvik KO, Hegedus L (2004). Major genetic influence on the regulation of the pituitary-thyroid axis: a study of healthy Danish twins. *J Clin Endocrinol Metab* **89**: 1181–1187.
- Illig T, Gieger C, Zhai G, Romisch-Margl W, Wang-Sattler R, Prehn C *et al.* (2010). A genome-wide perspective of genetic variation in human metabolism. *Nat Genet* **42**: 137–141.
- Kettunen J, Tukiainen T, Sarin AP, Ortega-Alonso A, Tikkanen E, Lytikainen LP *et al.* (2012). Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat Genet* **44**: 269–276.
- Keurentjes JJ, Fu J, de Vos CH, Lommen A, Hall RD, Bino RJ *et al.* (2006). The genetics of plant metabolism. *Nat Genet* **38**: 842–849.
- Khoury MJ, McCabe LL, McCabe ER (2003). Population screening in the age of genomic medicine. *N Engl J Med* **348**: 50–58.
- Mihalik SJ, Goodpaster BH, Kelley DE, Chace DH, Vockley J, Toledo FG *et al.* (2010). Increased levels of plasma acylcarnitines in obesity and type 2 diabetes and identification of a marker of glucolipototoxicity. *Obesity (Silver Spring)* **18**: 1695–1700.
- Nicholson G, Rantalainen M, Maher AD, Li JV, Malmmodin D, Ahmadi KR *et al.* (2011). Human metabolic profiles are stably controlled by genetic and environmental variation. *Mol Syst Biol* **7**: 525.
- Nyholt DR (2006). On the probability of dizygotic twins being concordant for two alleles at multiple polymorphic loci. *Twin Res Hum Genet* **9**: 194–197.
- Panicker V (2011). Genetics of thyroid function and disease. *Clin Biochem Rev* **32**: 165–175.
- Panicker V, Wilson SG, Spector TD, Brown SJ, Falchi M, Richards JB *et al.* (2008). Heritability of serum TSH, free T4 and free T3 concentrations: a study of a large UK twin cohort. *Clin Endocrinol (Oxf)* **68**: 652–659.
- Samollow PB, Perez G, Kammerer CM, Finegold D, Zwartjes PW, Havill LM *et al.* (2004). Genetic and environmental influences on thyroid hormone variation in Mexican Americans. *J Clin Endocrinol Metab* **89**: 3276–3284.
- Shah AA, Craig DM, Sebek JK, Haynes C, Stevens RC, Muehlbauer MJ *et al.* (2012). Metabolic profiles predict adverse events after coronary artery bypass grafting. *J Thorac Cardiovasc Surg* **143**: 873–878.
- Shah SH, Bain JR, Muehlbauer MJ, Stevens RD, Crosslin DR, Haynes C *et al.* (2010). Association of a peripheral blood metabolic profile with coronary artery disease and risk of subsequent cardiovascular events. *Circ Cardiovasc Genet* **3**: 207–214.
- Shah SH, Hauser ER, Bain JR, Muehlbauer MJ, Haynes C, Stevens RD *et al.* (2009). High heritability of metabolomic profiles in families burdened with premature cardiovascular disease. *Mol Syst Biol* **5**: 258.
- Srinivasan M, Patel MS (2008). Metabolic programming in the immediate postnatal period. *Trends Endocrinol Metab* **19**: 146–152.

- Suhre K, Shin SY, Petersen AK, Mohny RP, Meredith D, Wagele B *et al.* (2011a). Human metabolic individuality in biomedical and pharmaceutical research. *Nature* **477**: 54–60.
- Suhre K, Wallaschofski H, Raffler J, Friedrich N, Haring R, Michael K *et al.* (2011b). A genome-wide association study of metabolic traits in human urine. *Nat Genet* **43**: 565–569.
- Turgeon C, Magera MJ, Allard P, Tortorelli S, Gavrilov D, Oglesbee D *et al.* (2008). Combined newborn screening for succinylacetone, amino acids, and acylcarnitines in dried blood spots. *Clin Chem* **54**: 657–664.
- Votava F, Torok D, Kovacs J, Moslinger D, Baumgartner-Parzer SM, Solyom J *et al.* (2005). Estimation of the false-negative rate in newborn screening for congenital adrenal hyperplasia. *Eur J Endocrinol* **152**: 869–874.
- Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E *et al.* (2011). Metabolite profiles and the risk of developing diabetes. *Nat Med* **17**: 448–453.
- Wilcken B, Wiley V (2008). Newborn screening. *Pathology* **40**: 104–115.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)