

1 **Lifelong impacts of puberty timing on human plasma metabolic**
2 **profiles**

3 **Zengjun Li, MB¹; Si Fang, MS^{2,3}; Dong Liu, MPH⁴; Fei Li*, MD^{4,5,6}, Cairong**
4 **Zhu*, PhD¹; Jian Zhao*, PhD^{4,6}**

5 **Affiliations:**

6 1. Department of Epidemiology and Biostatistics, West China School of Public Health
7 and West China Fourth Hospital, Sichuan University, 610041, China.

8 2. Population Health Sciences, Bristol Medical School, University of Bristol, Bristol,
9 United Kingdom.

10 3. Medical Research Council (MRC) Integrative Epidemiology Unit (IEU), University
11 of Bristol, Bristol, United Kingdom.

12 4. Ministry of Education and Shanghai Key Laboratory of Children's Environmental
13 Health, Xinhua Hospital, Shanghai Jiao Tong University School of Medicine,
14 Shanghai, 200092, China

15 5. Department of Developmental and Behavioral Pediatric & Child Primary Care,
16 Brain and Behavioral Research Unit of Shanghai Institute for Pediatric Research,
17 Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai,
18 China

19 6. Department of Maternal and Child Health, School of Public Health, Shanghai Jiao
20 Tong University, 200092, China.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

21 ***Corresponding author:**

22 Dr Jian Zhao, Ministry of Education and Shanghai Key Laboratory of Children's
23 Environmental Health, Xinhua Hospital, Shanghai Jiao Tong University School of
24 Medicine, Shanghai, 200092, China (jzhao.epi@gmail.com).

25 Dr Cairong Zhu, Department of Epidemiology and Biostatistics, West China School
26 of Public Health and West China Fourth Hospital, Sichuan University, 610041, China
27 (cairong.zhu@hotmail.com).

28 Dr Fei Li, Department of Developmental and Behavioral Pediatric & Child Primary
29 Care, Brain and Behavioral Research Unit of Shanghai Institute for Pediatric Research,
30 Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai,
31 200092, China (feili@shsmu.edu.cn).

32

33 **Abstract**

34 There has been uncertainty regarding the long-term impact of puberty timing on
35 human plasma metabolites. This lack of clarity can be attributed to the influence of
36 confounding factors present in conventional observational studies. To determine the
37 causal effect of puberty timing on plasma metabolites, we employed a two-sample
38 Mendelian randomization (MR) analysis, complemented by MR mediation analysis
39 assessing the direct effect. We utilized data from a large-scale meta-analysis of
40 genome-wide association studies (GWAS) on puberty timing, consisting of 329,345
41 women of European ancestry, and a meta-analysis of GWAS on plasma metabolites,
42 involving up to 86,507 individuals. Our findings provide moderate evidence
43 supporting a causal effect of puberty timing on 23 out of 174 plasma metabolites.
44 After excluding 7 single nucleotide polymorphisms (SNPs) related to birth weight and
45 childhood adiposity, causal effects remained for 16 metabolites. Through two-step
46 MR analysis, we observed strong evidence that adulthood adiposity mediated the
47 causal relationships of puberty timing on 35 plasma metabolites. We also observed
48 moderate evidence for an independent causal effect of puberty timing on 10
49 metabolites through multivariable MR analysis. We further used metabolomic data
50 measured in the UK Biobank (UKB) to perform a replication analysis to validate the
51 causal effect estimated. Nine amino acids were identified in the UKB, and the
52 replication analysis supported our main findings.

53 **Introduction**

54 The onset of puberty, influenced by a combination of genetic and environmental
55 factors ^{1,2}, plays a pivotal role in shaping growth and development throughout an
56 individual's lifespan ³. Early puberty onset has been associated with various

57 unfavorable health outcomes, including childhood and adulthood obesity ^{4,5}, type 2
58 diabetes ⁶, hypertension ⁷, breast cancer ⁸, and mental disorders ⁹. The hormonal
59 fluctuations during puberty initiation trigger metabolic changes that can have
60 significant implications ³. The investigation of the causal relationship between puberty
61 timing and plasma metabolites in adulthood is of utmost importance, given the
62 increasing recognition of plasma metabolites as effective biomarkers for illness
63 diagnosis and prognosis ^{10,11}. This exploration will provide insights into the long-
64 lasting effects of early-life determinants, such as puberty onset, on human metabolism
65 later in life. Unfortunately, previous studies on this subject have been limited,
66 predominantly relying on conventional multivariable analysis methods that are
67 susceptible to confounding or reverse causation biases ^{5,12,13}. Additionally, the
68 utilization of longitudinal studies for the examination of the correlation between
69 puberty timing and the concentration of plasma metabolites necessitates an extended
70 duration of follow-up, resulting in limited empirical support.

71 Mendelian randomization (MR) is an increasingly used study design that employs
72 genetic variants as instrumental variables to draw causal inferences between an
73 exposure of interest (e.g., puberty timing) and an outcome (e.g., a metabolite) ¹⁴. MR
74 analyses are much less vulnerable to conventional confounding as a result of the
75 random assignment of alleles during meiosis ¹⁵, hence differentiating MR from normal
76 observational research. Moreover, MR analyses are less prone to reverse causation
77 because the assortment of alleles precedes the initiation of exposure (e.g., puberty
78 timing).

79 A recent MR study investigated if puberty timing has a distinct influence on adulthood
80 cardiometabolic traits ¹⁶. The results suggested that the effects of puberty timing on

81 adulthood adiposity and cardiometabolic traits are not likely driven by itself but by
82 childhood adiposity. Additionally, it is worth noting that the majority of the
83 cardiometabolic characteristics examined in this study mostly focused on lipids,
84 resulting in a very restricted range of metabolites investigated. Consequently, in this
85 study, we aimed to examine whether puberty timing is likely to, directly and/or
86 indirectly, impact a wide range of plasma metabolites, including amino acids,
87 biogenic amines, acylcarnitines, phosphatidylcholines, lysophosphatidylcholines,
88 sphingomyelins, and hexose in adulthood in a two-sample MR framework. Along with
89 conducting univariable MR analysis to infer the total effect of puberty timing on
90 plasma metabolites, we also performed MR mediation analysis to take the mediating
91 effect of adulthood adiposity into account.

92 **Results**

93 **Lifelong causal effects of puberty timing on human metabolites**

94 The study design is illustrated in **Fig 1**. In the primary analysis, all genetically
95 significant SNPs were used as instrument variables. Additionally, to minimize the bias
96 arising from horizontal pleiotropy, a secondary analysis was also performed by
97 excluding SNPs that were genetically associated with confounders (e.g., birth weight
98 and childhood adiposity) (**Fig 2**). Two-sample MR analysis suggested moderate
99 causal associations between genetically predicted puberty timing and 23 plasma
100 metabolites within the subclasses including acylcarnitines, amino acids, biogenic
101 amines, and lysophosphatidylcholines based on the primary analysis (i.e., inverse-
102 variance weighted (IVW) analysis) results (**Fig 3-4; Table S7**) with *P* values ranging
103 from 9.09×10^{-4} to 4.92×10^{-2} . The MR analysis results supported positive causal
104 associations of later puberty timing with higher levels of 20 metabolites. The strongest

105 causal evidence was found for dodecanedioylcarnitine (beta = 0.11, 95% CI: 0.04 to
106 0.17, $P = 9.09 \times 10^{-4}$), which means each year delayed puberty onset was associated
107 with 0.11 standard deviation (SD) units higher of dodecanedioylcarnitine. By contrast,
108 lower levels of three amino acids, aspartate (beta = -0.04, 95% CI: -7.19×10^{-2} to -
109 6.40×10^{-4} , $P = 4.60 \times 10^{-2}$), phenylalanine (beta = -0.03, 95% CI: -5.03×10^{-2} to -
110 1.20×10^{-3} , $P = 3.98 \times 10^{-2}$), and tyrosine (beta = -0.03, 95% CI: -5.17×10^{-2} to -8.46
111 $\times 10^{-3}$, $P = 6.41 \times 10^{-3}$), were observed per one year later onset of puberty.

112 For the 23 metabolites indicating moderate causal consequences of puberty timing,
113 causal effect estimates in sensitivity analyses were broadly consistent with the results
114 of the primary IVW analysis (**Table S8-S9**). Of note, intercept terms estimated in the
115 MR-Egger regression analysis suggested that some of the results might be biased by
116 horizontal pleiotropy including aspartate, phenylalanine, alpha-aminoadipic acid,
117 lysoPC a C16:0, PC aa C38:0 and PC aa C40:1, which was also supported by
118 evidence of heterogeneity tested via Cochran's Q statistic in the IVW analysis (**Table**
119 **S10-S11**).

120 After excluding 7 SNPs associated with birth weight and childhood body mass index
121 (BMI), the secondary IVW analysis confirmed moderate evidence for causal
122 associations of later puberty timing with higher levels of 16 metabolites out of the
123 aforementioned 23 metabolites in the primary two-sample MR analysis. It is
124 noteworthy that 3 additional metabolites, including dodecenoylcarnitine: beta = 0.06,
125 95% CI: 2.75×10^{-3} to 1.10×10^{-2} , $P = 3.94 \times 10^{-2}$; tryptophan: beta = 0.04, 95% CI:
126 5.61×10^{-3} to 8.25×10^{-2} , $P = 2.47 \times 10^{-2}$; and kynurenine: beta = 0.06, 95% CI: 1.57
127 $\times 10^{-2}$ to 9.58×10^{-2} , $P = 6.38 \times 10^{-3}$, appeared to be impacted by puberty timing in
128 the secondary IVW analysis, after controlling for potentially horizontal pleiotropic

129 effects via birth weight or childhood adiposity. On the contrary, 10 out of the 23
130 putative metabolites observed in the primary analysis turned to statistically non-
131 significant in the secondary IVW analysis (**Fig 4, Table S12**), which included
132 phenylalanine, threonine, asparagine, aspartate, tyrosine, trans-hydroxyproline,
133 lysoPC a C17:0, lysoPC a C18:0, lysoPC a C18:1, and lysoPC a C18:2. Among them,
134 intercept terms estimated in the MR-Egger regression analysis suggested possible
135 horizontal pleiotropy for phenylalanine and heterogeneity for aspartate and lysoPC a
136 C18:1 (**Table S13-S14**).

137 **Mediation analysis**

138 The estimated causal effects of adulthood BMI on 174 metabolites suggested that
139 adulthood BMI strongly influenced plasma concentrations of some metabolites
140 through two-step MR analysis, including 35 metabolites reaching nominal
141 significance and 13 reaching statistical significance after the multiple testing
142 Bonferroni correction (**Table S15**). Further, we found a causal effect of puberty
143 timing (beta = -0.13, 95% CI: -0.17 to -0.09, $P = 8.71 \times 10^{-12}$) on the mediator (i.e.,
144 adulthood BMI). By exploiting the "product of coefficients" method, we derived the
145 indirect effects of puberty timing on each plasma metabolite (**Table S15**). For
146 instance, we found that each year of delayed puberty onset was associated with 0.01
147 SD units increase in serine through the pathway of adulthood adiposity (beta = 0.01,
148 95% CI: 1.04×10^{-3} to 2.44×10^{-2} , $P = 0.03$), indicating that adulthood adiposity was
149 a putative factor that mediated the causal relationships of puberty timing on certain
150 metabolites. To further rule out the influence of horizontal pleiotropy, we removed a
151 total of 13 SNPs highly related to childhood adiposity and birth weight from
152 adulthood BMI and puberty timing genetic instrument sets. A total of 31 metabolites

153 reached nominal significance and 8 reached the multiple testing Bonferroni corrected
154 significance level. Moreover, the effect size of puberty timing on adulthood adiposity
155 attenuated from -0.13 kg/m^2 to -0.09 kg/m^2 (beta= -0.09 , 95% CI: -0.11 to -0.06 , $P =$
156 1.81×10^{-8}). Nevertheless, the mediating role of adulthood BMI between puberty
157 timing and adulthood metabolites persisted when the potential pleiotropic effects were
158 adjusted for (**Table S17**).

159 The MVMR analysis results supported causal evidence for direct causal effects of
160 puberty timing on 10 metabolites (i.e., carnitine, dodecanedioylcarnitine,
161 hydroxytetradecenoylcarnitine, tiglylcarnitine, lysine, ornithine, cis-hydroxyproline,
162 trans-hydroxyproline, lysoPC a C20:4, and PC aa C40:4) based on nominal
163 significance threshold of $P < 0.05$. Moreover, among the 10 significant metabolites,
164 the results suggested that carnitine, dodecanedioylcarnitine, tiglylcarnitine
165 (acylcarnitines), lysine (amino acids), and lysoPC a C20:4 (lysophosphatidylcholines)
166 were directly affected by puberty timing and had the same direction as results in the
167 primary analysis (**Fig 5; Table S16**). Notably, the effect sizes attenuated to varying
168 degrees, because adulthood BMI was adjusted for in the MVMR analysis. For
169 example, the coefficient for dodecanedioylcarnitine attenuated from 0.11 SD units to
170 0.07 SD units per year later onset of puberty, approximately half of the original
171 estimate. Compared with the results in the primary MR analysis, except for
172 phenylalanine and tyrosine (amino acids) which showed inconsistent directions, the
173 remaining 21 metabolites showed the same direction as the primary two-sample MR
174 analysis. Furthermore, the MVMR analysis suggested that the plasma levels of 17
175 metabolites (i.e., hydroxyhexadecenoylcarnitine, nonacylcarnitine, valerylcarnitine,
176 asparagine, aspartate, methionine, phenylalanine, serine, threonine, tyrosine, sarcosine,
177 trans-hydroxyproline, lysoPC a C16:0, lysoPC a C17:0, lysoPC a C18:0, lysoPC a

178 C18:1, lysoPC a C18:2) were mediated by adulthood BMI rather than directly affected
179 by puberty timing (**Table S16; Table S18**).

180 **Replication analysis**

181 We attempted to replicate the effect estimates for puberty timing on the identified nine
182 metabolites in the UKB. In the two-sample MR analysis, a nominally significant
183 causal relationship was revealed between puberty timing and two amino acids, namely
184 phenylalanine (beta = -0.03, 95% CI: -5.38×10^{-2} to -2.45×10^{-3} , $P = 3.18 \times 10^{-2}$) and
185 valine (beta = -0.04, 95% CI: -6.54×10^{-2} to -3.82×10^{-3} , $P = 2.77 \times 10^{-2}$, although
186 none of the nine amino acids reached the strictly adjusted P threshold of 5.56×10^{-3}
187 (0.05/9) (**Fig 6; Table S19**). Causal effect estimates in the replication analysis of the
188 UKB data were consistent with those estimated in the primary two-sample MR
189 analysis except for isoleucine and glutamine. Notably, consistent with the primary
190 MVMR analysis results, in the MVMR analysis using the UKB data, the causal
191 effects attenuated towards null after adjusting for adulthood adiposity, suggesting little
192 direct effects were observed between puberty timing and the nine amino acids (**Fig 6;**
193 **Table S20**).

194 **Discussion**

195 Plasma levels of metabolites have been investigated as promising biomarkers for
196 disease risk stratification and critical molecular targets for developing novel
197 treatments and interventions for a range of diseases and conditions¹⁷⁻²⁰. Given that
198 puberty timing is consistently associated with the risk of developing various chronic
199 diseases in adulthood²¹, further causal insights into the relationship between puberty
200 timing and plasma levels of metabolites will help understand underlying biological

201 mechanisms and identify possible putative target biomarkers for disease treatment or
202 intervention. Here, we systematically and extensively investigated the lifelong
203 impacts of puberty timing on human plasma metabolic profiles, by integrating data on
204 human puberty timing and plasma metabolism in the MR causal inference framework.
205 Genetically predicted puberty timing appeared to have causal effects on 23 plasma
206 metabolites in adulthood. After considering the potential mediating role of adulthood
207 adiposity, direct causal effects of puberty timing on 10 adulthood metabolites
208 remained, suggesting the causal effects of puberty timing on certain metabolites in
209 adulthood were mediated by adulthood adiposity. Our findings in the primary MR
210 analysis were further supported by the results from the mediation analysis and
211 external replication analysis in the UKB.

212 Among the 23 plasma metabolites identified in the two-sample MR analysis, seven
213 belong to acylcarnitines. In MVMR analysis, 5 of the 10 metabolites affected by
214 puberty timing were also confirmed within the subclass of acylcarnitines, including
215 carnitine, tiglylcarnitine (short-chain), dodecanedioylcarnitine,
216 hydroxytetradecenoylcarnitine, and octadecenoylcarnitine (medium- and long-chain).
217 With the causal effect estimates ranging from 0.05 SD to 0.07 SD per year of delayed
218 puberty onset, our findings suggested that certain acylcarnitine metabolites were
219 directly and independently impacted by puberty timing. Acylcarnitines were a group
220 of compounds formed from carnitine and acyl-CoAs in mitochondria or peroxisomal
221 by carnitine acyltransferases^{22,23}, which can cross mitochondrial and cell membranes.
222 Short-chain acylcarnitines are generated from the degradation of glucose, amino acids,
223 and fatty acids. Medium- and long-chain acylcarnitines are largely produced from
224 fatty acids participating in the transportation of fatty acids for the following fatty acids
225 oxidation²⁴. It has been well understood that increased plasma levels of acylcarnitine

226 are associated with a higher risk of cardiovascular disease²⁵⁻²⁷. The underlying
227 mechanism might be the accumulation of acylcarnitine was associated with a
228 disturbed β -oxidation rate and mitochondrial dysfunction^{28,29}. In addition, previous
229 studies have shown that a shorter reproductive lifespan is associated with an increased
230 risk of cardiovascular disease events³⁰. Furthermore, in our study, we found a positive
231 causal effect of delayed puberty timing on the accumulation of acylcarnitine. Taking
232 together, our findings may shed new light on the correlation between puberty timing
233 and cardiovascular diseases. However, whether puberty timing affects cardiovascular
234 diseases through the acylcarnitine metabolism pathway has not been illuminated, thus
235 more studies are warranted.

236 Of note, we found 8 amino acids in adulthood were significantly impacted by puberty
237 timing in the primary MR analysis, among which inverse causal effects of puberty
238 timing on the concentrations of tyrosine and phenylalanine were externally validated
239 using data from the UKB. After accounting for adulthood adiposity in the MVMR
240 analysis, the direct causal effects of puberty timing on these 8 amino acids attenuated
241 to null in both the primary and replication MVMR analysis. It suggests a mediating
242 role of adulthood adiposity in the causal pathway from puberty timing to the
243 aforementioned amino acids. Our findings were partly consistent with the findings
244 from a previous study¹⁶, where an inverse association between age at menarche and
245 phenylalanine measured at 18 years old was revealed in the conventional
246 observational analysis among the females in the Avon Longitudinal Study of Parents
247 and Children (ALSPAC), but a causal association was failed to be established in the
248 one-sample MR analysis using the same ALSPAC cohort data.

249 In our two-step MR analysis, we found that adulthood BMI influenced the levels of 35
250 metabolites, and adulthood BMI itself was also influenced by puberty timing. After
251 adjusting for adulthood BMI through MVMR analysis, the number of metabolites
252 influenced by puberty timing decreased from 23 to 10. Our findings suggested that
253 adulthood BMI may serve as a mediator between puberty timing and plasma
254 metabolites. The reason lies in the fact that early-maturing individuals may face a
255 higher risk of obesity and exhibit a higher likelihood of metabolic abnormalities in
256 adulthood³¹. Obesity can also significantly increase the risk of developing metabolic
257 disorders by affecting energy metabolism and altering pathways³², which could be the
258 reason for changes in acylcarnitine and amino acid metabolism.

259 The strengths of this study were as follows. First, the data we used were curated from
260 the latest and largest GWAS in European ancestry. Second, the MR mediation analysis
261 makes it possible to infer the direct and indirect effects of genetically predicted
262 puberty timing on adulthood plasma levels of metabolites. In the two-step MR
263 analysis, we used adulthood BMI as a mediator to investigate the indirect effects of
264 puberty timing on metabolites. As for MVMR analysis, we included both adulthood
265 BMI and puberty timing as exposures to explore the direct effect of puberty timing on
266 metabolites. Third, to minimize the risk of bias due to horizontal pleiotropy, we
267 excluded SNPs that had genome-wide significant associations with birth weight or
268 childhood obesity. Finally, external replication analysis using data from the UKB
269 supported our primary analysis results, suggesting the robustness of our findings.

270 Limitations were also noted. First, our study results were primarily based on GWAS
271 conducted in individuals of European descent. Consequently, it should be cautious
272 when generalizing the results to non-European populations. Second, despite extensive

273 analyses were performed to minimize the risk of bias due to horizontal pleiotropy, it
274 cannot be fully ruled out, thus the results should be interpreted with caution. Third,
275 the puberty timing GWAS were conducted in females exclusively, while the
276 metabolites GWAS included female and male individuals. As there may be sex-
277 specific effects on the levels of metabolites³³, this could potentially impact our causal
278 effect estimates. Thus, future studies taking into account the sex-specific effects with
279 regard to the relationship between puberty timing and human plasma metabolism are
280 warranted.

281 In summary, we explored the causal relationship of puberty timing with human
282 plasma metabolite levels using publicly available GWAS summary data. We found a
283 causal effect of puberty timing on 23 plasma metabolites through two-sample MR
284 analysis, and most of the effects were mediated by adulthood adiposity. The findings
285 provide novel insights into the causal inferences of puberty timing on human
286 metabolism. Additionally, the role of adulthood adiposity in mediating the causal
287 relationship between puberty timing and plasma metabolites is identified.

288 **Material and Methods**

289 **Data sources**

290 The information on all GWAS data used in the analyses is shown in **Table S21**.
291 Summary statistics on age at menarche (AAM) (measured in years) were obtained
292 from the largest meta-analysis of GWASs, including 329,345 women of European
293 ancestry from 42 cohort studies³⁴. A total of 389 SNPs reaching genome-wide
294 significance ($P < 5 \times 10^{-8}$) were identified. There exists a genetic correlation ($r_g=0.74$)
295 between the AAM of females and the age at voice breaking of males, which suggests

296 that many genetic variants have comparable effects on puberty timing in both males
297 and females. Consequently, there was credibility that genetic variants of AAM can
298 provide confident insights into puberty timing in males as well ³⁵. Thus, in our study,
299 these genetic variants were considered instrumental variables for puberty timing. We
300 also refer to AAM as puberty timing in the present study.

301 A total of 174 plasma metabolites in seven biochemical categories (i.e., amino acids,
302 biogenic amines, acylcarnitines, phosphatidylcholines, lysophosphatidylcholines,
303 sphingomyelins, and hexoses), from the largest GWAS of human plasma metabolites,
304 were used as the outcomes in this study ³⁶. Summary statistics on all metabolites
305 measured by high-throughput platforms were obtained from the cross-platform meta-
306 analysis GWAS of blood metabolites, which included several cohorts with sample
307 sizes for each metabolite ranging from 8,569 to 86,507 individuals of European
308 ancestry. The GWASs were undertaken in cohorts of Fenland study (N = 9,363),
309 EPIC-Norfolk (European Prospective Investigation into Cancer-Norfolk) (N = 5,841),
310 and INTERVAL study (Metabolon Discovery HD4 platform, N = 8,455 and ¹H
311 nuclear magnetic resonance (NMR), Nightingale, N = 40,905). The results were
312 further meta-analyzed with two published studies conducted by Kettunen et al. (N =
313 24,925) and Shin et al. (N = 7,824) ^{37,38}.

314 The adulthood adiposity GWAS conducted by the Genetic Investigation of
315 Anthropomorphic Traits (GIANT) consortium included 315,347 adults of European
316 ancestry, combining data from the Research Program on Genes, Environment, and
317 Health (RPGEH), Genetic Epidemiology Research on Aging (GERA) cohort and a
318 meta-analysis of GWASs conducted by Locke et al. ^{39,40}. A total of 147 genome-wide

319 significant variants associated with adulthood BMI were selected as genetic
320 instruments in the MR mediation analysis.

321 Since birth weight and childhood adiposity have been reported to be associated with
322 puberty timing and plasma metabolite level changes in previous studies ⁴¹⁻⁴³, we
323 excluded SNPs, which were genetically significantly associated with these two traits,
324 from the genetic instruments for puberty timing to minimize the potential risk of bias
325 due to horizontal pleiotropy. Full summary data on birth weight and childhood BMI
326 were obtained from the Early Growth Genetics (EGG) Consortium ^{43,44}. In the birth
327 weight GWAS, Warrington et al. conducted a meta-analysis of the fetal genetic effects
328 in up to 297,356 Europeans with their own birth weight and 210,248 Europeans with
329 offspring birth weight ⁴³. Bradfield et al. performed a trans-ancestral GWAS meta-
330 analysis of childhood BMI from 30 studies, but only data on Europeans were used in
331 the analyses ⁴⁴.

332 We then attempted to replicate the findings of our study by using data from
333 independent GWASs. Data on puberty timing were obtained from a GWAS conducted
334 on up to 182,416 females of European descent from 58 studies ⁴⁵. Summary statistics
335 on adulthood BMI were obtained from a meta-analysis of 322,154 European
336 ancestries from the GIANT consortium ⁴⁰. Furthermore, we obtained summary
337 statistics from a GWAS of metabolites measured in the UKB, which were conducted
338 on over 114,000 individuals of European ancestry ⁴⁶.

339 **Genetic instruments selection**

340 To ensure statistical independence between genetic instruments, we performed a
341 stringent linkage disequilibrium (LD) clumping (window = 10 MB and $r^2 < 0.001$ in

342 PLINK 1.9) using the European ancestry data set of the 1000 Genomes Project as a
343 reference panel and selected all the bi-allelic SNPs with minor allele frequency
344 (MAF) > 0.01 using the clump data function in the TwoSample MR R package ⁴⁷.
345 After LD clumping, 208 and 147 SNPs were selected as genetic instruments for
346 puberty timing and adulthood BMI, respectively. Of note, when a SNP instrumenting
347 for the exposure of interest was not available in the outcome summary data, a proxy
348 variant in high-LD with the SNP (window = 1MB and $r^2 > 0.80$ in European ancestry)
349 was looked up using the online platform LDlink (<https://ldlink.nci.nih.gov/>). SNPs
350 would be removed from the sets of genetic instruments if no available proxies can be
351 identified. As no qualified alternative proxy variants were found in each step, the
352 analyses did not use any proxy SNPs. Detailed information on IVs for puberty timing
353 and adulthood BMI is presented in **Table S1-S6**.

354 **Statistical analysis**

355 **Two-sample MR analysis**

356 As shown in **Fig 1**, a two-sample MR analysis was used as the primary analysis
357 approach to examine the causal effects of genetically predicted puberty timing on 174
358 human plasma metabolites. The IVW method, which assumes no correlations or
359 interactions between genetic instruments and generates average causal effect estimates
360 across multiple instruments, was used in the primary MR analysis ⁴⁸. The causal effect
361 estimates were interpreted as standardized mean differences in concentrations of
362 metabolites per year delayed onset of puberty. Heterogeneity between causal effect
363 estimates for each SNP was examined by calculating Cochran's Q statistic, which
364 indicated potentially horizontal pleiotropy.

365 **MR mediation analysis**

366 To investigate the extent, to which adulthood adiposity might mediate the causal
367 relationship between puberty timing and metabolite levels, we performed an MR
368 mediation analysis (**Fig 1-2**). The indirect causal effects of puberty timing on plasma
369 metabolite levels were estimated using the two-step MR method ⁴⁹. In the first step,
370 we tested the causal effects of adulthood BMI on all plasma metabolite levels. In the
371 second step, we tested the causal effect of puberty timing on adulthood BMI. Then we
372 used the “product of coefficients” approach ⁵⁰ to calculate the indirect causal effects of
373 puberty timing on each blood metabolite via adulthood BMI. The standard error of
374 each indirect causal effect was estimated by using the “multivariate delta method” ⁵¹.

375 Compared with conventional univariable two-sample MR, which estimates the total
376 effect of one exposure on the outcome, MVMR includes multiple risk factors as
377 exposures to estimate the direct causal effect on the outcome after adjusting for other
378 risk factors ⁵². We accomplished MVMR analysis to assess the direct causal effects of
379 both puberty timing and adulthood BMI on 174 plasma metabolites. Since only
380 participants involved in the ReproGen consortium were available in the full summary
381 statistics of puberty timing GWAS ³⁴, a total of 207 IVs for puberty timing used in the
382 MVMR were directly selected from the summary data following the same LD
383 clumping strategy.

384 **Sensitivity analysis**

385 To assess the robustness of our results in the two-sample MR analysis, we also
386 performed four other sensitivity analyses. MR-Egger regression provides an unbiased
387 estimate of the causal effect even if all SNPs used as instrumental variables are invalid

388 ⁵³. The weighted median model requires that at least 50% of the SNPs were valid
389 genetic instruments to provide unbiased estimates of causal effects in MR analyses ⁵⁴.
390 The weighted mode method provides consistent causal effect estimates even if the
391 majority of SNPs were invalid ⁵⁵. The Mendelian Randomization Pleiotropy RESidual
392 Sum and Outlier (MR-PRESSO) method detects outlying genetic variants that might
393 lead to horizontal pleiotropy and further provides causal effect estimates after
394 removing outliers ⁵⁶. Horizontal pleiotropy was tested via the intercept terms in the
395 MR-Egger regression analysis.

396 As birth weight and childhood obesity were investigated as common causes of
397 puberty timing and some of the adulthood metabolite levels ⁵⁷, to minimize the risk of
398 violating one of the MR core assumptions (i.e., exclusion restriction assumption) ⁵⁸,
399 we excluded SNPs that were genetically significantly associated with birth weight or
400 childhood BMI ($P < 5 \times 10^{-8}$) from genetic instruments. Based on this, a secondary
401 IVW analysis was performed with the reduced genetic instrument sets.

402 **Replication analysis in the UK Biobank**

403 The UKB, a population-based cohort, has recruited approximately 500,000
404 individuals aged 40 to 69 years across the UK. By utilizing the targeted high-
405 throughput nuclear magnetic resonance (NMR) metabolomics platform (Nightingale
406 Health Ltd; biomarker quantification version 2020), 249 metabolic measures (i.e., 165
407 metabolic measures and 84 derived ratios) were quantified simultaneously ⁴⁶.
408 However, given that only nine of the metabolites (i.e., alanine, glutamine, glycine,
409 histidine, isoleucine, leucine, phenylalanine, tyrosine, and valine) measured in the
410 UKB were identified to coincide with the metabolites included in the primary analysis,
411 we merely estimated the relationship between puberty timing and the nine amino

412 acids to inspect the robustness of our primary analysis results. Firstly, we conducted a
413 two-sample MR analysis to investigate the total effect of puberty timing on each
414 amino acid. Subsequently, we also performed MVMR analysis to obtain the direct
415 impact of puberty timing on amino acids after adjusting for adulthood adiposity. Of 68
416 SNPs identified as associated with puberty timing, 60 were eventually used in two-
417 sample MR, and the 8 SNPs were removed due to being palindromic with
418 intermediate allele frequencies. A total of 55 SNPs were employed in MVMR analysis
419 after LD clumping.

420 **Multiple testing correction**

421 Given that a large number of tests were performed, a conservative Bonferroni
422 multiple-test correction was applied. The P value threshold for statistical significance
423 becomes $P < 4.90 \times 10^{-4}$ (0.05/102) after correcting for 102 tests corresponding to the
424 number of principal components explained 95% of the variance of the 174 metabolites
425 in the GWAS by Lotta et al.⁵⁹, which indicated strong causal evidence for our results.
426 Meanwhile, the conventional P value threshold of 0.05 was also reported to avoid
427 ignoring potentially crucial findings after constructing a “suggestive significance
428 threshold”^{60,61}, which offered moderate causal evidence.

429 All statistical analyses were performed using R version 4.1.2, in which the
430 TwoSampleMR (version 0.5.6) and MRPRESSO (version 1.0) packages were used
431 ^{47,56}. The ggplot2 and circlize R packages were used to visualize statistical analysis
432 results ^{62,63}.

433 **Data availability**

434 The genetic instruments for puberty timing were obtained from the corresponding
435 literature at https://reprogen.org/data_download.html. Data on plasma metabolites
436 GWAS can be found at www.omicscience.org. Full summary data on adulthood BMI
437 were obtained from the MRC Integrative Epidemiology Unit OpenGWAS project
438 (<https://gwas.mrcieu.ac.uk/>). Full summary data on birth weight and childhood BMI
439 were obtained from the Early Growth Genetics Consortium (egg-consortium.org). The
440 replication analyses were conducted using data from the UKB
441 (<https://www.ukbiobank.ac.uk/>) and the summary level data were obtained at
442 <https://gwas.mrcieu.ac.uk/>.

443 **Code availability**

444 Scripts are available on GitHub at https://github.com/ZengjunLi/Puberty_metabolites.

445 **References**

- 446 1. Toppari, J. & Juul, A. Trends in puberty timing in humans and environmental modifiers.
447 *Molecular and Cellular Endocrinology* **324**, 39-44 (2010).
- 448 2. Day, F.R., Perry, J.R.B. & Ong, K.K. Genetic Regulation of Puberty Timing in Humans.
449 *Neuroendocrinology* **102**, 247-255 (2015).
- 450 3. Pinyerd, B. & Zipf, W.B. Puberty—Timing Is Everything! *Journal of Pediatric Nursing* **20**,
451 75-82 (2005).
- 452 4. Li, W. *et al.* Association between Obesity and Puberty Timing: A Systematic Review and
453 Meta-Analysis. in *International Journal of Environmental Research and Public Health* Vol. 14
454 (2017).
- 455 5. Prentice, P. & Viner, R.M. Pubertal timing and adult obesity and cardiometabolic risk in
456 women and men: a systematic review and meta-analysis. *International Journal of Obesity* **37**,
457 1036-1043 (2013).
- 458 6. Cheng, T.S., Day, F.R., Lakshman, R. & Ong, K.K. Association of puberty timing with type 2

- 459 diabetes: a systematic review and meta-analysis. *PLoS medicine* **17**, e1003017 (2020).
- 460 7. Hulanicka, B., Lipowicz, A., Koziel, S. & Kowalisko, A. Relationship between early puberty
461 and the risk of hypertension/overweight at age 50: Evidence for a modified Barker hypothesis
462 among Polish youth. *Economics & Human Biology* **5**, 48-60 (2007).
- 463 8. Biro, F.M., Huang, B., Wasserman, H., Gordon, C.M. & Pinney, S.M. Pubertal Growth, IGF-1,
464 and Windows of Susceptibility: Puberty and Future Breast Cancer Risk. *Journal of Adolescent*
465 *Health* **68**, 517-522 (2021).
- 466 9. Galvao, T.F. *et al.* Pubertal timing in girls and depression: A systematic review. *Journal of*
467 *Affective Disorders* **155**, 13-19 (2014).
- 468 10. Park, J., Shin, Y., Kim, T.H., Kim, D.-H. & Lee, A. Plasma metabolites as possible biomarkers
469 for diagnosis of breast cancer. *PLoS One* **14**, e0225129 (2019).
- 470 11. Zhao, Y., Lv, H., Qiu, S., Gao, L. & Ai, H. Plasma metabolic profiling and novel metabolite
471 biomarkers for diagnosing prostate cancer. *RSC advances* **7**, 30060-30069 (2017).
- 472 12. Heys, M. *et al.* Age of Menarche and the Metabolic Syndrome in China. *Epidemiology*
473 **18**(2007).
- 474 13. Feng, Y. *et al.* Effects of age at menarche, reproductive years, and menopause on metabolic
475 risk factors for cardiovascular diseases. *Atherosclerosis* **196**, 590-7 (2008).
- 476 14. Davey Smith, G. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology
477 contribute to understanding environmental determinants of disease?*. *International Journal of*
478 *Epidemiology* **32**, 1-22 (2003).
- 479 15. Lawlor, D.A., Harbord, R.M., Sterne, J.A.C., Timpson, N. & Davey Smith, G. Mendelian
480 randomization: Using genes as instruments for making causal inferences in epidemiology.
481 *Statistics in Medicine* **27**, 1133-1163 (2008).
- 482 16. Bell, J. *et al.* Influence of puberty timing on adiposity and cardiometabolic traits: A Mendelian
483 randomisation study. *PLOS Medicine* **15**, e1002641 (2018).
- 484 17. Yuan, B. *et al.* A plasma metabolite panel as biomarkers for early primary breast cancer
485 detection. *International Journal of Cancer* **144**, 2833-2842 (2019).
- 486 18. Knific, T. *et al.* Models including plasma levels of sphingomyelins and phosphatidylcholines
487 as diagnostic and prognostic biomarkers of endometrial cancer. *The Journal of Steroid*
488 *Biochemistry and Molecular Biology* **178**, 312-321 (2018).

- 489 19. Li, S., Gao, D. & Jiang, Y. Function, Detection and Alteration of Acylcarnitine Metabolism in
490 Hepatocellular Carcinoma. *Metabolites* **9**(2019).
- 491 20. Takaya, H. *et al.* Acylcarnitine: Useful biomarker for early diagnosis of hepatocellular
492 carcinoma in non-steatohepatitis patients. *World J Gastrointest Oncol* **11**, 887-897 (2019).
- 493 21. Day, F.R., Elks, C.E., Murray, A., Ong, K.K. & Perry, J.R. Puberty timing associated with
494 diabetes, cardiovascular disease and also diverse health outcomes in men and women: the UK
495 Biobank study. *Sci Rep* **5**, 11208 (2015).
- 496 22. Kerner, J. & Hoppel, C. Fatty acid import into mitochondria. *Biochimica et Biophysica Acta*
497 (*BBA*) - *Molecular and Cell Biology of Lipids* **1486**, 1-17 (2000).
- 498 23. Violante, S. *et al.* Peroxisomes contribute to the acylcarnitine production when the carnitine
499 shuttle is deficient. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of*
500 *Lipids* **1831**, 1467-1474 (2013).
- 501 24. Meierhofer, D. Acylcarnitine profiling by low-resolution LC-MS. *PLoS One* **14**, e0221342
502 (2019).
- 503 25. Guasch-Ferré, M. *et al.* Plasma acylcarnitines and risk of cardiovascular disease: effect of
504 Mediterranean diet interventions. *Am J Clin Nutr* **103**, 1408-16 (2016).
- 505 26. Kaul, P. *et al.* Incidence of heart failure and mortality after acute coronary syndromes. *Am*
506 *Heart J* **165**, 379-85.e2 (2013).
- 507 27. Kalim, S. *et al.* A plasma long-chain acylcarnitine predicts cardiovascular mortality in incident
508 dialysis patients. *J Am Heart Assoc* **2**, e000542 (2013).
- 509 28. Nicholas, D.A. *et al.* Fatty Acid Metabolites Combine with Reduced β Oxidation to Activate
510 Th17 Inflammation in Human Type 2 Diabetes. *Cell Metab* **30**, 447-461.e5 (2019).
- 511 29. Jarrell, Z.R. *et al.* Plasma acylcarnitine levels increase with healthy aging. *Aging (Albany NY)*
512 **12**, 13555-13570 (2020).
- 513 30. Mishra, S.R., Chung, H.F., Waller, M. & Mishra, G.D. Duration of estrogen exposure during
514 reproductive years, age at menarche and age at menopause, and risk of cardiovascular disease
515 events, all-cause and cardiovascular mortality: a systematic review and meta-analysis. *BJOG:*
516 *An International Journal of Obstetrics & Gynaecology* **128**, 809-821 (2021).
- 517 31. Richardson, T.G., Sanderson, E., Elsworth, B., Tilling, K. & Davey Smith, G. Use of genetic
518 variation to separate the effects of early and later life adiposity on disease risk: mendelian

- 519 randomisation study. *Bmj* **369**, m1203 (2020).
- 520 32. Zhou, Y. *et al.* Obesity and diabetes related plasma amino acid alterations. *Clinical*
521 *Biochemistry* **46**, 1447-1452 (2013).
- 522 33. Verri Hernandez, V. *et al.* Age, Sex, Body Mass Index, Diet and Menopause Related
523 Metabolites in a Large Homogeneous Alpine Cohort. *Metabolites* **12**(2022).
- 524 34. Day, F.R. *et al.* Genomic analyses identify hundreds of variants associated with age at
525 menarche and support a role for puberty timing in cancer risk. *Nat Genet* **49**, 834-841 (2017).
- 526 35. Day, F.R. *et al.* Shared genetic aetiology of puberty timing between sexes and with health-
527 related outcomes. *Nature Communications* **6**, 8842 (2015).
- 528 36. Lotta, L.A. *et al.* A cross-platform approach identifies genetic regulators of human metabolism
529 and health. *Nat Genet* **53**, 54-64 (2021).
- 530 37. Kettunen, J. *et al.* Genome-wide study for circulating metabolites identifies 62 loci and reveals
531 novel systemic effects of LPA. *Nat Commun* **7**, 11122 (2016).
- 532 38. Shin, S.Y. *et al.* An atlas of genetic influences on human blood metabolites. *Nat Genet* **46**,
533 543-550 (2014).
- 534 39. Hoffmann, T.J. *et al.* A Large Multiethnic Genome-Wide Association Study of Adult Body
535 Mass Index Identifies Novel Loci. *Genetics* **210**, 499-515 (2018).
- 536 40. Locke, A.E. *et al.* Genetic studies of body mass index yield new insights for obesity biology.
537 *Nature* **518**, 197-206 (2015).
- 538 41. Wang, Y., Dinse, G.E. & Rogan, W.J. Birth weight, early weight gain and pubertal maturation:
539 a longitudinal study. *Pediatr Obes* **7**, 101-9 (2012).
- 540 42. Bull, C.J. *et al.* Adiposity, metabolites, and colorectal cancer risk: Mendelian randomization
541 study. *BMC Med* **18**, 396 (2020).
- 542 43. Warrington, N.M. *et al.* Maternal and fetal genetic effects on birth weight and their relevance
543 to cardio-metabolic risk factors. *Nat Genet* **51**, 804-814 (2019).
- 544 44. Bradfield, J.P. *et al.* A trans-ancestral meta-analysis of genome-wide association studies
545 reveals loci associated with childhood obesity. *Hum Mol Genet* **28**, 3327-3338 (2019).
- 546 45. Perry, J.R. *et al.* Parent-of-origin-specific allelic associations among 106 genomic loci for age
547 at menarche. *Nature* **514**, 92-97 (2014).
- 548 46. Borges, M.C. *et al.* Role of circulating polyunsaturated fatty acids on cardiovascular diseases

- 549 risk: analysis using Mendelian randomization and fatty acid genetic association data from over
550 114,000 UK Biobank participants. *BMC Med* **20**, 210 (2022).
- 551 47. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the
552 human phenome. *Elife* **7**(2018).
- 553 48. Burgess, S., Butterworth, A. & Thompson, S.G. Mendelian randomization analysis with
554 multiple genetic variants using summarized data. *Genet Epidemiol* **37**, 658-65 (2013).
- 555 49. Relton, C.L. & Davey Smith, G. Two-step epigenetic Mendelian randomization: a strategy for
556 establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol* **41**,
557 161-76 (2012).
- 558 50. VanderWeele, T.J. Mediation Analysis: A Practitioner's Guide. *Annu Rev Public Health* **37**, 17-
559 32 (2016).
- 560 51. MacKinnon, D.P., Fairchild, A.J. & Fritz, M.S. Mediation analysis. *Annu Rev Psychol* **58**, 593-
561 614 (2007).
- 562 52. Sanderson, E., Davey Smith, G., Windmeijer, F. & Bowden, J. An examination of
563 multivariable Mendelian randomization in the single-sample and two-sample summary data
564 settings. *Int J Epidemiol* **48**, 713-727 (2019).
- 565 53. Bowden, J. *et al.* Assessing the suitability of summary data for two-sample Mendelian
566 randomization analyses using MR-Egger regression: the role of the I² statistic. *Int J Epidemiol*
567 **45**, 1961-1974 (2016).
- 568 54. Bowden, J., Davey Smith, G., Haycock, P.C. & Burgess, S. Consistent Estimation in
569 Mendelian Randomization with Some Invalid Instruments Using a Weighted Median
570 Estimator. *Genet Epidemiol* **40**, 304-14 (2016).
- 571 55. Hartwig, F.P., Davey Smith, G. & Bowden, J. Robust inference in summary data Mendelian
572 randomization via the zero modal pleiotropy assumption. *International Journal of*
573 *Epidemiology* **46**, 1985-1998 (2017).
- 574 56. Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy
575 in causal relationships inferred from Mendelian randomization between complex traits and
576 diseases. *Nature Genetics* **50**, 693-698 (2018).
- 577 57. Bell, J.A. *et al.* Influence of puberty timing on adiposity and cardiometabolic traits: A
578 Mendelian randomisation study. *PLoS Med* **15**, e1002641 (2018).

- 579 58. Boef, A.G.C., Dekkers, O.M. & le Cessie, S. Mendelian randomization studies: a review of the
580 approaches used and the quality of reporting. *International Journal of Epidemiology* **44**, 496-
581 511 (2015).
- 582 59. Lotta, L.A. *et al.* A cross-platform approach identifies genetic regulators of human metabolism
583 and health. *Nat Genet* **53**, 54-64 (2021).
- 584 60. Sterne, J.A.C. & Smith, G.D. Sifting the evidence—what's wrong with significance tests?
585 *Physical Therapy* **81**, 1464-1469 (2001).
- 586 61. Amrhein, V., Greenland, S. & McShane, B. Scientists rise up against statistical significance.
587 (Nature Publishing Group, 2019).
- 588 62. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*, (ggplot2: Elegant Graphics for
589 Data Analysis, 2009).
- 590 63. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular
591 visualization in R. *Bioinformatics* **30**, 2811-2 (2014).

592

593 **Acknowledgements**

594 This study was supported by grants from the National Natural Science Foundation of
595 China (82373588, 82125032, 81930095 and 81761128035), Xinhua Hospital,
596 Shanghai Jiao Tong University School of Medicine, Shanghai, China (2021YJRC02),
597 Innovative research team of high-level local universities in Shanghai (SHSMU-
598 ZDCX20211100), the Science and Technology Commission of Shanghai Municipality
599 (19410713500 and 2018SHZDZX01), the Shanghai Municipal Commission of Health
600 and Family Planning (GWV-10.1-XK07, 2020CXJQ01, 2018YJRC03) and Sichuan
601 Science and Technology Program (2021JDR0189). The funders of the study had no
602 role in the study design, data collection, data analysis, or data interpretation; writing
603 or reviewing of the manuscript; and decision to submit the manuscript for publication.

604 **Author information**

605 **Authors and Affiliations**

606 1. Department of Epidemiology and Biostatistics, West China School of Public Health
607 and West China Fourth Hospital, Sichuan University, 610041, China.

608 Zengjun Li & Cairong Zhu*

609 2. Population Health Sciences, Bristol Medical School, University of Bristol, Bristol,
610 United Kingdom.

611 Si Fang

612 3. Medical Research Council (MRC) Integrative Epidemiology Unit (IEU), University
613 of Bristol, Bristol, United Kingdom.

614 Si Fang

615 4. Ministry of Education and Shanghai Key Laboratory of Children's Environmental
616 Health, Xinhua Hospital, Shanghai Jiao Tong University School of Medicine,
617 Shanghai, 200092, China

618 Dong Liu, Fei Li* & Jian Zhao*

619 5. Department of Developmental and Behavioral Pediatric & Child Primary Care,
620 Brain and Behavioral Research Unit of Shanghai Institute for Pediatric Research,
621 Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai,
622 China

623 Fei Li*

624 6. Department of Maternal and Child Health, School of Public Health, Shanghai Jiao
625 Tong University, 200092, China.

626 Fei Li* & Jian Zhao*

627 **Contributions**

628 Cairong Zhu, Fei Li, and Jian Zhao initially conceived the study design. Zengjun Li
629 and Si Fang carried out statistical analyses, and Zengjun Li wrote the first draft of the
630 manuscript with input from Cairong Zhu, Fei Li, and Jian Zhao. Zengjun Li and Dong
631 Liu plotted the figures. All authors contributed to the interpretation of the results and
632 critical revision of the manuscript.

633 **Corresponding author**

634 Correspondence to Jian Zhao, Cairong Zhu, and Fei Li.

635

636 **Ethical approval**

637 This study used publicly available GWAS summary data and was not subject to

638 institutional review board approval.

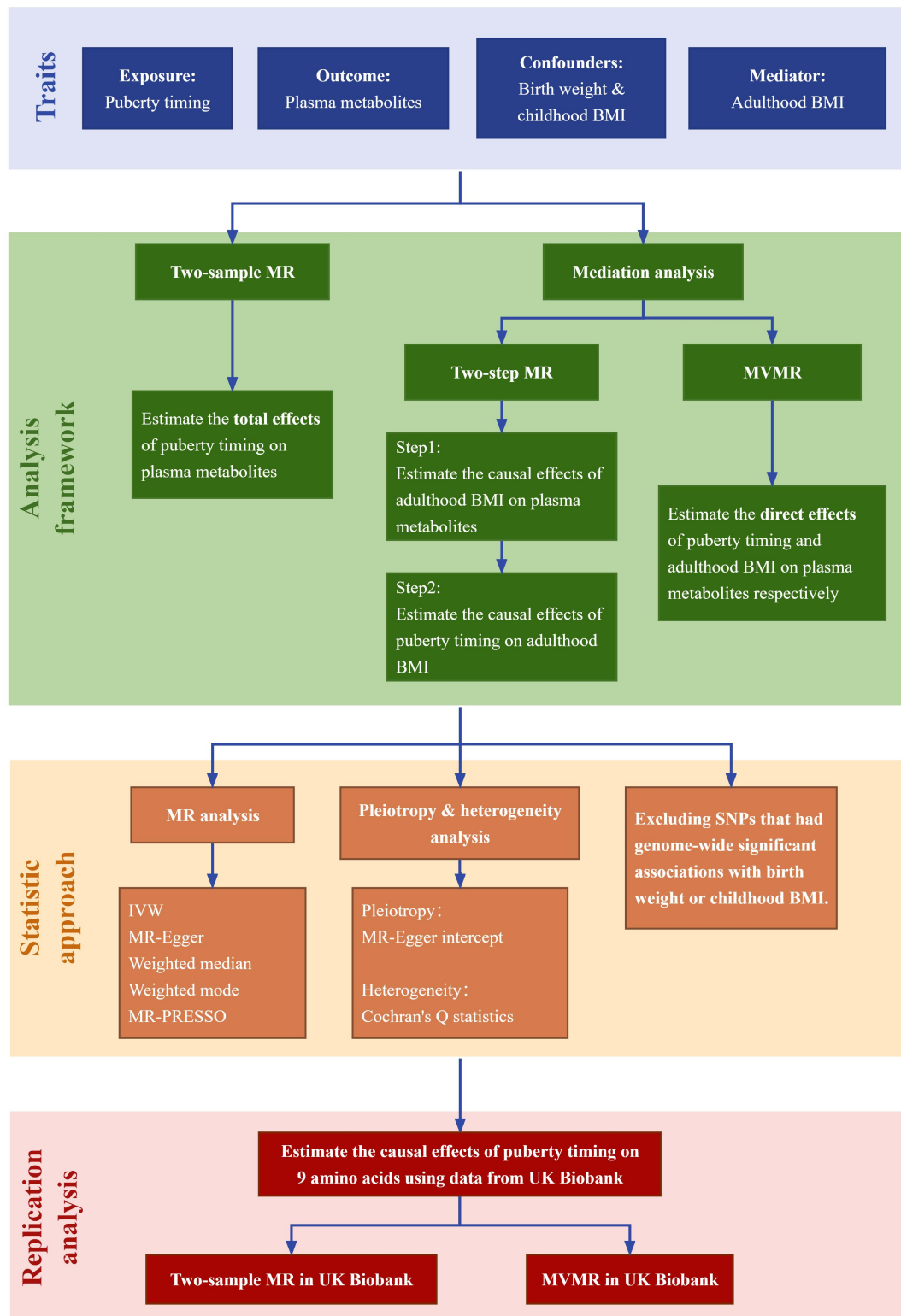
639 **Competing interests**

640 The authors declare no competing interests.

641 **Supplementary information**

642 Table S1-S21.

643

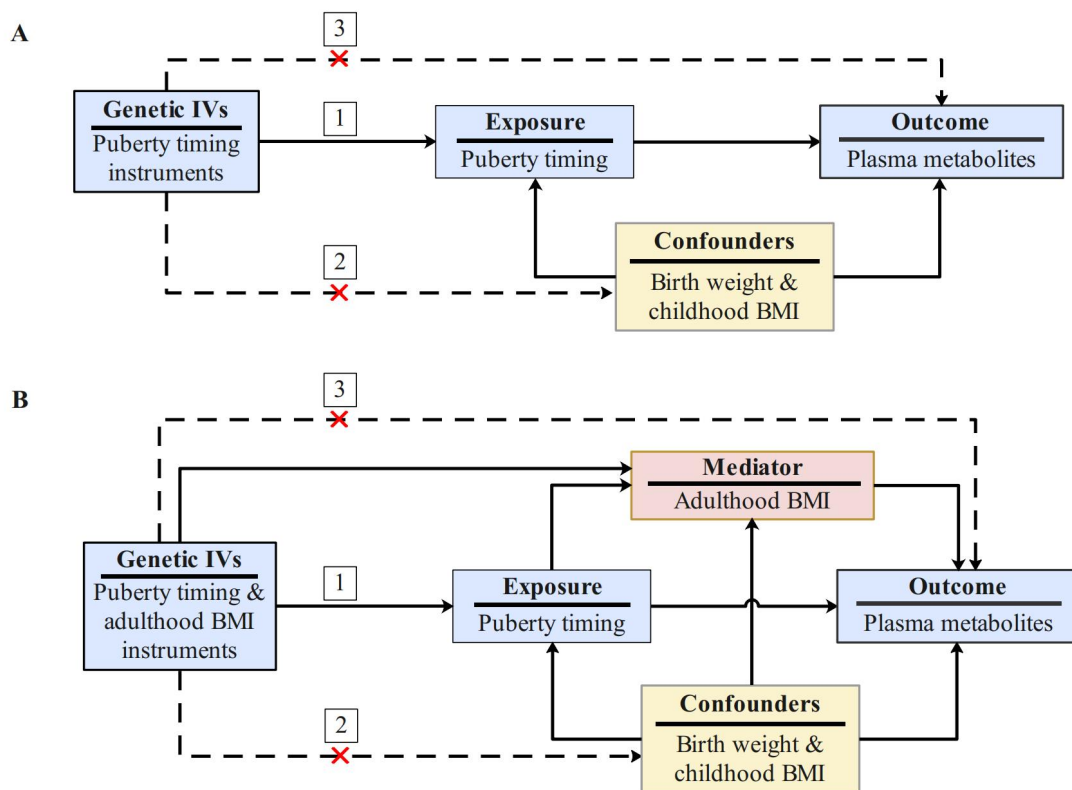


644

645 **Fig 1. The schematic diagram for Mendelian randomization of puberty timing on**
 646 **plasma metabolites. The primary analysis was conducted utilizing a two-sample MR**

647 (Mendelian randomization) approach to investigate the total effect of puberty timing
648 on plasma metabolites. In addition, we employed the two-step MR and MVMR
649 (multivariable Mendelian randomization) approach, taking the mediating effect of
650 adulthood BMI (body mass index) into account to infer the “direct” and “indirect”
651 effects. The IVW (inverse-variance weighted) approach was used as the primary
652 analysis, while MR-Egger, weighted median, weighted mode, and MR-PRESSO
653 (Mendelian randomization pleiotropy residual sum and outlier) were used as
654 sensitivity analyses. Subsequently, we excluded SNPs (single nucleotide
655 polymorphisms) that were genetically associated with birth weight and/or childhood
656 BMI from genetic instruments to minimize the risk of bias due to horizontal
657 pleiotropy. We further used metabolomic data obtained from the UK Biobank to
658 perform a replication analysis, aiming to validate the causal effect estimated.

659

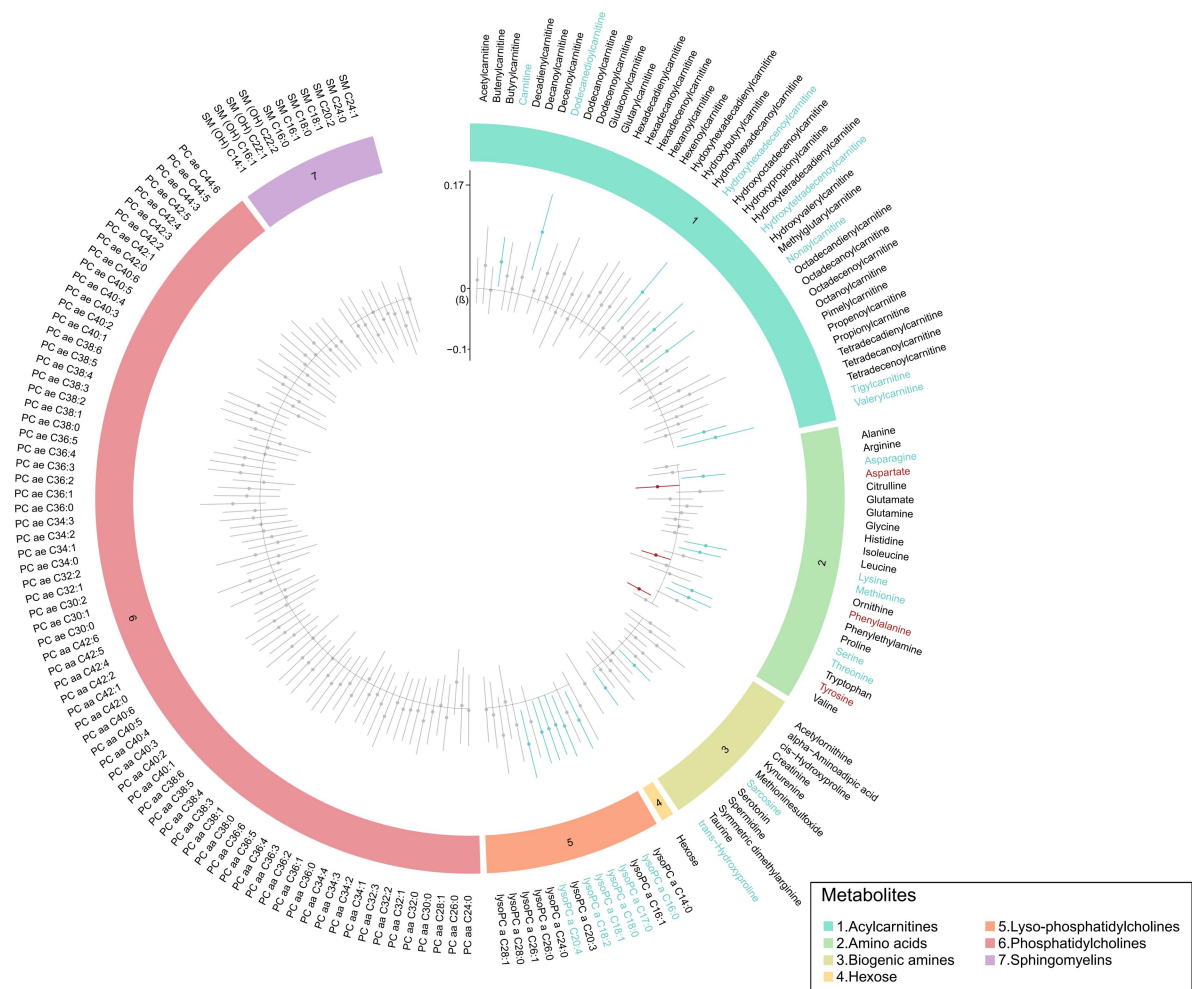


660

661 **Fig 2. The directed acyclic graphs of the Mendelian randomization study.** The figures
 662 depict directed acyclic graphs for univariable Mendelian randomization (MR) and mediation
 663 MR analysis in panel A and panel B respectively. The dashed lines represent potential
 664 horizontal pleiotropy or direct causal effects between variables, which would violate the MR
 665 assumption. MR approach is based on three assumptions: (1) the IVs (instrument variables)
 666 are strongly associated with the exposure (e.g., puberty timing); (2) the genetic IVs are not
 667 associated with any confounders (e.g., birth weight and childhood BMI (body mass index));
 668 (3) the genetic IVs should influence the outcome (e.g., a metabolite) through the exposure
 669 independently instead of through other pathways. The primary analysis was performed,
 670 wherein SNPs that exhibited genetic associations with puberty timing were selected as IVs.
 671 To minimize the risk of bias due to horizontal pleiotropy, a secondary analysis was conducted
 672 wherein SNPs that exhibited genetic associations with birth weight and/or childhood BMI

673 were excluded.

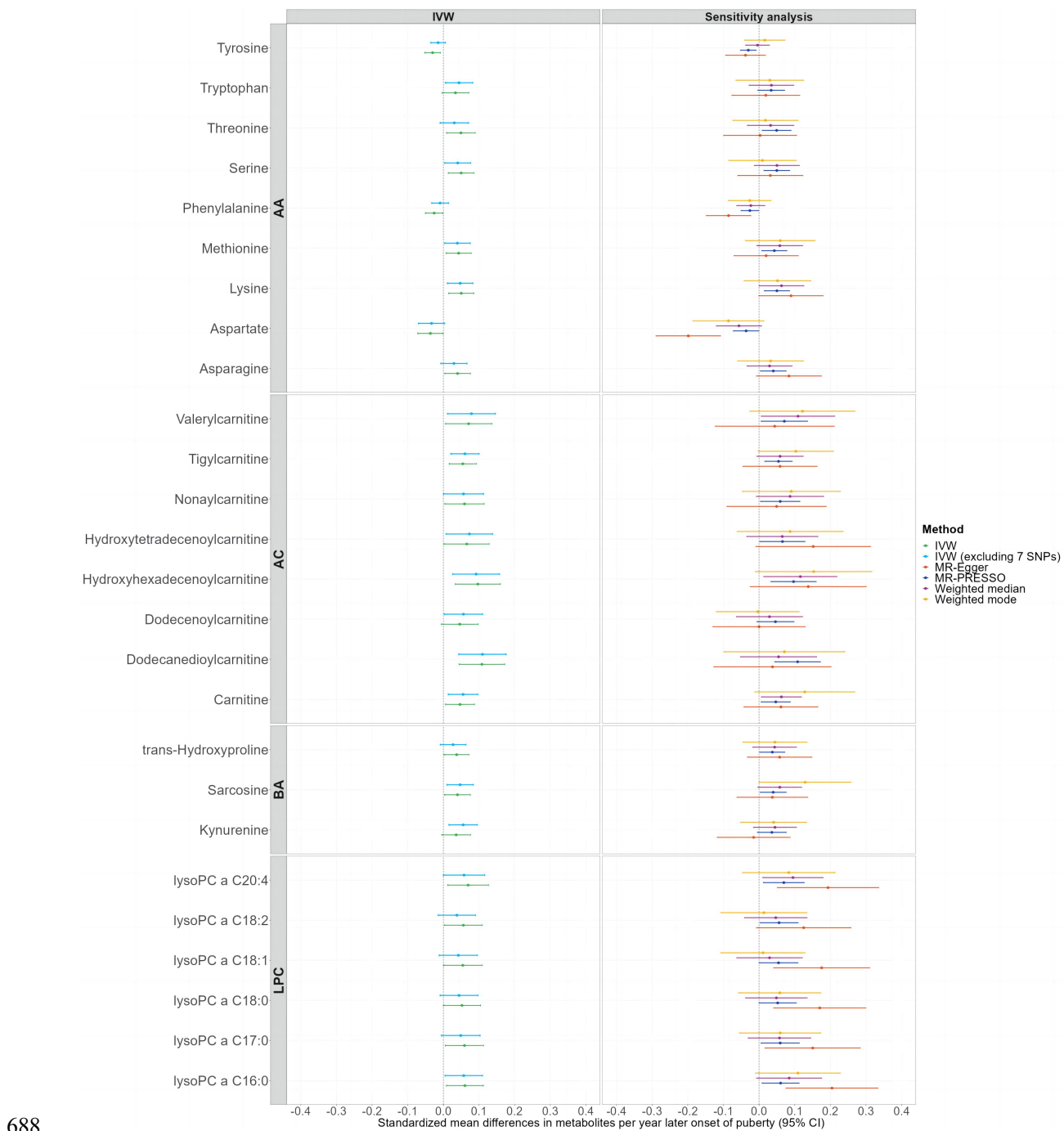
674



675

676 **Fig 3. Circos plot depicting the causal effect estimates of puberty timing on all 174**
 677 **plasma metabolites.** The legend located at the lower section of the figure provides a visual
 678 representation of distinct subclasses of plasma metabolites. The presence of metabolites
 679 highlighted with the color blue indicates a positive causal effect of puberty timing on the
 680 concentration of plasma metabolites. Specifically, each year later in puberty timing is
 681 associated with higher levels of metabolites. In contrast, the identification of metabolites
 682 highlighted with the color red suggests a negative causal effect of puberty timing on the
 683 concentration of plasma metabolites, that is each year later in puberty timing is associated
 684 with lower levels of metabolites. The presence of metabolites that had been designated as
 685 black indicates that there is insufficient evidence to establish a causal relationship. The

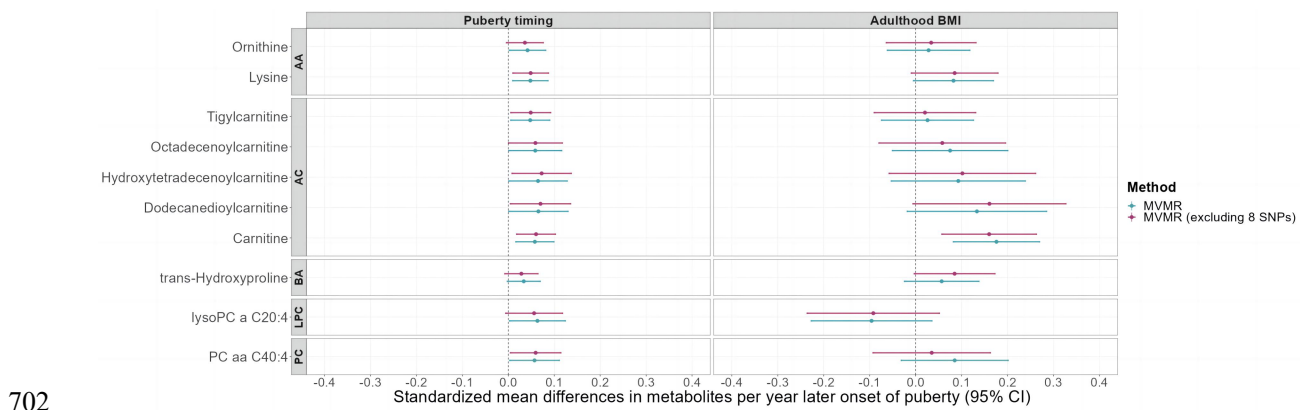
686 evidence indicates a causal effect of puberty timing on 23 plasma metabolites. Error bars
687 represent 95% confidence intervals. All statistical tests were two-sided.



688

689 **Fig 4. Forest plots showing the causal effect estimates of puberty timing on plasma**
 690 **metabolites.** The legend located on the right side of the forest plot indicates different analysis
 691 approaches. Plasma metabolites reaching nominal significance ($P < 0.05$) were shown
 692 according to the results of IVW (inverse-variance weighted). The results presented on the left

693 side of the plot were analyzed using the IVW approach. The colors represent green, using the
694 IVW method to estimate the effect size of puberty timing on plasma metabolites; blue; using
695 the IVW method, but excluding 7 genetic instruments genetically related to birth weight
696 and/or childhood BMI (body mass index). The results presented on the right were derived
697 from four distinct sensitivity analyses. The sensitivity analysis findings indicate that the
698 results of the primary analysis are largely robust. Error bars represent 95% confidence
699 intervals. All statistical tests were two-sided. The metabolites were divided according to
700 different subclasses. AA amino acids; AC acylcarnitines; BA biogenic amines; LPC
701 lysophosphatidylcholines.



702

703 **Fig 5. Forest plots comparing the effect estimates of puberty timing and adulthood BMI**

704 **on plasma metabolites in MVMR analyses.** The legend located on the right side of the

705 forest indicates the distinct analysis. Plasma metabolites reaching nominal significance ($P <$

706 0.05) in puberty timing were shown according to the results of the MVMR (multivariable

707 Mendelian randomization) approach. The colors represent blue, using the MVMR method to

708 estimate the effect size of puberty timing or adulthood BMI on plasma metabolites; purple;

709 using the MVMR method, but excluding 8 genetic instruments associated with birth weight

710 and/or childhood BMI (body mass index). The results on the left show the effect size of

711 puberty timing on each plasma metabolite, whereas the results on the right present the effect

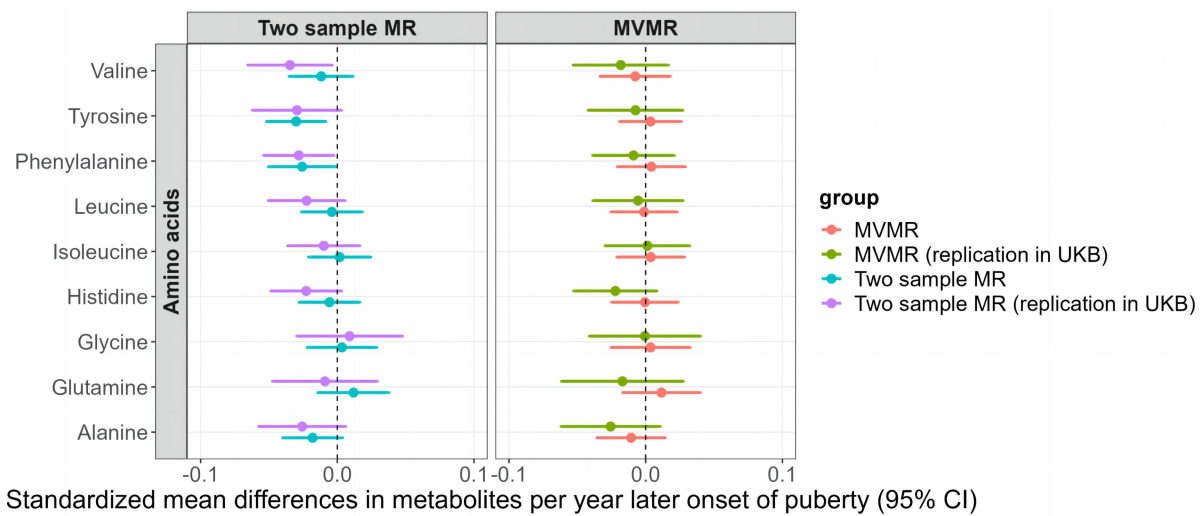
712 size of adulthood BMI on each plasma metabolite. Error bars represent 95% confidence

713 intervals. All statistical tests were two-sided. The metabolites were also divided according to

714 different subclasses. AA amino acids; AC acylcarnitines; BA biogenic amines; LPC

715 lysophosphatidylcholines; PC phosphatidylcholines.

716



717

718 **Fig 6. Forest plots comparing the effect estimates of puberty timing on nine amino acids**

719 **in replication analyses.** The legend located on the right side of the forest plot indicates

720 different analysis methods. The results of two-sample MR (Mendelian randomization) and

721 MVMR (multivariable Mendelian randomization) represented the total (blue) and direct

722 impact (orange) of puberty timing on plasma metabolites, respectively. To inspect the

723 robustness of our two-sample MR and MVMR analysis results (purple and green), we

724 estimated the causal relationship of puberty timing on nine identified amino acids using data

725 from the UKB (UK Biobank). Our findings in the primary MR analysis were further

726 supported by the findings from external replication analysis in the UKB. Error bars represent

727 95% confidence intervals. All statistical tests were two-sided.