Heritability of Urinary Amines, Organic Acids, and Steroid Hormones in Children

Fiona A. Hagenbeek[1,2,*], Jenny van Dongen[1,2,3], René Pool[1,2], Amy C. Harms[4], Peter J. Roetman[5], Vassilios Fanos[6], Britt J. van Keulen[7], Brian R. Walker[8,9], Naama Karu[4], Hilleke E. Hulshoff Pol[10], Joost Rotteveel[7], Martijn J.J. Finken[7], Robert R. J. M. Vermeiren[5,11], Cornelis Kluft[12], Meike Bartels[1,2], Thomas Hankemeier[4], Dorret I. Boomsma[1,2,3]

[1] Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands.

[2] Amsterdam Public Health research institute, Amsterdam, the Netherlands.

[3] Amsterdam Reproduction & Development (AR&D) research institute, Amsterdam, the Netherlands.

[4] Division of Analytical Biosciences, Leiden Academic Center for Drug Research, Leiden University, Leiden, the Netherlands.

[5] LUMC-Curium, Department of Child and Adolescent Psychiatry, Leiden University Medical Center, Leiden, the Netherlands.

[6] Department of Surgical Sciences, University of Cagliari and Neonatal Intensive Care Unit, Cagliari, Italy.

[7] Emma Children's Hospital, Amsterdam UMC, Vrije Universiteit Amsterdam, Pediatrics, Amsterdam, The Netherlands.

[8] Centre for Cardiovascular Science, University of Edinburgh, Queen's Medical Research Institute, Edinburgh, United Kingdom.

[9] Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, United Kingdom.

[10] Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands.

[11] Youz, Parnassia Psychiatric Institute, the Hague, the Netherlands.

[12] Good Biomarker Sciences, Leiden, the Netherlands.

[*] Correspondence: E-mail: f.a.hagenbeek@vu.nl, Telephone: +31 (0)20 59 82974,

Abstract

Variation in metabolite levels reflects individual differences in genetic and environmental factors. Here, we investigated the role of these factors in urinary metabolomics data in children. We examined the effects of sex and age on 86 metabolites, as measured on three metabolomics platforms that target amines, organic acids, and steroid hormones. Next, we estimated their heritability in a twin cohort of 1300 twins (age range: 5.7 - 12.9 years). We observed associations between age and 50 metabolites and between sex and 21 metabolites. The mean monozygotic (MZ) and dizygotic (DZ) correlations for urinary metabolites were 0.51 (range: 0.25-0.75) and 0.16 (range: 0.01-0.46) for the amines, 0.52 (range: 0.33-0.64) and 0.23 (range: 0.07-0.35) for the organic acids, and 0.61 (range: 0.43-0.81) and 0.25 (range: 0.11-0.44) for the steroids. Broad-sense heritability was 0.49 (range: 0.25-0.64), 0.50 (range: 0.33-0.62), and 0.64 (range: 0.43-0.81) for 50 amines, 13 organic acids, and 6 steroids, and narrow-sense heritability was 0.50 (range: 0.37-0.68), 0.50 (0.23-0.61), and 0.47 (range: 0.32-0.70) for 6 amines, 7 organic acids, and 4 steroids. We conclude that urinary metabolites in children have substantial heritability, with similar estimates for amines and organic acids, and higher estimates for steroid hormones.

Keywords: heritability; classical twin design; metabolites; urine; children; amines; organic acids; steroid hormones.

1. Introduction

Metabolites are small, low molecular weight (<1.5 kDa) molecules that are involved in cellular metabolism [1]. Metabolites have many functions in the human body, including signaling, energy storage, and structure formation [2]. The extensive profiling of metabolites in biofluids, cells, or organisms, i.e., metabolomics, provides the means to investigate metabolic regulation [3], and has been successfully applied to elucidate biological mechanisms and to identify drug targets or disease biomarkers [4]. Individual differences in metabolite levels reflect variation in exogenous factors, such as diet, medication use, or smoking behavior, and endogenous factors, such as sex, and age [5]. Genetics is an endogenous factor that has both direct and indirect influences through lifestyle and behavior.

Previous heritability studies of metabolomics data in various biofluids found that about 50% of the variance in metabolite levels is due to additive genetic factors [6,7]. Non-additive (i.e., dominant genetic effects) and common (shared) environmental factors (i.e., shared by family members) generally have a minor influence on metabolite levels [8–13]. Some nuance on the contribution of genetic factors is possible: the contribution to individual differences in metabolite levels differs per metabolite class and subclass, with generally somewhat higher heritability estimates observed for lipids and lipid-like molecules than for organic acids and derivatives [14–21]. The genetic architecture of many subclasses of metabolites has yet to be studied in detail. For example, steroid hormones, compared to other lipids, are generally not well represented on broad metabolomics platforms, and are less frequently investigated with targeted analytical platforms. Targeted metabolomics studies that include steroid hormones show a substantial contribution of genetic factors to cortisol and cortisol metabolism, androgens, and estrogens in various biofluids [22–25] and sometimes report sex differences in the contribution of genetic or shared environmental factors to steroid hormones [26].

With the exception of steroid hormones, the earlier studies have mainly focused on metabolite levels in serum or plasma samples in (young) adults of European ancestry. Relatively little is known about the metabolome in children, in different tissues, or in different ancestries. Here, we aimed to elucidate the genetic etiology of the urinary metabolome in children. Urine is an interesting biological matrix for metabolomics studies as it can be collected non-invasively in large quantities. Urine metabolites provide a snapshot of an individual's exposure and endogenous (water-soluble) waste clearance and reflect filtration of blood in the kidneys at approximately 120mL per minute [27]. Urine contains a wealth of information on metabolic activity, and to date more than 3000 metabolites have been identified in urine [28,29]. One investigation of the urinary metabolome in

children reported high estimates for SNP-based heritability [30], suggesting a substantial contribution of genetic factors to urinary metabolite levels in children.

In the ACTION Biomarker Study [31–33], urinary metabolomics data for 86 metabolites were generated on a liquid chromatography mass spectrometry (LC-MS) platform targeting amines, a gas chromatography (GC) MS platform targeting organic acids, and an LC-MS platform targeting steroid hormones. We first investigate the effects of sex and age on these 86 urinary metabolites in the ACTION Biomarker Study [31–33]. This study includes 1068 monozygotic (MZ) twins (531 pairs, M age = 9.6, SD age = 1.9, 47.4% females) and 232 dizygotic (DZ) twins (114 pairs, M age = 9.9, SD age = 1.6, 51.7% females, **Table 1**). We undertook to replicate the sex and ages associations in an independent clinical cohort of 179 children (M age = 10.2, SD age = 1.8, age range: 6.3-13.4, 25.1% females). Second, we aim to characterize the reliability of our measures by considering the MZ twin correlations. The MZ correlation forms an alternative for test-retest correlation, and, as, such, can serve as a reliability index. Specifically, a high MZ correlation is consistent with high reliability of the measure. However, a low MZ correlation does not necessarily mean that the reliability is low [34]. Third, we quantified the familial resemblance for urinary metabolites in children by estimating the correlations in MZ and DZ twin pairs and applied genetic structural equation modeling of MZ and DZ data to obtain maximum likelihood estimates of heritability. We found that sex and age significantly influence urinary metabolites in young twins and replicated the associations with age in the independent clinical cohort. The MZ twin correlations implied that we could measure approximately 90% of the urinary metabolites with moderate to high reliably, with more steroids rating high reliability as compared to amines and organic acids. Genetic modelling revealed that genetic factors account for nearly all familial resemblance in urinary metabolites.

Table 1. Participant characteristics by the LC-MS amine, GC-MS organic acid, and LC-MS steroid platforms for the NTR-ACTION cohort.

	N (N pairs)	Mean (SD) [range] age	N (%) females
Total	1300 (645)	9.6 (1.8) [5.7 - 12.9]	626 (48.2%)
MZ	1068 (531)	9.6 (1.9) [6.0 - 12.9]	506 (47.4%)
DZ	232 (114)	9.9 (1.6) [5.7 – 12.0]	120 (51.7%)

Notes: LC-MS, liquid chromatography mass spectrometry; GC-MS, gas chromatography mass spectrometry; NTR, Netherlands Twin Register; MZ, monozygotic twin pairs; DZ, dizygotic twin pairs.

2. Results

2.1. Urinary metabolites in children have robust associations with age not with sex

To explore the effects of sex and age on urinary metabolites, we ran mixed-effect models with sex and age as fixed effects and family as a random effect in 1300 participants (*M* age = 9.6 years, *SD* age = 1.8, age range: 5.7 - 12.9, 48.2% females) of the NTR-ACTION cohort (**Table 1**). Sex and age effects were significant for 13 of the 56 amines (23%), for 2 of the 20 organic acids (10%), and for 4 of the 10 steroids (40%) (**Figure 1**, **Table S1**). Age effects, but not sex effects, were significant for 39 (70%), 14 (70%), and 5 (50%) amines, organic acids, and steroid hormones, respectively. For one organic acid (5%) and one steroid (10%), sex effects, but no age effects, were significant. For the remaining 4 amines (7%) and 3 organic acids (15%), we observed no significant effects of sex or age. We replicated the significant associations of age with 30 metabolites in 179 children of the LUMC-Curium cohort (*M* age = 10.2 years, *SD* age = 1.8, age range: 6.3-13.4, 25.1% females, **Figure 1**, **Table S2**). The remaining associations for age, and none of the associations for sex were replicated. There were high correlations among the strength and direction of association between the discovery and replication results (**Figure S1a**).

We expect that certain metabolites, especially steroid hormones, have lower levels in younger children and thus may not show sex differences at younger ages. Therefore, we performed a series of sensitivity analyses by stratifying the sample by sex and into two age groups, consisting of children age 10 years or older and children younger than 10 years (see **Supplementary Results**). The association of sex in the older age group (age≥10) with putrescine showed higher levels in girls, but all other associations with sex or age did not replicate across the sensitivity analyses.
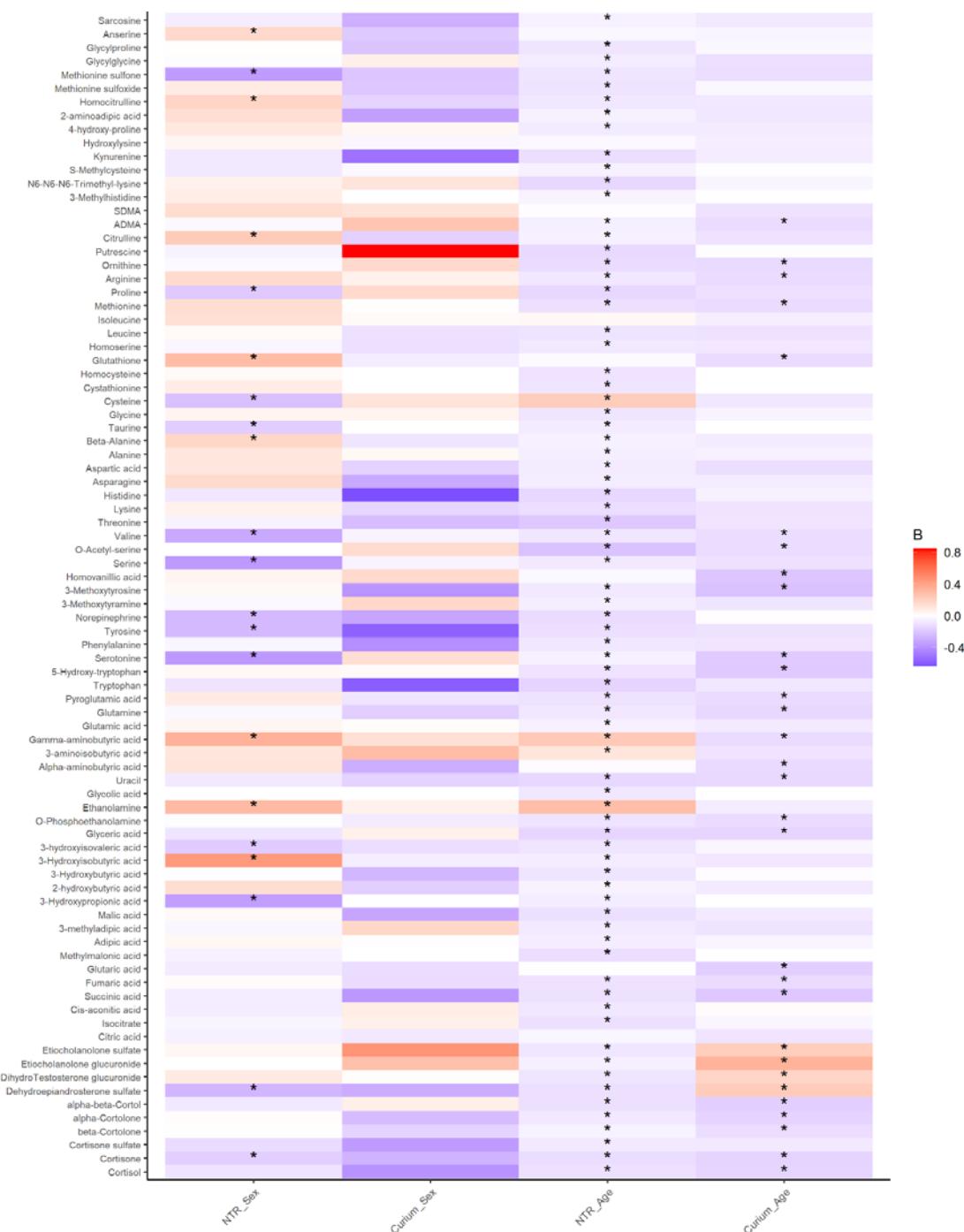
Figure 1. Comparison of the associations of sex and age with metabolites in the NTR and LUMC-Curium ACTION cohorts. The vertical axis lists the metabolites and the horizontal axis depicts the associations for sex or age in the discovery NTR-ACTION cohort (NTR) and the replication LUM-Curium ACTION cohort (Curium). All associations marked with a star (*) where significant after correcting for multiple testing (see **Methods**). The color gradient indicates the strength of the associations (beta coefficient [B]), where dark blue indicates strong negative associations and dark red indicates strong positive associations. All association results from the discovery are reported in **Table S1** and all association results from the replication are reported in **Table S2**.

2.2. Ninety percent of urinary metabolites in children have moderate to high reliability

To estimate the lower bound of reliability of the urinary metabolites in children, as obtained with the metabolomics platforms for the ACTION Biomarker Study, we obtained MZ twin correlations for 1068 MZ twins (531 pairs, M age = 9.6, SD age = 1.9, 47.4% females) (**Table 1**). The MZ correlations, corrected for sex and age effects (**Table S3**), were on average 0.51 for the amines (range: 0.25-0.75), 0.52 (range: 0.33-0.64) for the organic acids, and 0.61 (range: 0.43-0.81) for the steroids (**Figure 2**, **Table S4**). MZ correlations may be interpreted as a form of test-retest reliability [34], and these results thus suggest that approximately 90% of the measures of urinary metabolites in the ACTION Biomarker Study were measured with adequate reliability. One (1.8%) amine and 4 (40.0%) steroids were obtained with excellent reliability ($r_{MZ} \geq 0.75$), 8 (14.3%) amines and 3 (15.0%) organic acids had good reliability (r_{MZ} 0.60- 0.74), and 43 (76.8%) amines, 13 (65.0%) organic acids, and 6 (60.0%) steroids had moderate reliability (r_{MZ} 0.40-0.59) [35]. For the remaining 4 amines (7.1%) and 4 organic acids (20%, r_{MZ} <0.40) we could not infer their reliability, as low MZ correlations do not necessarily imply low reliability [34].
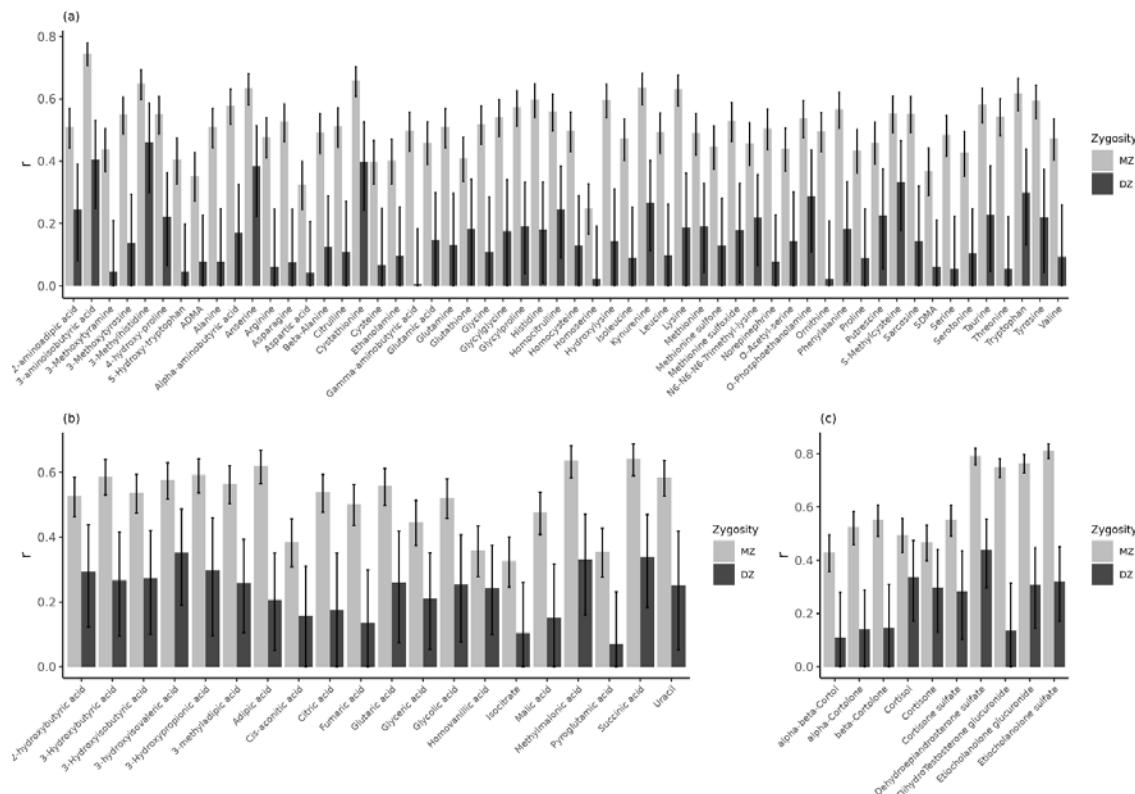

Figure 2. Monozygotic (MZ) and dizygotic (DZ) twin correlations of the (**a**) LC-MS amines, (**b**) GC-MS organic acids, and (**c**) LC-MS steroids corrected for sex and/or age in the NTR-ACTION cohort. See **Table S3** for an overview of the correction for sex and/or age and for all correlations see **Table S4**.

2.3. Genetic inluences explain familial resemblance in urinary metabolites in children

MZ twin correlations reflect both the lower bound of measurement reliability, and the upper limit of the broad-sense heritability (h^2) of these metabolites, where broad-sense heritability refers to the variance explained by all, i.e., additive and non-additive genetic, factors. When correlations in MZ twins are higher than in DZ twins, this is consistent with a genetic influence on a trait. The more detailed pattern of MZ and DZ correlations is informative to assess the contribution of additive and non-additive genetic factors and of shared environmental contributions to their resemblance.

We obtained MZ and DZ twin correlation for 1068 MZ twins (531 pairs, *M* age = 9.6, *SD* age = 1.9, 47.4% females) and 232 DZ twins (114 pairs, *M* age = 9.9, *SD* age = 1.6, 51.7% females) from the NTR-ACTION cohort (**Table 1**). **Figure 2** clearly shows that for nearly all metabolites, the MZ correlations are substantially larger than the DZ correlations, indicating that genetic factors are an important source of phenotypic individual differences. The mean MZ correlation was 0.51 for the amines (range: 0.25-0.75), 0.52 (range: 0.33-0.64) for the organic acids, and 0.61 (range: 0.43-0.81) for the steroids (**Figure 2**, **Table S4**). The mean DZ correlations were 0.16 (range: 0.01-0.46), 0.23 (range: 0.07-0.35), and 0.25 (range: 0.11-0.44), for the amines, organic acids, and steroids, respectively. Thus, the MZ correlations indicated substantial broad-sense heritability ($h^2>0.40$) for 52 (92.9%) amines, 16 (80.0%) organic acids, and 10 (100%) steroids.

Based on the twin correlations, we estimated the contribution of genetic and environmental contributions to the phenotypic variance by genetic structural equation modelling, by estimating additive genetic effects (A), unshared environmental effects (E), and dominance effects (D) or common (shared) environmental effects (C). Because the ACDE model is not identified in the classical twin design, we applied a common rule of thumb for inspecting twin correlations [36] and fitted the ADE model given $r_{MZ}>2*r_{DZ}$, or the ACE model given $r_{MZ}<2*r_{DZ}$ (**Figure 2**, **Table S4**). We estimated variance components in ADE models for 50 (89.3%) amines, 13 (65.0%) organic acids, and 6 (60.0%) steroids, and ACE models for 6 (10.7%) amines, 7 (35.0%) organic acids, and 4 (40.0%) steroids (**Table S5**).

For the 69 (80.2%) metabolites where, the ADE model was the model of choice, we reported the broad-sense h^2, not the separate estimates of the standardized variance for A and D. Given the relatively small number of complete DZ pairs (114) compared to the number of MZ pairs (645) (**Table 1**), the power to detect these variance components is low [37,38]. Here, the broad-sense h^2 was 0.49 (range: 0.25-0.64), 0.50 (range: 0.33-0.62), and 0.64 (range: 0.43-0.81) for the amines, organic acids, and steroids, respectively (**Figure 3**, **Table S6**). In the ADE models, the mean contribution of unique

environmental factors was 0.51 (range: 0.36-0.75) for the amines, 0.50 (range: 0.38-0.67) for the organic acids, and 0.36 (range: 0.19-0.57) for the steroids. For the 17 (19.8%) metabolites where an ACE model fitted best, the narrow-sense h^2 (i.e., the standardized effect of A in the ACE model) was 0.50 (range: 0.37-0.68) for the amines, 0.50 (0.23-0.61) for the organic acids, and 0.47 (range: 0.32-0.70) for the steroids (**Figure 3**, **Table S6**). Here, common environmental factors explained on average 13% (range: 0.04-0.28), 6% (range: 0-0.13), and 10% (range: 0.02-0.18) of the variance, while unique environmental factors accounted for approximately 37% (range: 0.25-0.46), 45% (range: 0.36-0.64), and 42% (range: 0.21-0.53) of the variance in amines, organic acids, and steroids, respectively.
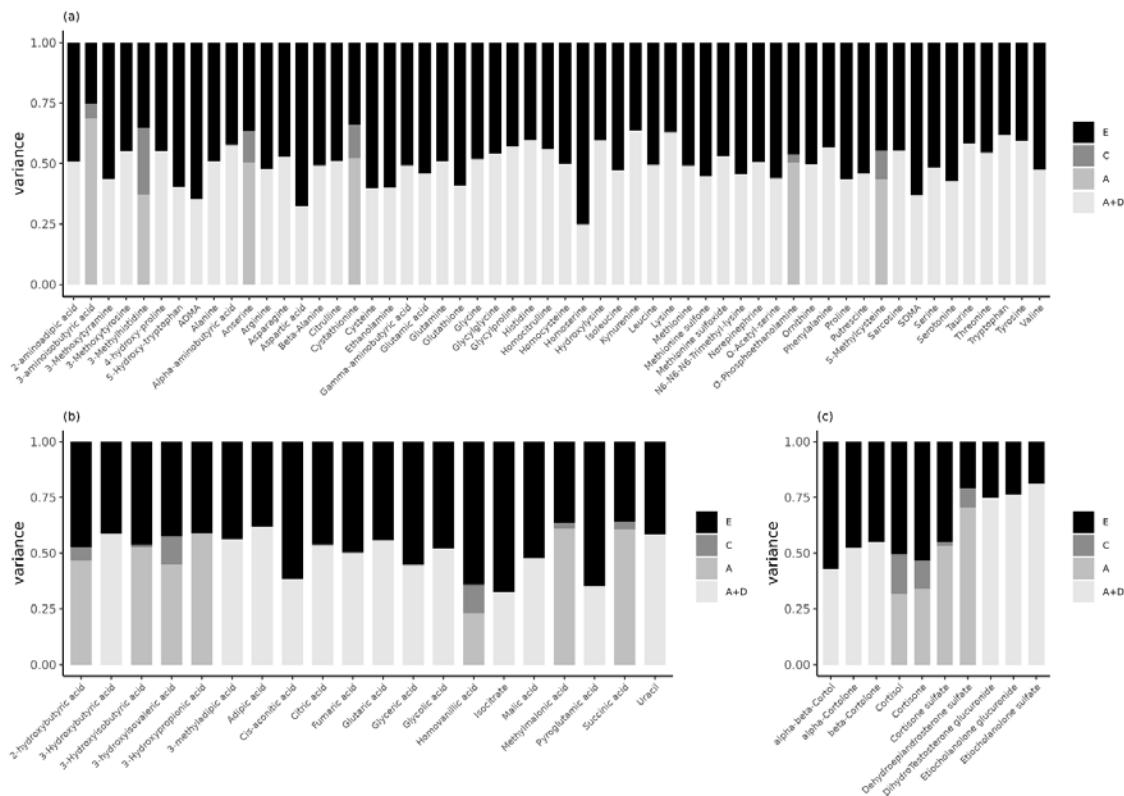
Figure 3. Estimates of the standardized variance components of the ADE or ACE models for the (**a**) LC-MS amines, (**b**) GC-MS organic acids, and (**c**) LC-MS steroids corrected for sex and/or age in the NTR-ACTION cohort. For the ADE models we provide the estimates for the broad-sense heritability, i.e., the sum of the additive (A) and dominant (D) genetic variance components. See **Table S3** for an overview of the correction for sex and/or age and for all estimates and confidence intervals see **Table S6**.

3. Discussion

We investigated the influence of sex and age on 86 urinary metabolites and estimated the heritability of these metabolites in 5 to 13-year-old twins. To assess the association of age and sex we tested their significance in a discovery cohort of twins. Even in this sample with a relatively small

age range, age was significantly associated with 50 metabolites (58%) and sex with 21 metabolites (35%). Replication in an independent cohort showed that the associations with age were the most robust, as 30 metabolites retained their significant association with age, while the association of putrescine with sex only replicated in the older age group (age≥10) after stratification. To address the question if metabolites can be reliably assessed in urine samples collected at home from young children, we determined the lower bound of the reliability, by estimating the MZ correlations and observed a range of 0.25-0.81, with 90% of the urinary metabolites reliably measured. We addressed the etiology of individual differences by the application of the classical twin design and genetic structural equation modeling of the MZ and DZ data. The main result is that individual differences in urinary metabolites are mainly explained by genetic factors. In 69 (80.2%) metabolites the broad-sense heritability was estimated at approximately 51% (range: 25-81%) and in 17 metabolites the narrow-sense heritability was on average 49% (range: 23-70%) with 9% (range: 0.3-28%) of the variance attributable to common environmental factors.

We reported robust associations of urinary metabolites with age but not with sex. We had anticipated sex differences in steroid hormones [26], but that sex differences in metabolites did not replicate in an independent cohort, may be due to a lack of power given the relatively low number of girls (N = 45, 25%) in the independent cohort compared to the discovery cohort. We observed the most robust associations for age with dehydroepiandrosterone sulfate (DHEA-S), conjugated etiocholanolone (sulfate and glucuronide), and homovanillic acid, as these associations were observed in the discovery and replication cohorts regardless of dichotomization of the age variable (**Figure S2**). Age affects the activity of endocrine systems, including the hypothalamic-pituitary-adrenal (HPA) axis and the hypothalamic-pituitary-gonadal (HPG) axis. Four conjugates of gonadal androgens were positively associated with age (**Figure 4**), which reflects expected physiological changes with age. Conversely, four adrenal-produced glucocorticoids were negatively associated with age of children in both cohorts (**Figure 4**). Higher levels of circulating cortisol are not unexpected in younger children (aged 39–106 months) [39]. High secretion of the corticosteroids is associated with delayed pubertal development [40,41], but is also affected by exposure factors, such as high-protein diet [42]. Therefore, the observed associations of steroidal hormones with age may reflect both the direct impact of age on the secretion of these metabolites, and indirectly reflect age effects on exposure-related factors that change the secretion, metabolism, and urine excretion rates. Additional associations with age were exclusively negative (**Figure 1**), and many of these metabolites are expected to change along the course of life [43–46]. They included energy production intermediates and neuroactive compounds (along with their metabolites or precursors), as well as products of lipid and protein metabolism which are routinely measured in urine [46].
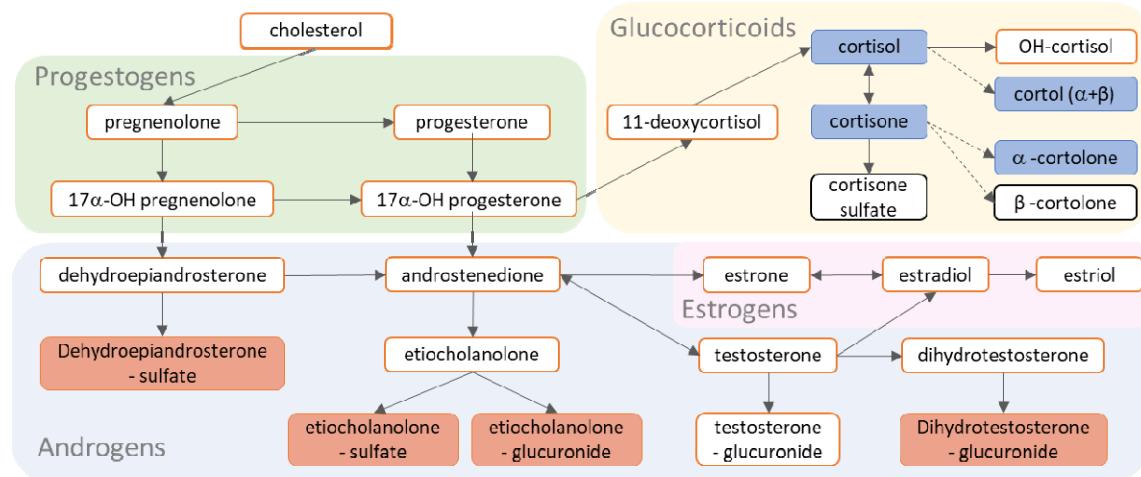
Figure 4. Metabolism of steroid hormones covered in this study. Metabolite boxes with red background indicate metabolites that were positively associated with age, a dark blue background indicates negative association with age. Metabolite boxes with white background in red frames were not covered in this study, while those in black frames were but did not show significant association with age in both cohorts. Arrows indicate direct reactions while broken arrows indicate multi-step conversion. All association results for all models in the discovery are reported in **Table S1** and **Tables S9-S12** and all association results for all models in the replication are reported in **Table S2** and **Tables S13-S16**.

In contrast to our findings, a previous study found no associations of age with 44 urinary metabolites, and only 3 significant associations with sex in a meta-analysis of 1192 children [47]. Part of the discrepancies between our study and the meta-analysis is due to the low overlap in urinary metabolites. Only 16 of the 44 urinary metabolites (36%) included in the meta-analysis were also included on the metabolomics platforms comprising the ACTION Biomarker Study. Of the 16 overlapping metabolites 14 had significant associations with age and 2 with sex, in the discovery cohort of the current study. Another likely explanation for the discrepancy is that the children in the meta-analysis were younger (mean = 7.4, range: 6.5–8.9) than those included in our discovery (mean = 9.6, range: 5.7-12.9) and replication cohorts (mean = 10.2, range: 6.3-13.4). Thus, the metabolites that are significantly associated with age in our study might reflect the onset of puberty. This would explain why the associations with the steroids, DHEA-S and conjugated etiocholanolone were among the most robust.

Genetic modelling showed that all metabolites, irrespective of platform, are subject to genetic influences. We might have expected that in this sample of young children (who live at home with their parents, and share diet, exposures to the same household, and neighborhood), there would be a detectable influence of the shared environment. However, we found that mainly genetic and unique environmental factors accounted for the individual differences in urinary metabolites in children. The substantial contribution of genetic factors to urinary metabolites in children is in line

with previously reported SNP-based heritability estimates for the urinary metabolome in children [30]. A previous small twin study (N = 128) investigating the urinary metabolome in adults showed a moderate contribution of genetic factors to 14 diet-associated urinary metabolites, and substantial contributions of shared environmental influences [48]. This study in adults only investigated the heritability of 14 diet-associated metabolites, already implying these metabolites are under environmental influence. In this same sample of twins, it was shown that 20% of the urinary NMR metabolomic profile was stable over a 2-month period and that both common genetic and shared environmental factors contribute to the short-term stability of the urinary metabolome in adults [49]. Previous studies investigating the blood metabolome in adult populations [8,10–13], also obtained very limited evidence for contributions of the common environment, while the influence of genetics is substantial. To date, genome-wide association studies for blood metabolites have identified more than 800 metabolite loci, with an average heritability of metabolite loci of 6% for lipids and lipid-like molecules and of 1% for organic acids and derivatives [18]. Here we extend these findings to show substantial broad-sense heritability for the majority of urinary amines, organic acids, and steroids in children. Together these results suggests that the majority of the urinary metabolites in children and blood metabolites in adults are determined by genetic and unique environmental influences and not by shared environmental influences.

It should be noted that the twin pairs in this study were selected for their concordance or discordance for childhood aggressive behavior, though we found no replicable associations between the urinary amines and organic acids with childhood aggression [33]. A limitation of the current study is the underrepresentation of DZ twin pairs. As a consequence, the power to detect significant dominant genetic and common environmental variance components was low [37,38]. It should also be noted that we often observed very large differences between the MZ and DZ twin correlations. This included instances where the MZ correlations were greater than four times the DZ correlations ($4*r_{DZ}<r_{MZ}$), which would result in negative values of the additive genetic variance in the ADE models [50]. Therefore, we chose to report the broad-sense heritability estimates for the metabolites with ADE models as opposed to the separate estimates for the additive and dominant genetic variance components. Based on the twin correlations, we can infer that dominant genetic effects likely contribute to the broad-sense heritability. This is in contrast to previous studies for serum metabolites in adults that suggested most genetic effects on metabolites were additive [8,9]. The pattern of very high MZ and very low DZ correlations could also indicate emergenesis (i.e., interactions between relatively infrequent alleles at different loci) [34,51], though it has been argued that duplicate gene interactions (i.e., complete dominance at both gene pairs, where the effect of

one gene is hidden if the other is dominant) between moderately frequent alleles would also account for very low correlations between DZ twins [52].

In summary, this is the largest heritability study for urinary metabolites in children to date, benefitting from a large sample size (N = 1300), and the breath of coverage both in terms of the number of analyzed metabolites (86) and coverage across the metabolome by targeting amines, organic acids, and steroid hormones. We found substantial heritability for urinary metabolites in children, and genetic factors and not common environmental factors explain most of the familial resemblance in urinary metabolites. Age, but not sex, showed robust associations with urinary metabolites. These results are promising for future genome-wide association studies on urinary metabolites, which can elucidate the genetic architecture of the urinary metabolome in children.

4. Materials and Methods

4.1. Study population and procedures

4.1.1. NTR-ACTION cohort

Twin pairs from the Netherlands Twin Register [53] were invited for participation in the NTR-ACTION Biomarker Study (Aggression in Children: Unraveling gene-environment interplay to inform Treatment and InterventiON strategies) [31,32] based on their longitudinal concordance or discordance for aggressive behavior rated by either the mother (93%) or teacher (7%) (see [33]). Parents of participating twin pairs were asked to collect first-morning urine and buccal-cell samples for their children using the standardized protocols as developed for the large-scale collection in the home situation (http://www.action-euproject.eu/content/data-protocols). Urine samples were stored in a home freezer at −18°C, transported to the lab in a mobile freezer unit at −18°C, and stored in the lab at −80°C until further processing. Zygosity assessment in the twins was based on SNP arrays as described in Odintsova et al. [54].

First-morning urine was collected for 1362 twins, for whom metabolite profiling was performed on three metabolomics platforms targeting amines, steroids, and organic acids. We excluded twins if there was an insufficient amount of urine to analyze all three platforms (N = 2), if the collected urine was not the first-morning urine (e.g., parent-reported time of urine collection was after 12:00 in the afternoon; N = 13), if the time between collection and freezing of the urine sample was more than 2 hours (N = 25), or if the twins were the second pair of multiples in the same family (N = 22). We performed the analyses in 1300 twins (**Table 1**). We give participant characteristics per aggression concordance group in **Table S7**.

4.1.2. LUMC-Curium cohort

Children referred to the LUMC-Curium academic center for child and youth psychiatry and whose parents had consented to participation in the ongoing biobank protocol of LUMC-Curium were included in the ACTION Biomarker Study. Parents of participating children collected first-morning urine and buccal-cell swaps with the protocols as developed for the NTR-ACTION cohort and all biological samples were analyzed simultaneously with the NTR samples. Information on psychiatric disorders in the LUMC-Curium sample is provided in Hagenbeek et al. (2020) [33]. In the LUMC-Curium cohort, we performed the analyses in 179 children (*M* age = 10.2, *SD* age = 1.8, age range: 6.3-13.4, 25.1% females, including 8 sibling pairs and 1 sibling trio), after exclusion of one child who also participated in NTR-ACTION, and exclusion of urine samples that were not first-morning (*N* = 3), or if the time between collection and freezing of the urine sample was more than 2 hours (*N* = 3).

4.2. Creatinine measurement

Creatinine was measured using a colorimetric assay kit according to manufacturer's instructions (Cayman, Ann Harbor, MI, USA).

4.3. Metabolite profiling

4.3.1. Measurement protocol

The Metabolomics Facility of the University of Leiden (Leiden, the Netherlands) assessed metabolites in urine on three platforms: an ultra-performance liquid chromatography mass spectrometry (UPLC-MS) platform targeting amines, a gas chromatography mass spectrometry (GC-MS) platform targeting organic acids, and a UPLC-MS platform targeting steroids. To ensure random distribution of aggression cases and controls across batches, we randomized NTR and LUMC-Curium ACTION subjects across batches. In each batch, we retained twins from the same pair on the same plate. Each batch included a calibration line, quality control (QC) samples (every 10 samples), sample replicates and blanks. The QC samples comprised pooled aliquots of all urine samples from all children. Using the pooled QC samples, the lab applied in-house developed algorithms to compensate for shifts in the sensitivity of the mass spectrometer across batches. Relative Standard Deviation (RSD) of the Quality Control samples (RSDqc) were used to evaluate the performance and reproducibility of individual metabolites. We report metabolites as 'relative response ratios' (target area/area of internal standard) after QC correction.

4.3.2. LC-MS amine platform

Prior to chromatic separation, they added methanol to 5 μL of spiked (with internal standards, see **Table S8**) urine for protein precipitation and centrifuged the supernatant. After sample

14

evaporation (speedvac), samples were reconstituted in borate buffer (pH 8.5) and derivatized employing an AccQ-Tag derivatization strategy adapted from the protocol supplied by Waters (Waters, Etten-Leur, The Netherlands). An Agilent 1290 Infinity II LC system (1290 Multicolumn Thermostat and 1290 High Speed Pump, Agilent Technologies, Waldbronn, Germany) with an Accq-Tag Ultra column (Waters Chromatography B.V., Etten-Leur, The Netherlands) achieved chromatographic separation. The UPLC was coupled to electrospray ionization on a quadrupole-ion trap (SCIEX Qtrap 6500; Sciex, Massachusetts, USA). Analytes were monitored in Multiple Reaction Monitoring (MRM) using nominal mass resolution and detected in the positive ion mode.

4.3.3. GC-MS organic acid platform

In order to extract the organic acids and remove urea, twice liquid-liquid extraction with ethyl acetate was applied to 50 μL of spiked (with internal standards: Succinic acid-2,2,3,3-d_4, Fumaric acid-2,3-d_2, and Citric acid-2,2,4,4-d_4) urine. The two-step online derivatization procedure comprised oximation with methoxyamine hydrochloride (MeOX, 15 mg/mL in pyridine), followed by N-Methyl-N-(trimethylsilyl)- trifluoroacetamide (MSTFA) silylation. After derivatization, sample (1 μL) was injected into the GC-MS. With helium as a carrier gas (1,7 mL/min), chromatic separation was performed on a 25m (HP-5MS UI) film thickness 30 x 0.25m ID column. The mass spectrometer (MSD 5975C, Agilent Technologies, Waldbronn, Germany) was operated using a single quadrupole with electron impact ionization (70 eV) in SCAN mode (mass range m/z 50–500).

4.3.4. LC-MS steroid platform

A LC-MS steroid platform was developed for the current study. Sample preparation comprised adding internal standards (Cortisone-2,2,4,6,6,9,12,12-d_8, and Testosterone-16,16,17-d_3) to 90 μL of urine and filtering the samples with a 0.2μm PTFE membrane. Using an Acquity UPLC CSH C18 column, 130Å, 1.7 μm, 2.1 mm X 100 mm, (Waters, Etten-Leur, The Netherlands), with a flow of 0.4 mL/min over a 15 min gradient, chromatographic separation was achieved by UPLC (Agilent 1290, San Jose, CA, USA). Samples were analyzed using a triple quadrupole mass spectrometer (Agilent 6460, San Jose, CA, USA) with electrospray ionization. By switching positive and negative ion mode, analytes were detected in MRM using nominal mass resolution.

4.3.5. Metabolomics data preprocessing

We removed 4 amines with a missing rate of >10% and excluded metabolites with a relative standard deviation of the quality control samples (RSDqc) equal to or larger than 15% (excluded 6 amines, 1 organic acid, and 3 LC-MS steroids). For the remaining metabolites, we imputed metabolite measurements that fell below the limit of detection/quantification with half of the value of this limit, or when this limit was unknown with half of the lowest observed level for this

metabolite. All metabolites were normalized by dividing by the sample creatinine levels and inverse normal rank transformed.

4.4. Statistical analyses

4.4.1. Associations with sex and age

To assess the influence of sex and age on urinary metabolites we ran mixed-effect models with sex and age as fixed effects and family and zygosity as random effects (to correct for family clustering in the data) (R v. 4.1.1, nlme v. 3.1-153) [55,56]. We expect that some metabolites have lower levels in younger children and will therefore not show sex differences at younger ages. Therefore, we performed a series of sensitivity and stratified analyses to elucidate the associations with sex and age further. We repeated the mixed-effect models with sex and age after dichotomizing age (age<10 vs. age≥10). We chose age 10 as cut-off for dichotomizing the age variable as 10 is the median age in the NTR and LUMC-Curium cohorts. Next, we performed stratified analyses. First, we investigated whether different metabolites show an association with age amongst males and females with mixed-effect models with age as fixed effect and family and zygosity as random effects. Again, these analyses were conducted for age as a continuous measure and after dichotomizing. Second, we employed mixed-effect models with sex as fixed effect and family and zygosity as random effects to investigate whether different metabolites show an association with sex amongst the younger (age<10) and older (age≥10) children. For each analysis we corrected for multiple testing separately with a False Discovery Rate (FDR) of 5% for 86 metabolites setting the significance threshold to $q \leq 0.05$, i.e., 5% of the significant results will be false positives [57]. Thus, in the stratified analyses we correct separately in males and females or younger and older children. We replicated the analyses for sex and age for the urinary metabolites in 179 children of the LUMC-Curium cohort, by using mixed-effect models with family as random effects. We employed FDRs of 5% ($q \leq 0.05$) for 79 (sex + continuous age), 76 (sex + dichotomized age), 73 (sex-stratified continuous age in males), 67 (sex-stratified continuous age in females), 63 (sex-stratified dichotomized aged in males), 51 (sex-stratified dichotomized aged in females), 11 (age-stratified sex in children age<10), and 13 (age-stratified sex in children age≥10) metabolites, respectively.

4.4.2. Genetic analyses

To quantify the contribution of genetic and environmental influences to complex traits, such as metabolites, the classical twin design (CTD) is often used [58]. The CTD hinges on comparing the resemblance between monozygotic (MZ) and dizygotic (DZ) twin pairs. MZ twin pairs derive from the same fertilized egg and share 100% of their genetic material. In contrast, DZ twin pairs derive from separately fertilized eggs and share on average 50% of their genetic material. Thus, by examining the

difference between MZ and DZ covariances for a phenotype of interest, we can derive estimates for the proportions of variance attributable to genetic factors, i.e., heritability (h^2), and to environmental factors. The CTD aims to decompose the phenotypic covariance between MZ and DZ twin pairs into additive genetic (A), non-additive or dominant genetic (D), shared or common environmental (C), and unshared or unique environmental factors (E) [59]. From models that estimate both additive and dominant genetic factors, we can obtain broad-sense h^2 estimates, while models without dominant genetic factors allow us to estimate the narrow-sense h^2 (or additive genetic).

We first fitted two group (MZ + DZ twin pairs) saturated models with sex and age as covariates and equal means and variances across birth order and zygosity by raw-data maximum likelihood (ML) in openMx (v. 2.19.8) [60] to obtain ML estimates of the twin correlations. We compared the saturated model with models that removed the age and/or sex covariates via likelihood ratio-tests. We obtained the MZ and DZ twin correlations from the best-fitting model for each urinary metabolite (**Table S3**). Next, we applied genetic structural equation modelling with raw-data ML estimation in openMx to estimate the additive genetic, dominant genetic or common environmental, and unique environmental variances for the urinary metabolites in the NTR-ACTION cohort based on the twin correlations (**Figure 2**, **Table S4**). We implemented ADE models if the MZ correlations were greater than two times the DZ correlations ($r_{MZ}>2*r_{DZ}$), and ACE models if this was not the case ($r_{MZ}<2*r_{DZ}$). In our genetic models we did not constrain the variance components to take values between zero and one, thus, we allowed for negative variance estimates [50]. ADE or ACE models were estimated including sex and/or age as covariates depending on best-fitting saturated model (**Table S3**). In the ADE model, we focus on the broad-sense h^2, rather than reporting the separate estimates of the standardized variance of A and D. We did not test the significance of the D or C variance components, as the power to detect these variance components is low, given the relatively small number of complete DZ pairs (114) compared to the number of MZ pairs (645) (**Table 1**) [37,38].

Author Contributions

Conceptualization, F.A.H., A.C.H., V.F., B.J.vK., N.K., H.E.H.P., J.R., M.J.J.F., R.R.J.M.V., C.K., M.B., T.H. and D.I.B.; formal analysis, F.A.H.; investigation, F.A.H., J.vD., A.C.H., P.J.R., R.R.J.M.V., C.K., M.B., T.H. and D.I.B.; data curation, R.P.; writing—original draft, F.A.H., J.vD., and D.I.B.; writing— review and editing, F.A.H., J.vD., R.P., A.C.H. P.J.R., B.J.vK., B.R.W., N.K., H.E.H.P., J.R., M.J.J.F., C.K., M.B., and D.I.B.; visualization, F.A.H. and N.K.; supervision, J.vD., M.B. and D.I.B.; funding acquisition,

A.C.H., V.F., H.E.H.P., J.R., M.J.J.F., R.R.J.M.V., C.K., M.B., T.H. and D.I.B. All authors have read and agreed to the published version of the manuscript.

Funding

Institutional Review Board Statement

The study was conducted according to the guidelines of the Declaration of Helsinki, and the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Center, Amsterdam (NTR 25th of May 2007 and ACTION 2013/41 and 2014.252), an Institutional Review Board certified by the U.S. Office of Human Research Protections (IRB number IRB00002991 under Federal-wide Assurance- FWA00017598; IRB/institute codes), and the and the Medical Ethical Committee of Leiden University Medical Center (B15.017, B17.031, B17.032 and B17.040).

Informed Consent Statement

Written informed consent to participate in the ACTION Biomarker Study was provided by the participants' legal guardian for all subjects involved in the studies.

Data Availability Statement

The standardized protocol for large scale collection of urine and buccal-cell samples in the home situation as developed for the ACTION Biomarker Study in children is available at http://www.action-euproject.eu/content/data-protocols. The data of the Netherlands Twin Register (NTR) ACTION Biomarker Study may be accessed, upon approval of the data access committee, through the NTR (https://tweelingenregister.vu.nl/information_for_researchers/working-with-ntr-data).

Acknowledgments

Conflicts of Interest

C.K. was employed by Good Biomarker Sciences (Leiden, the Netherlands). The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Patti, G.J.; Yanes, O.; Siuzdak, G. Innovation: Metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 263–269, doi:10.1038/nrm3314.

2. Dunn, W.B.; Broadhurst, D.I.; Atherton, H.J.; Goodacre, R.; Griffin, J.L. Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chem. Soc. Rev.* **2011**, *40*, 387–426, doi:10.1039/b906712b.

3. Chaleckis, R.; Murakami, I.; Takada, J.; Kondoh, H.; Yanagida, M. Individual variability in human blood metabolites identifies age-related differences. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 4252–4259, doi:10.1073/pnas.1603023113.

4.	Pang, H.; Jia, W.; Hu, Z. Emerging Applications of Metabolomics in Clinical Pharmacology. *Clin. Pharmacol. Ther.* **2019**, *106*, 544–556, doi:10.1002/cpt.1538.

5.	Fiehn, O. Metabolomics--the link between genotypes and phenotypes. *Plant Mol. Biol.* **2002**, *48*, 155–171, doi:10.1023/A:1013713905833.

6.	Shin, S.-Y.; Fauman, E.B.; Petersen, A.-K.; Krumsiek, J.; Santos, R.; Huang, J.; Arnold, M.; Erte, I.; Forgetta, V.; Yang, T.-P.; et al. An atlas of genetic influences on human blood metabolites. *Nat. Genet.* **2014**, *46*, 543–50, doi:10.1038/ng.2982.

7.	Yet, I.; Menni, C.; Shin, S.Y.; Mangino, M.; Soranzo, N.; Adamski, J.; Suhre, K.; Spector, T.D.; Kastenmüller, G.; Bell, J.T. Genetic influences on metabolite levels: A comparison across metabolomic platforms. *PLoS One* **2016**, *11*, doi:10.1371/journal.pone.0153672.

8.	Kettunen, J.; Tukiainen, T.; Sarin, A.-P.; Ortega-Alonso, A.; Tikkanen, E.; Lyytikäinen, L.-P.; Kangas, A.J.; Soininen, P.; Würtz, P.; Silander, K.; et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat. Genet.* **2012**, *44*, 269–276, doi:10.1038/ng.1073.

9.	Tsepilov, Y.A.; Shin, S.Y.; Soranzo, N.; Spector, T.D.; Prehn, C.; Adamski, J.; Kastenmüller, G.; Wang-Sattler, R.; Strauch, K.; Gieger, C.; et al. Nonadditive effects of genes in human metabolomics. *Genetics* **2015**, *200*, 707–718, doi:10.1534/genetics.115.175760.

10.	Pool, R.; Hagenbeek, F.A.; Hendriks, A.M.; van Dongen, J.; Willemsen, G.; de Geus, E.; Willems van Dijk, K.; Verhoeven, A.; Suchiman, H.E.; Beekman, M.; et al. Genetics and Not Shared Environment Explains Familial Resemblance in Adult Metabolomics Data. *Twin Res. Hum. Genet.* **2020**, *23*, 145–155, doi:10.1017/thg.2020.53.

11.	Frahnow, T.; Osterhoff, M.A.; Hornemann, S.; Kruse, M.; Surma, M.A.; Klose, C.; Simons, K.; Pfeiffer, A.F.H. Heritability and responses to high fat diet of plasma lipidomics in a twin study. *Sci. Rep.* **2017**, *7*, 1–11, doi:10.1038/s41598-017-03965-6.

12.	Wong, M.W.K.; Thalamuthu, A.; Braidy, N.; Mather, K.A.; Liu, Y.; Ciobanu, L.; Baune, B.T.; Armstrong, N.J.; Kwok, J.; Schofield, P.; et al. Genetic and environmental determinants of variation in the plasma lipidome of older Australian twins. *Elife* **2020**, *9*, 1–28, doi:10.7554/eLife.58954.

13.	Bermingham, K.M.; Brennan, L.; Segurado, R.; Gray, I.J.; Barron, R.E.; Gibney, E.R.; Ryan, M.F.; Gibney, M.J.; Newman, J.W.; O'Sullivan, D.A.M. Genetic and environmental influences on serum oxylipins, endocannabinoids, bile acids and steroids. *Prostaglandins, Leukot. Essent.*

Fat. Acids **2021**, *173*, 102338, doi:10.1016/j.plefa.2021.102338.

14. Bellis, C.; Kulkarni, H.; Mamtani, M.; Kent, J.W.; Wong, G.; Weir, J.M.; Barlow, C.K.; Diego, V.; Almeida, M.; Dyer, T.D.; et al. Human Plasma Lipidome Is Pleiotropically Associated With Cardiovascular Risk Factors and Death. *Circ. Cardiovasc. Genet.* **2014**, *7*, 854–863, doi:10.1161/CIRCGENETICS.114.000600.

15. Darst, B.F.; Koscik, R.L.; Hogan, K.J.; Johnson, S.C.; Engelman, C.D. Longitudinal plasma metabolomics of aging and sex. *Aging (Albany. NY).* **2019**, *11*, 1262–1282, doi:10.18632/aging.101837.

16. Rhee, E.P.; Ho, J.E.; Chen, M.-H.; Shen, D.; Cheng, S.; Larson, M.G.; Ghorbani, A.; Shi, X.; Helenius, I.T.; O'Donnell, C.J.; et al. A genome-wide association study of the human metabolome in a community-based cohort. *Cell Metab.* **2013**, *18*, 130–143, doi:10.1016/j.cmet.2013.06.013.

17. Cadby, G.; Melton, P.E.; McCarthy, N.S.; Giles, C.; Mellett, N.A.; Huynh, K.; Hung, J.; Beilby, J.; Dubé, M.-P.; Watts, G.F.; et al. Heritability of 596 lipid species and genetic correlation with cardiovascular traits in the Busselton Family Heart Study. *J. Lipid Res.* **2020**, *61*, 537–545, doi:10.1194/jlr.RA119000594.

18. Hagenbeek, F.A.; Pool, R.; van Dongen, J.; Draisma, H.H.M.; Jan Hottenga, J.; Willemsen, G.; Abdellaoui, A.; Fedko, I.O.; den Braber, A.; Visser, P.J.; et al. Heritability estimates for 361 blood metabolites across 40 genome-wide association studies. *Nat. Commun.* **2020**, *11*, 39, doi:10.1038/s41467-019-13770-6.

19. Feofanova, E. V.; Chen, H.; Dai, Y.; Jia, P.; Grove, M.L.; Morrison, A.C.; Qi, Q.; Daviglus, M.; Cai, J.; North, K.E.; et al. A Genome-wide Association Study Discovers 46 Loci of the Human Metabolome in the Hispanic Community Health Study/Study of Latinos. *Am. J. Hum. Genet.* **2020**, *107*, 849–863, doi:10.1016/j.ajhg.2020.09.003.

20. Rhee, E.P.; Yang, Q.; Yu, B.; Liu, X.; Cheng, S.; Deik, A.; Pierce, K.A.; Bullock, K.; Ho, J.E.; Levy, D.; et al. An exome array study of the plasma metabolome. *Nat. Commun.* **2016**, *7*, 12360, doi:10.1038/ncomms12360.

21. Tabassum, R.; Rämö, J.T.; Ripatti, P.; Koskela, J.T.; Kurki, M.; Karjalainen, J.; Palta, P.; Hassan, S.; Nunez-Fontarnau, J.; Kiiskinen, T.T.J.; et al. Genetic architecture of human plasma lipidome and its link to cardiovascular disease. *Nat. Commun.* **2019**, *10*, 4329, doi:10.1038/s41467-019-11954-8.

22. Rietschel, L.; Streit, F.; Zhu, G.; McAloney, K.; Frank, J.; Couvy-Duchesne, B.; Witt, S.H.; Binz, T.M.; McGrath, J.; Hickie, I.B.; et al. Hair Cortisol in Twins: Heritability and Genetic Overlap with Psychological Variables and Stress-System Genes. *Sci. Rep.* **2017**, *7*, 15351, doi:10.1038/s41598-017-11852-3.

23. van Keulen, B.J.; Dolan, C. V; Andrew, R.; Walker, B.R.; Hulshoff Pol, H.E.; Boomsma, D.I.; Rotteveel, J.; Finken, M.J.J. Heritability of cortisol production and metabolism throughout adolescence: a twin study. *J. Clin. Endocrinol. Metab.* **2019**, 1–28, doi:10.1210/clinem/dgz016.

24. Coviello, A.D.; Zhuang, W. V.; Lunetta, K.L.; Bhasin, S.; Ulloor, J.; Zhang, A.; Karasik, D.; Kiel, D.P.; Vasan, R.S.; Murabito, J.M. Circulating Testosterone and SHBG Concentrations Are Heritable in Women: The Framingham Heart Study. *J. Clin. Endocrinol. Metab.* **2011**, *96*, E1491–E1495, doi:10.1210/jc.2011-0050.

25. Travison, T.G.; Zhuang, W. V.; Lunetta, K.L.; Karasik, D.; Bhasin, S.; Kiel, D.P.; Coviello, A.D.; Murabito, J.M. The heritability of circulating testosterone, oestradiol, oestrone and sex hormone binding globulin concentrations in men: the Framingham Heart Study. *Clin. Endocrinol. (Oxf).* **2014**, *80*, 277–282, doi:10.1111/cen.12260.

26. Grotzinger, A.D.; Briley, D.A.; Engelhardt, L.E.; Mann, F.D.; Patterson, M.W.; Tackett, J.L.; Tucker-Drob, E.M.; Harden, K.P. Genetic and environmental influences on pubertal hormones in human hair across development. *Psychoneuroendocrinology* **2018**, *90*, 76–84, doi:10.1016/j.psyneuen.2018.02.005.

27. Miller, I.J.; Peters, S.R.; Overmyer, K.A.; Paulson, B.R.; Westphall, M.S.; Coon, J.J. Real-time health monitoring through urine metabolomics. *npj Digit. Med.* **2019**, *2*, 109, doi:10.1038/s41746-019-0185-y.

28. Bouatra, S.; Aziat, F.; Mandal, R.; Guo, A.C.; Wilson, M.R.; Knox, C.; Bjorndahl, T.C.; Krishnamurthy, R.; Saleem, F.; Liu, P.; et al. The human urine metabolome. *PLoS One* **2013**, *8*, e73076, doi:10.1371/journal.pone.0073076.

29. Wishart, D.S.; Guo, A.; Oler, E.; Wang, F.; Anjum, A.; Peters, H.; Dizon, R.; Sayeeda, Z.; Tian, S.; Lee, B.L.; et al. HMDB 5.0: the Human Metabolome Database for 2022. *Nucleic Acids Res.* **2022**, *50*, D622–D631, doi:10.1093/nar/gkab1062.

30. Calvo-Serra, B.; Maitre, L.; Lau, C.H.E.; Siskos, A.P.; Gützkow, K.B.; Andrušaitytė, S.; Casas, M.; Cadiou, S.; Chatzi, L.; González, J.R.; et al. Urinary metabolite quantitative trait loci in children

and their interaction with dietary factors. *Hum. Mol. Genet.* **2021**, *29*, 3830–3844, doi:10.1093/hmg/ddaa257.

31. Bartels, M.; Hendriks, A.; Mauri, M.; Krapohl, E.; Whipp, A.; Bolhuis, K.; Conde, L.C.; Luningham, J.; Fung Ip, H.; Hagenbeek, F.; et al. Childhood aggression and the co-occurrence of behavioural and emotional problems: results across ages 3–16 years from multiple raters in six cohorts in the EU-ACTION project. *Eur. Child Adolesc. Psychiatry* **2018**, *27*, 1105–1121, doi:10.1007/s00787-018-1169-1.

32. Boomsma, D.I. Aggression in children: unravelling the interplay of genes and environment through (epi)genetics and metabolomics. *J. Pediatr. Neonatal Individ. Med.* **2015**, *4*, e040251, doi:10.7363/040251.

33. Hagenbeek, F.A.; Roetman, P.J.; Pool, R.; Kluft, C.; Harms, A.C.; van Dongen, J.; Colins, O.F.; Talens, S.; van Beijsterveldt, C.E.M.; Vandenbosch, M.M.L.J.Z.; et al. Urinary Amine and Organic Acid Metabolites Evaluated as Markers for Childhood Aggression: The ACTION Biomarker Study. *Front. Psychiatry* **2020**, *11*, doi:10.3389/fpsyt.2020.00165.

34. Lykken, D.T. Research With Twins: The Concept of Emergenesis. *Psychophysiology* **1982**, *19*, 361–372, doi:10.1111/j.1469-8986.1982.tb02489.x.

35. Cicchetti, D. V. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol. Assess.* **1994**, *6*, 284–290, doi:10.1037/1040-3590.6.4.284.

36. Posthuma, D.; Beem, a L.; de Geus, E.J.C.; van Baal, G.C.M.; von Hjelmborg, J.B.; Iachine, I.; Boomsma, D.I. Theory and Practice in Quantitative Genetics. *Twin Res.* **2003**, *6*, 361–376, doi:10.1375/136905203770326367.

37. Visscher, P.M. Power of the classical twin design revisited. *Twin Res.* **2004**, *7*, 505–12, doi:10.1375/1369052042335250.

38. Visscher, P.M.; Gordon, S.; Neale, M.C. Power of the Classical Twin Design Revisited: II Detection of Common Environmental Variance. *Twin Res. Hum. Genet.* **2008**, *11*, 48–54, doi:10.1375/twin.11.1.48.

39. Dettling, A.C.; Gunnar, M.R.; Donzella, B. Cortisol levels of young children in full-day childcare centers: relations with age and temperament. *Psychoneuroendocrinology* **1999**, *24*, 519–536, doi:10.1016/S0306-4530(99)00009-8.

40. Shi, L.; Wudy, S.A.; Buyken, A.E.; Maser-Gluth, C.; Hartmann, M.F.; Remer, T. Prepubertal

glucocorticoid status and pubertal timing. *J. Clin. Endocrinol. Metab.* **2011**, *96*, 891–898, doi:10.1210/jc.2010-2935.

41.   van Keulen, B.J.; Dolan, C. V.; van der Voorn, B.; Andrew, R.; Walker, B.R.; Pol, H.H.; Boomsma, D.I.; Rotteveel, J.; Finken, M.J.J. Sexual dimorphism in cortisol metabolism throughout pubertal development: A longitudinal study. *Endocr. Connect.* **2020**, *9*, 542–551, doi:10.1530/EC-20-0123.

42.   Akalestou, E.; Genser, L.; Rutter, G.A. Glucocorticoid Metabolism in Obesity and Following Weight Loss. *Front. Endocrinol. (Lausanne).* **2020**, *11*, 1–9, doi:10.3389/fendo.2020.00059.

43.   Guneral, F.; Bachmann, C. Age-related reference values for urinary organic acids in a healthy Turkish pediatric population. *Clin. Chem.* **1994**, *40*, 862–868.

44.   Chugani, D.C.; Muzik, O.; Behen, M.; Rothermel, R.; Janisse, J.J.; Lee, J.; Chugani, H.T. Developmental changes in brain serotonin synthesis capacity in autistic and nonautistic children. *Ann. Neurol.* **1999**, *45*, 287–295, doi:10.1002/1531-8249(199903)45:3<287::AID-ANA3>3.0.CO;2-9.

45.   Gu, H.; Pan, Z.; Xi, B.; Hainline, B.E.; Shanaiah, N.; Asiago, V.; Gowda, G.A.N.; Raftery, D. 1H NMR metabolomics study of age profiling in children. *NMR Biomed.* **2009**, *22*, 826–833, doi:10.1002/nbm.1395.

46.   Mayo Clinic Laboratories Amino Acids, Quantitative, Random, Urine Available online: https://www.mayocliniclabs.com/test-catalog/overview/60475#Clinical-and-Interpretive (accessed on Apr 25, 2022).

47.   Lau, C.-H.E.; Siskos, A.P.; Maitre, L.; Robinson, O.; Athersuch, T.J.; Want, E.J.; Urquiza, J.; Casas, M.; Vafeiadi, M.; Roumeliotaki, T.; et al. Determinants of the urinary and serum metabolome in children from six European populations. *BMC Med.* **2018**, *16*, 202, doi:10.1186/s12916-018-1190-8.

48.   Bermingham, K.M.; Brennan, L.; Segurado, R.; Barron, R.E.; Gibney, E.R.; Ryan, M.F.; Gibney, M.J.; O'Sullivan, A.M. Genetic and environmental influences on covariation in reproducible diet–metabolite associations. *Am. J. Clin. Nutr.* **2021**, *113*, 1232–1240, doi:10.1093/ajcn/nqaa378.

49.   Bermingham, K.M.; Brennan, L.; Segurado, R.; Barron, R.E.; Gibney, E.R.; Ryan, M.F.; Gibney, M.J.; O'Sullivan, A.M. Genetic and Environmental Contributions to Variation in the Stable Urinary NMR Metabolome over Time: A Classic Twin Study. *J. Proteome Res.* **2021**, *20*, 3992–

4000, doi:10.1021/acs.jproteome.1c00319.

50. Verhulst, B.; Prom-Wormley, E.; Keller, M.; Medland, S.; Neale, M.C. Type I Error Rates and Parameter Bias in Multivariate Behavioral Genetic Models. *Behav. Genet.* **2019**, *49*, 99–111, doi:10.1007/s10519-018-9942-y.

51. Lykken, D.T. The mechanism of emergenesis. *Genes, Brain Behav.* **2006**, *5*, 306–310, doi:10.1111/j.1601-183X.2006.00233.x.

52. Eaves, L.J. Dominance alone is not enough. *Behav. Genet.* **1988**, *18*, 27–33, doi:10.1007/BF01067073.

53. Ligthart, L.; van Beijsterveldt, C.E.M.; Kevenaar, S.T.; de Zeeuw, E.; van Bergen, E.; Bruins, S.; Pool, R.; Helmer, Q.; van Dongen, J.; Hottenga, J.-J.; et al. The Netherlands Twin Register: Longitudinal Research Based on Twin and Twin-Family Designs. *Twin Res. Hum. Genet.* **2019**, *22*, 623–636, doi:10.1017/thg.2019.93.

54. Odintsova, V. V.; Willemsen, G.; Dolan, C. V.; Hottenga, J.J.; Martin, N.G.; Slagboom, P.E.; Ordoñana, J.R.; Boomsma, D.I. Establishing a Twin Register: An Invaluable Resource for (Behavior) Genetic, Epidemiological, Biomarker, and "Omics" Studies. *Twin Res. Hum. Genet.* **2018**, *21*, 239–252, doi:10.1017/thg.2018.23.

55. R Core Team R: A language and environment for statistical computing Available online: https://www.r-project.org/.

56. Lindstrom, M.J.; Bates, D.M. Nonlinear Mixed Effects Models for Repeated Measures Data. *Biometrics* **1990**, *46*, 673–687.

57. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 1995, *57*, 289–300.

58. Boomsma, D.I.; Busjahn, A.; Peltonen, L. Classical twin studies and beyond. *Nat. Rev. Genet.* **2002**, *3*, 872–882, doi:10.1038/nrg932.

59. Rijsdijk, F. V; Sham, P.C. Analytic approaches to twin data using structural equation models. *Brief. Bioinform.* **2002**, *3*, 119–33.

60. Boker, S.; Neale, M.; Maes, H.; Wilde, M.; Spiegel, M.; Brick, T.; Spies, J.; Estabrook, R.; Kenny, S.; Bates, T.; et al. OpenMx: An Open Source Extended Structural Equation Modeling Framework. *Psychometrika* **2011**, *76*, 306–317, doi:10.1007/s11336-010-9200-6.