Early View

Original research article

Does DNA methylation mediate the association of age at puberty with FVC or FEV₁?

Liang Li, Hongmei Zhang, John W. Holloway, Susan Ewart, Caroline L. Relton, S. Hasan Arshad, Wilfried Karmaus

Please cite this article as: Li L, Zhang H, Holloway JW, *et al.* Does DNA methylation mediate the association of age at puberty with FVC or FEV₁?. *ERJ Open Res* 2022; in press (<https://doi.org/10.1183/23120541.00476-2021>).

This manuscript has recently been accepted for publication in the *ERJ Open Research*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJOR online.

Copyright ©The authors 2022. This version is distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0. For commercial reproduction rights and permissions contact permissions@ersnet.org

Does DNA methylation mediate the association of age at puberty with FVC or FEV₁?

Liang Li¹, Hongmei Zhang^{1*}, John W. Holloway², Susan Ewart³, Caroline L. Relton⁴, S. Hasan Arshad^{5,6}, Wilfried Karmaus¹

Affiliations: ¹ Division of Epidemiology, Biostatistics, and Environmental Health, School of Public Health, University of Memphis, Memphis, TN, USA. ² Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, SO16 6YD, UK. ³ College of Veterinary Medicine, Michigan State University, East Lansing, MI, USA. ⁴ MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN, UK. ⁵ Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, SO16 6YD, UK. ⁶ The David Hide Asthma and Allergy Research Centre, St Mary's, Hospital, Parkhurst Road, Newport, Isle of Wight PO30 5TG, UK.

***Correspondence:** Hongmei Zhang, Division of Epidemiology, Biostatistics, and Environmental Health, School of Public Health, University of Memphis, Memphis, TN, USA. E-mail:

hzhang6@memphis.edu

Take home message: The association between age at pubertal onset and lung function parameters FVC or FEV₁ in young adulthood was likely not to be mediated by DNA methylation in peripheral blood.

ABSTRACT

Background: Age of pubertal onset is associated with lung function in adulthood. However, the underlying role of epigenetics as a mediator of this association remains unknown.

Methods: DNA methylation (DNAm) in peripheral blood was measured at age 18 years in the Isle of Wight birth cohort (IOWBC) along with data on age of pubertal events, forced vital capacity (FVC), and forced expiratory volume in 1s (FEV₁) at 26 years. Structural equation models were applied to examine mediation effects of DNAm on the association of age at pubertal events with FVC and FEV₁.

Findings were further tested in the Avon Longitudinal Study of Parents and Children cohort (ALSPAC).

Results: In the IOWBC, for females, 21 cytosine-phosphate-guanine sites (CpGs) were shown to mediate the association of age at puberty with FVC or FEV₁ at 26 years ($p < 0.05$). In males, DNAm at 20 CpGs was found to mediate the association of age at puberty with FVC ($p < 0.05$). At almost all these CpGs, indirect effects (effects of age at pubertal events on FVC or FEV₁ via DNAm) contributed a smaller portion to the total effects compared to direct effects (e.g., at cg08680129, ~22% of the estimated total effect of age at menarche on FVC at age 26 was contributed by an indirect effect). Among the IOWBC-discovered CpGs available in ALSPAC, none of them were replicated in ALSPAC ($p > 0.05$).

Conclusions: Our findings suggest that post-adolescence DNAm in peripheral blood is likely not to mediate the association of age at pubertal onset with young adulthood FVC or FEV₁.

Introduction

Lung function parameters, including forced vital capacity (FVC) and forced expiratory volume in one second (FEV₁), are shown to be associated with the age of menarche [1-4]. Subjects with early menarche have lower FVC and FEV₁ in adulthood [1, 2, 4, 5]. A recent Mendelian randomization study suggests that pubertal timing in general rather than the time of a single pubertal event is associated with lung function [2], implying the necessity to examine the contribution of other pubertal events onset on lung function development in both males and females. Our recent study has shown that age at puberty was associated with FVC in young adulthood in both sexes [6].

A covalent addition of a methyl group to the DNA at a cytosine residue followed by a guanine (CpG site), i.e., DNA methylation (DNAm) at CpG sites, is an important epigenetic mechanism that regulates gene transcription. Multiple studies have demonstrated that DNAm, as an epigenetic mechanism, links aging, environment exposure, and lifestyle factors to disease and health status [7-10]. Recent studies examined such impact of DNAm via mediation analyses, which offers an

opportunity to better understand the role that DNA methylation plays between exposures and health outcomes [11-14]. Studies have shown that DNAm is associated with various traits including age of pubertal onset [15-19]. For females, one study suggested that DNA methylation level might be modulated by age at menarche with later puberty associated with lower DNAm [18]. An association between DNA methylation patterns and lung function measures has also been demonstrated [20, 21].

Given the association between age at pubertal events and lung function, and the association of DNAm with both age of pubertal events and with lung function, this study aimed to ascertain whether DNAm at post-adolescence or close to post-adolescence mediates the association between age of pubertal events and young adulthood lung function parameters FVC and FEV₁ (Figure 1). To test this, we applied mediation analyses using structural equation modeling [22], using data in the Isle of Wight birth cohort (IOWBC; discovery cohort) [23] and the Avon Longitudinal Study of Parents and Children birth cohort (ALSPAC; replication cohort) [24-26]. Mediation analysis is a statistical model with a minimum of three variables, in which there is one hypothesized mediator variable (M) transmitting the effect of independent (X) to the dependent variable (Y) [22]. In our study, mediation analysis allows evaluation of indirect and direct effects of an exposure (e.g., age at puberty) on an outcome of interest (e.g., a lung function parameter) (Figure 1), where the expected value of the mediator (e.g., DNAm) and the outcome can be expressed as a function of the exposure via a generalized linear model [27].

Methods

Study population in IOWBC

A population-based birth cohort study was initiated in 1989 on the Isle of Wight (IOWBC), UK, to prospectively study the natural history of allergic diseases and associated risk factors [23]. After exclusion of adoptions, perinatal deaths, and refusal to follow up, 1,456 newborns were enrolled in the study. In total, 1,261 children were followed up at 10, 18, or 26 years (Figure S1). The study was approved by the Isle of Wight Local Research Ethics Committee (06/Q1701/34). The internal review board of the University of Memphis approved the project (FWA00006815). At each follow-up, demographic information was collected, allergic status was assessed, and height, weight, and FVC and FEV₁ were measured.

Age at onset of pubertal events

The age at onset of pubertal events was obtained through questionnaires when subjects were interviewed at age 18 years. Questions relating to pubertal events came from the National Institute of Child and Human Development (NICHD) questionnaire from the Study of Early Child Care and Youth Development based on the Pubertal Development Scale (PDS) [28]. For females, pubertal events included body hair growth, breast growth, menarche, skin changes, and growth spurt, and for males, pubertal events body hair growth, facial hair growth, voice deepening, skin changes, and growth spurt were the focus. The detailed measurements of pubertal events are described elsewhere [29].

Lung function assessment and potential confounders

Assessment of lung function at 26 years of age was conducted using the KoKo spirometry software package on portable desktop device (PDS instrumentation, Louisville, KY, USS) according to the guidelines of the American Thoracic Society and the European Respiratory Society [30]. The lung

function measurements included in this study were forced vital capacity (FVC) and forced expiratory volume in one second (FEV₁).

Low birth weight (<2,500 grams), maternal smoking during pregnancy, asthma status at age 10 years, body mass index (BMI) at age 10 years, socioeconomic status (SES), atopy at age 10 years, and personal smoking at age 18 years were included in the analyses as potential confounders (Supplemental Material section S.2.).

DNA methylation (DNAm) and cell type composition

In the IOWBC, DNA was extracted from whole blood using a standard procedure, and DNAm was measured using Infinium HumanMethylation450 or MethylationEPIC BeadChips. Details of DNAm assessment and preprocessing are described in Supplemental Material section S.2. CpGs on autosome chromosomes and M values, calculated as base 2 logit transformed DNAm β values at each CpG, were used in the analyses. In this study, DNAm at age 18 years was included in the analyses.

Heterogeneity in white blood cell composition may confound DNA methylation measurements [31]. Proportions of six cell types, CD4+ T cells, natural killer cells, neutrophil, B cells, monocytes, and eosinophils, were inferred using the R package *minfi* [32, 33] and included in the analyses as confounders.

Statistical analyses

Given the differences in age of pubertal onset and in lung function growth between males and females [6, 34], all analyses were stratified by sex. To assess whether the sub-sample in the study was representative of the IOWBC, one sample t-tests were applied to the key variables, including ages of pubertal events onset, and lung function measures FVC and FEV₁ at age 26 years.

The study aimed to examine whether DNAm at age 18 years mediates the association between age of puberty and age 26-year FVC or FEV₁ in young adults (Figure 1). A set of analyses was carried out to achieve this goal (Figure S2). Firstly, the association of FVC and FEV₁ with age of

pubertal event onset was assessed using linear regressions. Only pubertal events that passed this screening were included in subsequent analyses. Next, we screened CpG sites from two directions in an effort to include as many CpGs as possible in the assessment of their mediation effects while avoiding significant power loss should we analyze at the genome-scale. One direction is to identify candidate CpGs potentially associated with age at onset of a pubertal event (forward screening) and the other direction is to screen CpGs associated with the two lung function parameters (backward screening). Screening was carried out using the R package *ttScreening* [35], and CpGs showing association with age at puberty or lung function at age 26 were included in subsequent path analyses. To assess the potential of mediation effects of DNAm, PROC CALIS (SAS 9.4) was applied to evaluate direct and indirect effects (via DNAm) of ages at pubertal events on FVC or FEV₁ (Figure 2); statistical significance level was set at 0.05.

Replication study in ALSPAC

Pubertal events showing statistically significant indirect effects through DNAm on FVC or FEV₁ were further tested in the ALSPAC birth cohort [24-26] (Supplemental Material section S.2.). Complete data on demographics, ages of puberty onset (menarche and body hair growth for females, and body hair growth and voice deepening for males), DNAm at ages 15 or 17 years, along with FVC and FEV₁ at ages 24 years were analyzed (n=353 females and 335 males) [36]. The same path analysis modeling with comparable covariates was applied. Statistical significance was set at p-value <0.05.

Results

Population characteristics

The IOWBC subsamples reasonably represented the complete cohort, indicated by statistically insignificant differences between the subsamples and the complete cohort on the variables of interest (all p-values $\geq$ 0.05, Table 1).

Initial screening for pubertal events

Based on linear regressions, ages at menarche and age at body hair growth in females were associated with both FVC and FEV₁, and for males, age of body hair growth, growth spurt, voice deepening, and facial hair growth were shown to be associated with FVC only (Table S1). Age of onset for these pubertal events was included in the subsequent mediation analyses.

Mediation analysis through DNAm at CpGs passed forward screening

In total, 321 unique CpGs for females and 489 unique CpGs for males survived forward screening and were included in path analyses (Supplemental Material section S.3. and Figure 3). Results from path analyses in IOWBC indicated that, for females, at 6 and 7 CpGs, DNAm mediated the association of age at menarche with FVC and FEV₁ at age 26 years, respectively (p-value<0.05; Tables S2A). For FEV₁ at age 26 years, 4 CpGs showed a mediation effect on the association of age at body hair growth with FEV₁ (p-value<0.05; Table S2A), but we did not identify any mediating CpGs for FVC at age 26 years. In total, 14 unique CpGs were identified. At these CpGs, of the total effects (direct + indirect effects), direct effects accounted for a large fraction of the total effect e.g., at cg08680129, ~ 78% of the estimated total effect of age at menarche on FVC at age 26 was contributed by direct effect, and an indirect effect through DNAm only contributed a small portion (Table S2A, Figure S3). The small contributions from indirect effects were also reflected by the small ratios of indirect effect over direct effect, e.g., at cg08680128, the ratio is 0.04/0.14=2/7. It is worth noting that using ratios of indirect effect over direct effect is especially informative when the two effects were in opposite directions (e.g., a negative direct effect but positive indirect effect), in which case the ratio was calculated using absolute values of estimated coefficients.

In males, through path analyses, we found that DNAm at 7, 3, 2, and 4 candidate CpGs mediated the association of ages of body hair growth, growth spurt, voice deepening, and facial hair growth, respectively, with FVC at age 26 years (p-values<0.05) (Table S2A). In total, 16 unique CpGs demonstrated statistically significant mediation effects. At 10 of the 16 identified CpGs, ratios

of indirect effect over direct effect were less than 0.5, indicating an overall small contribution of DNAm as a mediator on the association of age of pubertal events with FVC.

Mediation analysis through DNAm at CpGs passed backward screening

Path analyses were applied to 321 unique CpGs that passed backward screening (259 in females; Supplemental Material section S.3. and Figure 3). In IOWBC, for females, DNAm at three of the 259 candidate CpGs was shown to mediate the association between age at menarche and FEV₁ at age 26 years (p-value<0.05) (Table S2B). Mediation effects were not identified at any of the candidate CpGs for FVC at age 26 years. For age at body hair growth, DNAm at 3 and 4 CpGs mediated its association with FVC and FEV₁, respectively (p-value<0.05) (Table S2B). Altogether, 7 unique CpGs showed statistically significant mediation effects, but in most mediation analyses, magnitudes of indirect effects were lower than a half of that for direct effect (in 8 of the 10 mediation analyses; Table S2B).

In males, for ages of body hair growth, growth spurt, voice deepening, and facial hair growth, mediation effects of DNAm on the association with FVC at age 26 years was found at one CpG site of the 62 candidate CpGs for each of the four pubertal events (p-values<0.05) (Table S2B). At all these four CpGs, contributions of indirect effects to total effects were small, ranging from 25% to 38% (Table S2B) and the ratios of indirect over direct effects were all <0.6.

Comparison between forward screening and backward screening

In females, no common CpGs was observed between the 321 CpGs that survived forward screening and the 259 CpGs that survived backward screening. Similarly, for males, no common CpGs were identified between 489 CpGs that passed forward screening and the 62 CpGs that passed backward screening. When comparing common genes to which these CpGs were mapped between forward screening and backward screening, 7 and 1 common genes were observed in females and males, respectively.

Further assessment of the IOWBC discovered CpGs in ALSPAC

In females, of the 21 discovered CpGs, indirect effects at 10 CpGs (Table 2A and 2B) were consistent with those observed in IOWBC, but none of them were statistically significant. In males, ages of body hair growth and voice deepening were available in ALSPAC. Of the 11 IOWBC-discovered CpGs for these two pubertal events, consistent mediation effects were observed at 6 CpGs but were statistically insignificant (Table 2A and 2B). Overall, indirect effects of these two pubertal events were negligible in ALSPAC, consistent with the findings in the IOWBC. Detailed results were in the supplementary material S.4.

Discussion

Combining findings in both the discovery and replication cohorts, post-adolescence or close-to-post-adolescence DNA methylation showed negligible mediation effects on the association of age at pubertal events with young adulthood FVC or FEV₁. Specifically, indirect effects only represented a small proportion of total effects of age at pubertal events on the two lung function measurements in young adults. Overall, the findings in both cohorts did not support a role for DNAm in mediating the association of pubertal events with FVC or FEV₁.

Our findings were drawn based on a rigorous study design with the potential to maximize statistical power. We used a comprehensive screening process to screen pubertal events and CpG sites to maximize statistical power. The two ways of screening for CpG sites, age at puberty → DNAm (forward), DNAm → lung function (backward), allowed us to include as many potentially informative CpG sites (with respect to potential mediation effects) as possible. These carefully designed screening procedures provided us an opportunity to identify mediation CpGs, if any, with high power. Even with this carefully designed screening process, our data in both cohorts did not identify CpGs showing strong potential as mediators. In addition, the findings between the discovery and replication cohort were consistent and not affected by cell type compositions. Cell-type could introduce multicollinearity

in path analyses [37], which could lead to an apparent reversal of the direction of indirect effects and changes in the size of indirect effects. In IOWBC, the direction of associations with or without cell type compositions was intact for the indirect effects. For the indirect effects estimated in the ALSPAC cohort, although the direction of associations changed at two CpGs when removing the cell types in the models, since the size of indirect effects was very trivial regardless of the involvement of cell types, we expect the change in direction of association was due to random error.

Although our data did not support the role of DNAm as a mediator between age at puberty and young adult lung function parameters, we cannot rule out the existence of CpGs (but we have not discovered) acting as mediators. One explanation of the lack of association might be that current methods for arrays (Illumina 450K and EPIC arrays) cover only 2-4% CpG sites across the whole genome [38, 39]. In addition, we used DNAm at CpGs from peripheral blood. DNAm shows strong tissue-specific effects and DNAm in another tissue (e.g., lung) might mediate the association of pubertal event timing with lung function [40]. Our mediation analyses assumed a linear relationship between exposure and outcome variables. However, biological relationships can be complex and non-linear with unknown forms. More advanced analytical approaches in this area are certainly needed.

Our study has multiple strengths. The first is the utilization of discovery and replication cohorts and the results from these two independent cohorts suggest that our identified insignificant mediation effects of DNAm were potentially conclusive. Second, to avoid reverse causation, the mediation analyses were built upon a longitudinal setting, offering an insight to future investigations on possible “causal paths” and helping open the “black box” from age of pubertal events onset to lung function later in life [41]. Third, the carefully designed screening process in the detection of candidate pubertal events and candidate CpGs offered a strong potential to improve statistical power and at the same time reduced the risk of excluding potentially important epigenetic mediators. Via currently available epigenetic assessment platforms of Illumina, results from our carefully designed study support the absence of epigenetic mediation on the association of age at puberty with lung function measurements FVC and FEV₁, based on DNAm measured in whole blood. Further assessment is certainly deserved, accompanied by a much denser assessment of DNAm such as pyrosequencing, and

by novel analytical approaches such as analyzing a set of CpGs in one region as a unified entity. However, genome-scale pyrosequencing with all CpG sites covered is challenging. Novel analytical tools are greatly needed and will allow an in-depth assessment of the role of DNAm in the association of age at puberty with lung function measurements. Furthermore, instead of DNAm in peripheral blood, using DNAm measured in lung tissues or tissues in the respiratory system will enable us to assess whether our insignificant findings were due to tissue-specificity and in the meantime to better reveal underlying biological mechanisms from pubertal events to lung function in young adulthood. Also, in the screening process based on data in the IOWBC, some pubertal events were not associated with FEV₁ or FVC. The lack of association between age at those pubertal events and lung function parameters was likely due to the variations caused by recall bias. In IOWBC, ages of onset for each pubertal event were obtained by retrospective assessment at age 18 through questionnaire. The timing order of different pubertal ages in IOWBC had been evaluated in our earlier study and we showed that the timing order was consistent with other studies [29]. However, compared to pubertal events such as menarche that have less recall bias on age of onset [42], ages at other pubertal events, e.g., skin changes, were likely to have more recall bias, potentially leading to insignificant associations with FEV₁ or FVC.

Conclusion

Our study suggests that DNA methylation in peripheral blood in post- or late-adolescences might not play a role in mediating the association between age of pubertal events and lung function in young adults.

Acknowledgments: We are extremely grateful to all participants and their families who are followed up in the Isle of Wight 1989 birth cohort and the Avon Longitudinal Study of Parents and Children cohort, which include interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, midwives, and nurses. We also appreciate the hard

efforts of these two research teams in maintaining the cohorts and collecting data. Our special thanks also go to the High Performance Computing facility provided by the University of Memphis.

Author contributions: The original concept and design were initiated by Hongmei Zhang and Wilfried Karmaus. Data analyses were performed by: Liang Li and Hongmei Zhang. Data interpretation and manuscript writing was led by Liang Li and Hongmei Zhang. Critical comments on the manuscript and final approval of the manuscript were given by all the authors: Liang Li, Hongmei Zhang, John W. Holloway, Susan Ewart, Caroline L. Relton, S. Hasan Arshad, Wilfried Karmaus.

Conflict of interest: All authors are nothing to disclose.

Support statement: This work was supported by the National Institutes of Health research funds R01AI121226 (MPI: H Zhang and JW Holloway) and R01 AI091905 (PI: W Karmaus). The 10-year follow-up of IOWBC was funded by National Asthma Campaign, UK (Grant No 364) and the 18-year follow-up by R01HL082925 (PI: SH Arshad). The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding is available on the ALSPAC website (<http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>). The sponsor had no role in study design, data collection, and analysis, or the preparation of the manuscript. This publication is the work of the authors and Hongmei Zhang will serve as guarantors for the contents of this paper.

References

- 1 Macsali F, Real FG, Plana E, et al. Early age at menarche, lung function, and adult asthma. *Am J Respir Crit Care Med* 2011; 183: 8-14.
- 2 Gill D, Sheehan NA, Wielscher M, et al. Age at menarche and lung function: a Mendelian randomization study. *European journal of epidemiology* 2017; 32: 701-710.
- 3 Murri V, Antoniazzi F, Piazza M, et al. Lung Function in Women with Idiopathic Central Precocious Puberty: A Pilot Study. *Horm Res Paediatr* 2017; 87: 95-102.
- 4 Mahmoud O, Granell R, Tilling K, et al. Association of Height Growth in Puberty with Lung Function: A Longitudinal Study. *Am J Respir Crit Care Med* 2018; 198: 1539-1548.
- 5 Campbell B, Simpson JA, Bui DS, et al. Early menarche is associated with lower adult lung function: A longitudinal cohort study from the first to sixth decade of life. *Respirology (Carlton, Vic)* 2020; 25: 289-297.
- 6 Li L, Zhang H, Holloway JW, et al. Pubertal onset with adulthood lung function mediated by height growth in adolescence. *ERJ Open Res* 2020: 00535-02020.
- 7 Li S, Hursting SD, Davis BJ, et al. Environmental exposure, DNA methylation, and gene regulation: lessons from diethylstilbestrol-induced cancers. *Annals of the New York Academy of Sciences* 2003; 983: 161-169.
- 8 Zhao W, Ammous F, Ratliff S, et al. Education and Lifestyle Factors Are Associated with DNA Methylation Clocks in Older African Americans. *International journal of environmental research and public health* 2019; 16.
- 9 Needham BL, Smith JA, Zhao W, et al. Life course socioeconomic status and DNA methylation in genes related to stress reactivity and inflammation: The multi-ethnic study of atherosclerosis. *Epigenetics* 2015; 10: 958-969.
- 10 Smith JA, Zhao W, Wang X, et al. Neighborhood characteristics influence DNA methylation of genes involved in stress response and inflammation: The Multi-Ethnic Study of Atherosclerosis. *Epigenetics* 2017; 12: 662-673.
- 11 Tobi EW, Slieker RC, Luijk R, et al. DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood. *Sci Adv* 2018; 4: eaao4364.
- 12 Fujii R, Sato S, Tsuboi Y, et al. DNA methylation as a mediator of associations between the environment and chronic diseases: A scoping review on application of mediation analysis. *Epigenetics* 2021: 1-27.
- 13 McDade TW, Ryan C, Jones MJ, et al. Social and physical environments early in development predict DNA methylation of inflammatory genes in young adulthood. *Proceedings of the National Academy of Sciences of the United States of America* 2017; 114: 7611-7616.
- 14 Barker ED, Walton E, Cecil CAM. Annual Research Review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology. *Journal of child psychology and psychiatry, and allied disciplines* 2018; 59: 303-322.
- 15 Lomniczi A, Loche A, Castellano JM, et al. Epigenetic control of female puberty. *Nature neuroscience* 2013; 16: 281-289.
- 16 Luo L, Yao Z, Ye J, et al. Identification of differential genomic DNA Methylation in the hypothalamus of pubertal rat using reduced representation Bisulfite sequencing. *Reproductive biology and endocrinology : RB&E* 2017; 15: 81.
- 17 Almstrup K, Lindhardt Johansen M, Busch AS, et al. Pubertal development in healthy children is mirrored by DNA methylation patterns in peripheral blood. *Scientific reports* 2016; 6: 28657.
- 18 Demetriou CA, Chen J, Polidoro S, et al. Methylome analysis and epigenetic changes associated with menarcheal age. *PloS one* 2013; 8: e79391.

- 19 Chen S, Refaey H, Mukherjee N, et al. Age at onset of different pubertal signs in boys and girls and differential DNA methylation at age 10 and 18 years: an epigenome-wide follow-up study. *Hum Reprod Open* 2020; 2020: hoaa006.
- 20 Machin M, Amaral AF, Wielscher M, et al. Systematic review of lung function and COPD with peripheral blood DNA methylation in population based studies. *BMC pulmonary medicine* 2017; 17: 54.
- 21 Imboden M, Wielscher M, Rezwan FI, et al. Epigenome-wide association study of lung function level and its change. *Eur Respir J* 2019; 54.
- 22 MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu Rev Psychol* 2007; 58: 593-614.
- 23 Arshad SH, Holloway JW, Karmaus W, et al. Cohort Profile: The Isle Of Wight Whole Population Birth Cohort (IOWBC). *Int J Epidemiol* 2018; 47: 1043-1044i.
- 24 Boyd A, Golding J, Macleod J, et al. Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2013; 42: 111-127.
- 25 Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* 2013; 42: 97-110.
- 26 Northstone K, Lewcock M, Groom A, et al. The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019. *Wellcome Open Res* 2019; 4: 51-51.
- 27 Albert JM, Nelson S. Generalized causal mediation analysis. *Biometrics* 2011; 67: 1028-1038.
- 28 Petersen AC, Crockett L, Richards M, et al. A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of youth and adolescence* 1988; 17: 117-133.
- 29 Yousefi M, Karmaus W, Zhang H, et al. Relationships between age of puberty onset and height at age 18 years in girls and boys. *World journal of pediatrics : WJP* 2013; 9: 230-238.
- 30 Standardization of Spirometry, 1994 Update. American Thoracic Society. *Am J Respir Crit Care Med* 1995; 152: 1107-1136.
- 31 Adalsteinsson BT, Gudnason H, Aspelund T, et al. Heterogeneity in white blood cells has potential to confound DNA methylation measurements. *PloS one* 2012; 7: e46705.
- 32 Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics (Oxford, England)* 2014; 30: 1363-1369.
- 33 Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC bioinformatics* 2012; 13: 86.
- 34 Mahmoud O, Granell R, Tilling K, et al. Association of Height Growth in Puberty with Lung Function. A Longitudinal Study. *American journal of respiratory and critical care medicine* 2018; 198: 1539-1548.
- 35 Ray MA, Tong X, Lockett GA, et al. An Efficient Approach to Screening Epigenome-Wide Data. *BioMed research international* 2016; 2016: 2615348.
- 36 Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of biomedical informatics* 2009; 42: 377-381.
- 37 Barton SJ, Melton PE, Titcombe P, et al. In Epigenomic Studies, Including Cell-Type Adjustments in Regression Models Can Introduce Multicollinearity, Resulting in Apparent Reversal of Direction of Association. *Frontiers in genetics* 2019; 10: 816-816.
- 38 Fan S, Li C, Ai R, et al. Computationally expanding infinium HumanMethylation450 BeadChip array data to reveal distinct DNA methylation patterns of rheumatoid arthritis. *Bioinformatics (Oxford, England)* 2016; 32: 1773-1778.
- 39 Pidsley R, Zotenko E, Peters TJ, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome biology* 2016; 17: 208.

- 40 de Vries M, van der Plaat DA, Nedeljkovic I, et al. From blood to lung tissue: effect of cigarette smoke on DNA methylation and lung function. *Respiratory research* 2018; 19: 212.
- 41 Hafeman DM, Schwartz S. Opening the Black Box: a motivation for the assessment of mediation. *Int J Epidemiol* 2009; 38: 838-845.
- 42 Coleman L, Coleman J. The measurement of puberty: a review. *Journal of Adolescence* 2002; 25: 535-550.

TABLE 1 Comparison of the analytical subsample (n=434) with the whole IOWBC (n=1,261) on ages of pubertal events and lung function.

Variables		Cohort samples	Subsamples	p-value
		mean±SD n=1,261	mean±SD n=434	
Age of pubertal events (years)				
Female				
	Breast growth	12.43±1.58	12.29±1.60	0.20
	Body hair growth	12.25±1.43	12.21±1.42	0.66
	Growth spurt	12.52±1.70	12.36±1.80	0.21
	Skin changes	13.11±1.50	13.08±1.59	0.79
	Menarche	12.72±1.42	12.60±1.37	0.18
Male				
	Body hair growth	13.41±1.37	13.26±1.40	0.14
	Growth spurt	13.72±1.67	13.54±1.63	0.14
	Voice deepening	14.24±1.24	14.07±1.19	0.05
	Facial hair growth	15.38±1.16	15.32±1.09	0.45
	Skin changes	13.99±1.38	13.94±1.29	0.61
Lung function (litres) (age 26)				
Female				
	FVC	4.24±0.54	4.29±0.55	0.24
	FEV ₁	3.42±0.43	3.45±0.42	0.45
Male				
	FVC	5.85±0.82	5.88±0.82	0.89
	FEV ₁	4.61±0.72	4.66±0.72	0.50

TABLE 2A Associations of age of puberty with lung function FVC at age 26 years through DNAm at CpGs survived forward screening at age 18 years in the IOWBC, further tested in the ALSPAC cohort (only direction-consistent indirect effect included). The complete results are in Table S2A.

Pubertal events (sex)	Lung function at age 26	CpGs	Total effect ^{\$} (IOWBC)		Direct effect ^{\$} (IOWBC)		Indirect effect ^{\$} (IOWBC)		Total effect (ALSAPC)		Direct effect (ALSAPC)		Indirect effect (ALSAPC)	
			Est [#]	p-val [#]	Est	p-val	Est	p-val	Est	p-val	Est	p-val	Est	p-val
Menarche (female)	FVC	cg15711028	0.19	0.01	0.15	0.04	0.04	0.04	0.21	8.00×10 ⁻⁴	0.21	9.33×10 ⁻⁴	2.68×10 ⁻³	0.51
		cg08680129	0.18	0.01	0.14	0.06	0.04	0.03	0.22	5.62×10 ⁻⁴	0.20	1.28×10 ⁻³	1.41×10 ⁻²	0.15
		cg01315063	0.19	0.01	0.13	0.07	0.06	0.02	0.21	7.98×10 ⁻⁴	0.21	8.82×10 ⁻⁴	1.74×10 ⁻³	0.59
	FEV ₁	cg17394530	0.23	9.22×10 ⁻⁴	0.28	5.54×10 ⁻⁵	-0.05	0.02	0.23	3.61×10 ⁻⁴	0.23	3.44×10 ⁻⁴	-4.66×10 ⁻⁴	0.90
		cg08680129	0.24	5.92×10 ⁻⁴	0.20	0.01	0.05	0.02	0.23	3.01×10 ⁻⁴	0.22	5.81×10 ⁻⁴	1.02×10 ⁻²	0.25
		cg01315063	0.24	5.85×10 ⁻⁴	0.20	0.01	0.04	0.05	0.23	3.81×10 ⁻⁴	0.23	4.40×10 ⁻⁴	2.58×10 ⁻³	0.54
		cg10219771	0.24	4.91×10 ⁻⁴	0.20	0.01	0.05	0.03	0.23	3.97×10 ⁻⁴	0.23	4.03×10 ⁻⁴	1.84×10 ⁻⁴	0.95
Body hair growth (female)	FEV ₁	cg16332936	0.14	0.11	0.09	0.30	0.05	0.04	0.18	3.10×10 ⁻³	0.18	4.29×10 ⁻³	7.04×10 ⁻³	0.34
		cg04827020	0.17	0.04	0.12	0.16	0.05	0.03	0.18	3.30×10 ⁻³	0.17	5.91×10 ⁻³	1.13×10 ⁻²	0.17
Body hair* growth (males)	FVC	cg13636404	0.14	0.10	0.08	0.37	0.06	0.02	1.0×10 ⁻³	0.99	-4.1×10 ⁻⁴	1.00	1.40×10 ⁻³	0.71
		cg04130455	0.15	0.08	0.23	0.01	-0.08	0.01	3.9×10 ⁻³	0.95	0.01	0.93	-1.78×10 ⁻³	0.83
		cg21276535	0.16	0.05	0.11	0.20	0.06	0.02	4.8×10 ⁻³	0.94	-2.3×10 ⁻⁴	1.00	5.04×10 ⁻³	0.36
		cg27626110	0.11	0.21	0.17	0.05	-0.06	0.04	-7.3×10 ⁻⁴	0.99	-5.6×10 ⁻⁴	0.99	-1.70×10 ⁻⁴	0.96
Voice deepening (males)	FVC	cg01482375	0.17	0.04	0.12	0.16	0.05	0.04	-0.02	0.77	-0.03	0.68	7.88×10 ⁻³	0.28

Note: [#] Est: Regression coefficient estimates with unit in litres (L), it represents expected lung function change for one year increase in pubertal age. p-val: p-value. ^{\$} Total effect is the sum of direct and indirect effects of age of pubertal events on lung function at age 26 years. Indirect effect is the effect of age of pubertal events on lung function via DNAm at CpGs. All the effects are represented by regression coefficients. *Age at body hair growth is identified by age at attainment of Tanner stage>2 in ALSAPC cohort.

TABLE 2B Associations of age of puberty with lung function FVC at age 26 years through DNAm at CpGs survived backward screening at age 18 years in the IOWBC, further tested in the ALSPAC cohort (only direction-consistent indirect effect included). The complete results are in Table S2B.

Pubertal events (sex)	Lung function at age 26	CpGs	Total effect ^{\$} (IOWBC)		Direct effect ^{\$} (IOWBC)		Indirect effect ^{\$} (IOWBC)		Total effect (ALSAPC)		Direct effect (ALSAPC)		Indirect effect (ALSAPC)	
			Est [#]	p-val [#]	Est	p-val	Est	p-val	Est	p-val	Est	p-val	Est	p-val
Body hair growth (females)	FVC	cg04988476	0.14	0.09	0.1	0.22	0.04	0.05	0.14	0.02	0.14	0.02	1.55×10 ⁻³	0.64
Body hair growth (female)	FEV ₁	cg25601319	0.12	0.15	0.08	0.36	0.04	0.03	0.19	2.52×10 ⁻³	0.19	2.50×10 ⁻³	5.74×10 ⁻⁶	1.00
		cg04988476	0.13	0.11	0.09	0.28	0.04	0.04	0.19	2.23×10 ⁻³	0.19	2.52×10 ⁻³	2.39×10 ⁻³	0.54
		cg07143296	0.13	0.12	0.1	0.25	0.04	0.05	0.18	2.97 ×10 ⁻³	0.18	3.52×10 ⁻³	3.86×10 ⁻³	0.52
Body hair* growth (males)	FVC	cg06998210	0.13	0.13	0.09	0.32	0.05	0.05	0.005	0.94	4.31×10 ⁻³	0.95	1.03×10 ⁻³	0.79

Note: [#] Est: Regression coefficient estimates with unit in litres (L), it represents expected lung function change for one year increase in pubertal age. p-val: p-value. ^{\$} Total effect is the sum of direct and indirect effects of age of pubertal events on lung function at age 26 years. Indirect effect is the effect of age of pubertal events on lung function via DNAm at CpGs. All the effects are represented by regression coefficients. *Age at body hair growth is identified by age at attainment of Tanner stage>2 in ALSAPC cohort.

FIGURE 2 The path diagram through DNAm at CpGs at ages 18 years.

*Asthma status at age 10 years, height at age 10 years, BMI at age 10 years, SES at age 10 years, low birth weight status, maternal smoking during pregnancy, atopy status at age 10 years, and cell types

** Smoking at age 18 years

*** Asthma status at age 10 years, height at age 10 years, BMI at age 10 years, SES at age 10 years, low birth weight status, maternal smoking during pregnancy and atopy status at age 10 years

The Principle of mediation analysis is to analyze that an initial variable X may influence an outcome variable Y through a mediating variable M or multiple mediators. In our study, the X is age of onset of puberty; Y is lung function parameters at age 26 years; and M is DNAm at age 18 years. The direct effect of age of pubertal on lung function at age 26 years is assessed by regressing lung function parameter at age 26 years on age at puberty and the regression coefficient represents the direct effect. Indirect effect is assessed by age at puberty indirectly affecting lung function parameter at age 26 through DNAm at age 18 years. A total effect is the sum of direct and indirect effects.

FIGURE 3 Flow chart of identified CpG sites from screening CpGs through mediation analysis.

*: FS: Forward screening; BS: Backward screening.

Does DNA methylation mediate the association of age at puberty with FVC or FEV₁?

Liang Li¹, Hongmei Zhang^{1*}, John W. Holloway², Susan Ewart³, Caroline L. Relton⁴, S. Hasan Arshad^{5,6}, Wilfried Karmaus¹

Affiliations: ¹ Division of Epidemiology, Biostatistics, and Environmental Health, School of Public Health, University of Memphis, Memphis, TN, USA. ² Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK. ³ College of Veterinary Medicine, Michigan State University, East Lansing, MI, USA. ⁴ MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN, UK. ⁵ Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, SO16 6YD, UK. ⁶ The David Hide Asthma and Allergy Research Centre, St Mary's, Hospital, Parkhurst Road, Newport, Isle of Wight PO30 5TG, UK.

***Correspondence:** Hongmei Zhang, Division of Epidemiology, Biostatistics, and Environmental Health, School of Public Health, University of Memphis, Memphis, TN, USA. E-mail: hzhang6@memphis.edu

Supplementary materials

Contents

S. 1. Sample sizes at each follow-up for the study variables.

S. 2. Methods for potential confounders, test of DNAm, cell types, flow chart of screening CpGs using forward and backward screening, and replication cohort ALSPAC.

S. 3. Identification of pubertal events, lung function parameters, and CpG sites to be included in path analyses (Figure 2 and 3 in the main text).

S. 4. DNAm at age 18 years mediated the association of age of pubertal onset with lung function parameters in young adulthood.

Appendix:

S. 1. Sample sizes at each follow-up for the study variables.

S. 2. Methods for potential confounders, test of DNAm, cell types, flow chart of screening CpGs using forward and backward screening, and replication cohort ALSPAC

Potential confounders

Factors related to growth, demographic features, and environmental exposures might confound the association of age at pubertal events with height growth and lung function. Status of low birth weight (<2,500 grams), maternal smoking status during pregnancy, asthma status at age 10 years, body mass index (BMI) at age 10 years, socioeconomic status (SES), atopy at age 10 years, and personal smoking status at age 18 years were included in the analyses as potential confounders. Birth weight and maternal smoking during pregnancy were obtained at birth of the child. Asthma was defined as having ‘physician diagnosed asthma’ and either ‘wheezing or whistling in the chest in the last 12 months’ and/or ‘current treatment for asthma’, using the ISAAC questionnaire [1]. SES was ascertained by ‘low’, ‘medium’ and ‘high’ according to assessment of parental occupation, level of household income and number of rooms in

the house [2]. Atopy was defined by having a skin prick testing (SPT) to at least one allergen test with mean wheal diameter ≥ 3 mm greater than the negative saline control [3].

DNA methylation (DNAm)

In the Isle of Wight birth cohort, peripheral blood samples were collected at ages 18 years. DNA was extracted from whole blood using a standard salting out procedure [4]. DNA concentration was estimated by fluorometric quantitation. For each sample, one microgram DNA was bisulfite-treated for cytosine to thymine conversion using the EZ 96-DNA methylation kit (Zymo Research, Irvine, CA, USA), following the manufacturer's protocol. The HumanMethylation450 BeadChip interrogates >484,000 CpGs and MethylationEPIC BeadChip has > 850,000 CpGs both associated with over 24,000 genes. Arrays were processed using a standard protocol with multiple identical control samples assigned to each bisulfite conversion batch to assess assay variability. DNA samples were randomly distributed on the arrays aiming to eliminate batch effects. DNAm was calculated by a ratio of methylated (M) over the sum of methylated and unmethylated (U) probes ($\beta = M/[c+M+U]$), where c is used as a constant to prevent zero in the denominator[5], denoted as β values for each CpG site. In our study, β values were logit transformed with base 2 (denoted as M values), and M values were used in all the analyses.

DNAm quality control and pre-processing

Probes not reaching a detection p-value of 10^{-16} in at least 95% of samples were excluded. Quantile normalization was applied to DNA methylation intensities as suggested in the CPACOR pipeline [6]. Autosomal probes were then extracted and converted to β values. ComBat was used to logit transformed β values to remove batch effects [7]. CpG sites with probe-SNPs within ten base pairs of the CpG site and with minor allele frequency (MAF) greater than 0.007 in the IOWBC (i.e., $\sim \geq 10$ out of 1,456 subjects expected to have the minor allele in the cohort) were excluded.

Cell-type proportions

DNAm is influenced by heterogeneity of cell type compositions in peripheral blood, which may lead to biased inferences. Proportions of six cell types, CD4+ T cells, natural killer cells, neutrophil, B cells, monocytes, and eosinophils, were inferred using the R package *minfi* [8, 9] and included in the analyses as confounders.

Replication cohort ALSPAC

Pregnant women resident in Avon, UK with expected dates of delivery 1st April 1991 to 31st December 1992 were invited to take part in the study. The initial number of pregnancies enrolled is 14,541 (for these at least one questionnaire has been returned or a “Children in Focus” clinic had been attended by 19th July 1999). Of these initial pregnancies, there was a total of 14,676 fetuses, resulting in 14,062 live births and 13,988 children who were alive at 1 year of age. When the oldest children were approximately 7 years of age, an attempt was made to bolster the initial sample with eligible cases who had failed to join

the study originally. As a result, when considering variables collected from the age of seven onwards (and potentially abstracted from obstetric notes) there are data available for more than the 14,541 pregnancies mentioned above. The number of new pregnancies not in the initial sample (known as Phase I enrolment) that are currently represented on the built files and reflecting enrolment status at the age of 24 is 913 (456, 262 and 195 recruited during Phases II, III and IV respectively), resulting in an additional 913 children being enrolled. The phases of enrolment are described in more detail in the cohort profile paper and its update (see footnote 4 below). The total sample size for analyses using any data collected after the age of seven is therefore 15,454 pregnancies, resulting in 15,589 fetuses. Of these 14,901 were alive at 1 year of age. A 10% sample of the ALSPAC cohort, known as the Children in Focus (CiF) group, attended clinics at the University of Bristol at various time intervals between 4 to 61 months of age. The CiF group were chosen at random from the last 6 months of ALSPAC births (1432 families attended at least one clinic). Excluded were those mothers who had moved out of the area or were lost to follow-up, and those partaking in another study of infant development in Avon. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Consent for biological samples has been collected in accordance with the Human Tissue Act (2004). Please note that the study website contains details of all the data that is available through a fully searchable data dictionary and variable search tool: <http://www.bristol.ac.uk/alspac/researchers/our-data/>.

S. 3. Identification of pubertal events, lung function parameters, and CpG sites to be included in path analyses (Figure 2 and 3 in the main text).

To identify whether age of onset of pubertal events and lung function parameters were included in path analyses, we examined the association of ages of these pubertal events with lung function parameters FVC and FEV₁ at age 26 years in the IOWBC. Statistical significance is set at 0.05.

In females, at age 26 years lung function parameters, early ages of onset of body hair growth and menarche onset showed statistically significant associations with lower FVC and FEV₁ (Table S1). For males, ages of body hair growth, growth spurt, voice deepening, and appearance of facial hair growth

were associated with FVC at age 26 years. No other statistically significant associations were identified at these two ages. For males, FEV₁ was not included in path analyses, since it was not associated the age of any pubertal events.

Table S1 The association of age of pubertal events and lung function measures at age 26 years. Potential confounders adjusted were listed in Figure 2 in the main text, including: asthma status at age 10 years, height at age 10 years, BMI at age 10 years, SES at age 10 years, low birth weight status, maternal smoking during pregnancy and atopy status at age 10 years.

Lung function	Female			Male		
	Pubertal events	Estimate [#]	p-value	Pubertal events	Estimate	p-value
FVC	Breast growth	0.041	0.062	Body hair growth	0.104	0.015
	Body hair growth	0.051	0.032	Growth spurt	0.086	0.019
	Growth spurt	0.040	0.059	Voice deepening	0.110	0.021
	Skin changes	0.031	0.211	Facial hair growth	0.112	0.028
	Menarche	0.076	0.001	Skin changes	0.057	0.257
FEV ₁	Breast growth	0.019	0.271	Body hair growth	0.056	0.140
	Body hair growth	0.042	0.027	Growth spurt	0.041	0.217
	Growth spurt	0.025	0.127	Voice deepening	0.050	0.247
	Skin changes	0.030	0.135	Facial hair growth	0.032	0.484
	Menarche	0.066	<0.001	Skin changes	0.037	0.411
	Menarche	0.001	0.629	Skin changes	-0.001	0.809

[#]: Coeff. Estimate: Regression coefficient estimates, and the unit presented in litres (L). It represents expected lung function change for one year increase in pubertal age.

Next, we screened CpGs from two directions, which comprised of puberty-related CpGs (forward screening) and lung function-related CpGs (backward screening). (1) Forward screening was used to identify candidate CpGs such that DNAm at those CpGs was associated with ages of pubertal events. The pubertal events that passed the initial screening on their association with FVC or FEV₁ were included in this forward screening. For females, at 198 and 134 CpGs (in total, 321 unique CpGs with 11 overlappings), DNAm was associated with age at menarche and body hair growth, respectively (Figure 3). For males, at 152, 105, 131, and 109 CpGs (in total, 489 unique CpGs with 8 overlappings), DNAm

was associated with age at body hair growth, growth spurt, voice deepening, and facial hair growth, respectively (Figure 3). There were no common CpGs between males and females. CpGs (in total 321 unique CpGs for females and 489 unique CpGs for males) that survived this forward screening were included in subsequent path analyses. (2) In the backward screening of CpGs, for females, lung function parameters FVC and FEV₁ were used. For females, at 102 and 174 CpGs (in total, 259 unique CpGs with 17 overlappings), DNAm was associated with FVC and FEV₁ at age 26, respectively (Figure 3). For males, since in the initial screening, no pubertal events were associated with FEV₁, this lung function parameter was not used to screen CpG sites. In total, 62 CpGs were associated with FVC at age 26 (Figure 3). As in the forward screening, no CpGs overlapped between males and females. CpGs that survived this backward screening were treated as candidate CpGs and included in subsequent path analyses (Figure 3).

S. 4. DNAm at age 18 years mediated the association of age of pubertal onset with lung function parameters in young adulthood.

Further assessment of the IOWBC discovered CpGs in ALSPAC

In total, 21 CpGs for females and 20 CpGs for males discovered in IOWBC were assessed in ALSPAC. Lung function measurement at age 24 years and DNAm at age 15 or 17 years in ALSPAC were analyzed.

In females, of the 21 unique CpGs, indirect effects at 10 CpGs (Table 2A and 2B) were consistent with those observed in IOWBC, but none of them showed statistically significant mediation effects. A more striking observation was the low ratio of indirect effect over direct effect at these 10 CpGs; all ratios were <0.07 , substantially lower than those observed in the IOWBC. In males, ages of body hair growth and voice deepening were available in ALSPAC. Of the 11 IOWBC-discovered CpGs for these two pubertal events, consistent mediation effects were observed at 6 CpGs but were statistically insignificant (Table 2A and 2B). Overall, both direct and indirect effects of these two pubertal events were negligible in ALSPAC.

We also examined whether cell-type composition played a role in the above statistically insignificant findings potentially due to multicollinearity between certain cell types and other variables in the model [29]. In particular, we removed cell type proportions of the six cells from mediation analysis models in both cohorts. For the indirect effects discovered in IOWBC, all the directions were kept unchanged (Table S3A and S3B). In ALSPAC, the direction of indirect effects changed at only 2 of the 32 CpGs tested in ALSPAC (cg27626110 and cg25601319) (Table S3A and S3B). All the results in both cohorts indicated that cell types did not play any significant role in terms of the assessment of mediating effects (Tables S2A, S2B, S3A, and S3B). In summary, the IOWBC-discovered CpGs were not supported by data in ALSPAC regarding their role as a mediator, and such findings were not affected by cell type compositions.

Table S2A Association of age of pubertal events and lung function at age 26 through DNAm at CpGs survived forward screening at age 18 years in IOWBC. Potential confounders adjusted were listed in Figure 2 in the main text, including: asthma status at age 10 years, height at age 10 years, BMI at age 10 years, SES at age 10 years, low birth weight status, maternal smoking during pregnancy, atopy status at age 10 years, cell types, and smoking at age 18 years.

Pubertal events(sex)	Lung function at age 26	CpGs [®]	Gene	Region	Total effect [§] (IOWBC)		Direct effect [§] (IOWBC)		Indirect effect [§] (IOWBC)		Total effect (ALSAPC)		Direct effect (ALSAPC)		Indirect effect (ALSAPC)	
					Est [#]	p-val [#]	Est	p-val	Est	P-val	Est	p-val	Est	p-val	Est	P-val
Menarche (female)	FVC	cg15711028	<i>EIF4EBP2</i>	Body	0.19	0.01	0.15	0.04	0.04	0.04	0.21	8.00×10 ⁻⁴	0.21	9.33×10 ⁻⁴	2.68×10 ⁻³	0.51
		cg08680129	<i>PRX</i>	TSS1500	0.18	0.01	0.14	0.06	0.04	0.03	0.22	5.62×10 ⁻⁴	0.20	1.28×10 ⁻³	1.41×10 ⁻²	0.15
		cg05325369	<i>FBXO46</i>	5'UTR	0.17	0.02	0.13	0.07	0.04	0.04	0.21	8.49×10 ⁻⁴	0.21	8.37×10 ⁻⁴	-2.85×10 ⁻⁴	0.84
		cg09131512	<i>IER3</i>	TSS1500	0.17	0.02	0.21	5.60×10 ⁻³	-0.04	0.04	0.22	5.55×10 ⁻⁴	0.21	6.82×10 ⁻⁴	4.56×10 ⁻³	0.52
		cg12033187		intergenic	0.16	0.03	0.13	0.09	0.04	0.05	0.21	1.17×10 ⁻³	0.21	1.14×10 ⁻³	-3.33×10 ⁻⁴	0.90
		cg01315063	<i>C17orf96</i>	1stExon	0.19	0.01	0.13	0.07	0.06	0.02	0.21	7.98×10 ⁻⁴	0.21	8.82×10 ⁻⁴	1.74×10 ⁻³	0.59
Menarche (female)	FEV1	cg03341697	<i>FUT4</i>	5'UTR;1stExon	0.23	1.09×10 ⁻³	0.20	0.01	0.04	0.05	0.23	4.06×10 ⁻⁴	0.23	3.89×10 ⁻⁴	-7.40×10 ⁻⁴	0.74
		cg17394530		intergenic	0.23	9.22×10 ⁻⁴	0.28	5.54×10 ⁻⁵	-0.05	0.02	0.23	3.61×10 ⁻⁴	0.23	3.44×10 ⁻⁴	-4.66×10 ⁻⁴	0.90
		cg08894381		intergenic	0.23	1.01×10 ⁻³	0.28	1.08×10 ⁻⁴	-0.04	0.04	0.23	4.21×10 ⁻⁴	0.23	4.40×10 ⁻⁴	6.95×10 ⁻⁴	0.77
		cg08680129	<i>PRX</i>	TSS1500	0.24	5.92×10 ⁻⁴	0.20	0.01	0.05	0.02	0.23	3.01×10 ⁻⁴	0.22	5.81×10 ⁻⁴	1.02×10 ⁻²	0.25
		cg09131512	<i>IER3</i>	TSS1500	0.23	1.58×10 ⁻³	0.27	1.17×10 ⁻⁴	-0.05	0.03	0.23	2.90×10 ⁻⁴	0.23	3.40×10 ⁻⁴	2.91×10 ⁻³	0.54
		cg01315063	<i>C17orf96</i>	1stExon	0.24	5.85×10 ⁻⁴	0.20	0.01	0.04	0.05	0.23	3.81×10 ⁻⁴	0.23	4.40×10 ⁻⁴	2.58×10 ⁻³	0.54
		cg10219771		intergenic	0.24	4.91×10 ⁻⁴	0.20	0.01	0.05	0.03	0.23	3.97×10 ⁻⁴	0.23	4.03×10 ⁻⁴	1.84×10 ⁻⁴	0.95
Body hair growth (female)	FEV1	cg05414613	<i>SLC1A7</i>	Body	0.14	0.10	0.18	0.03	-0.05	0.03	0.19	2.35×10 ⁻³	0.19	2.44×10 ⁻³	8.90×10 ⁻⁴	0.72
		cg16332936	<i>TBX1</i>	Body	0.14	0.11	0.09	0.30	0.05	0.04	0.18	3.10×10 ⁻³	0.18	4.29×10 ⁻³	7.04×10 ⁻³	0.34
		cg04827020	<i>THBS1</i>	Body	0.17	0.04	0.12	0.16	0.05	0.03	0.18	3.30×10 ⁻³	0.17	5.91×10 ⁻³	1.13×10 ⁻²	0.17
		cg23328390	<i>C6orf134</i>	TSS200	0.15	0.08	0.19	0.03	-0.04	0.05	0.19	2.77×10 ⁻³	0.18	2.90×10 ⁻³	6.87×10 ⁻⁴	0.76
Body hair growth [□] (male)	FVC	cg19979646	<i>SUGT1</i>	TSS200	0.11	0.19	0.18	0.04	-0.07	0.03	-1.73×10 ⁻³	0.98	-0.01	0.93	4.17×10 ⁻³	0.53
		cg13636404	<i>ZNF655</i>	TSS1500	0.14	0.10	0.08	0.37	0.06	0.02	9.97×10 ⁻⁴	0.99	-4.05×10 ⁻⁴	1.00	1.40×10 ⁻³	0.71

Growth spurt* (male)	FVC	cg04130455		intergenic	0.15	0.08	0.23	0.01	-0.08	0.01	3.86×10 ⁻³	0.95	0.01	0.93	-1.78×10 ⁻³	0.83
		cg05679686	<i>CORO1C</i>	Body	0.15	0.09	0.10	0.25	0.05	0.04	1.39×10 ⁻²	0.84	0.02	0.74	-8.56×10 ⁻³	0.32
		cg21276535	<i>MAP2K1</i>	Body	0.16	0.05	0.11	0.20	0.06	0.02	4.81×10 ⁻³	0.94	-2.28×10 ⁻⁴	1.00	5.04×10 ⁻³	0.36
		cg09616494	<i>DTHD1</i>	Body	0.13	0.15	0.08	0.38	0.05	0.04	1.66×10 ⁻³	0.98	3.00×10 ⁻³	0.96	-1.34×10 ⁻³	0.69
		cg27626110	<i>SFXN1</i>	5'UTR	0.11	0.21	0.17	0.05	-0.06	0.04	-7.32×10 ⁻⁴	0.99	-5.63×10 ⁻⁴	0.99	-1.70×10 ⁻⁴	0.96
				5'UTR;TSS1												
		cg18278528	<i>SYNPO</i>	500	0.17	0.06	0.11	0.28	0.07	0.03						
		cg16965936	<i>PRDM15</i>	Body	0.18	0.04	0.24	0.01	-0.05	0.03						
		cg24011637	<i>FOSL2</i>	TSS1500	0.19	0.04	0.13	0.19	0.06	0.04						
Voice deepening (male)	FVC	cg06405341	<i>AXIN2</i>	Body	0.18	0.03	0.22	0.01	-0.04	0.04	-0.01	0.86	-0.01	0.84	1.50×10 ⁻³	0.70
		cg01482375	<i>LATS1</i>	5'UTR	0.17	0.04	0.12	0.16	0.05	0.04	-0.02	0.77	-0.03	0.68	7.88×10 ⁻³	0.28
Facial hair* (male)	FVC	cg01464985	<i>ZNF512</i>	TSS1500	0.15	0.07	0.22	0.01	-0.06	0.01						
		cg26065534	<i>ATP8A2</i>	Body	0.15	0.08	0.11	0.23	0.05	0.04						
		cg01808969	<i>FOXK1</i>	Body	0.16	0.07	0.22	0.02	-0.06	0.03						
		cg17959850	<i>EMB</i>	Body	0.16	0.06	0.10	0.26	0.06	0.03						

□: Age at body hair growth is identified by age at attainment of Tanner stage>2 in ALSAPC cohort. #: Est: Regression coefficient estimates, and the unit presented in litres(L). It represents the expected lung function change for one year increase in pubertal age. *: Growth spurt and facial hair in males are not available in ALSPAC. \$: Total effect is the sum of direct and indirect effects of age of pubertal events on lung function at age 26 years. Indirect effect is the effect of age of pubertal events on lung function via height growth and its related CpGs. @: CpGs name in bold font had direction-consistent mediation effects in IOWBC and ALSPAC.

Table S2B Association of age of pubertal events and lung function at age 26 through DNAm at CpGs survived backward screening at age 18 years in IOWBC. Potential confounders adjusted were listed in Figure 2 in the main text, including: asthma status at age 10 years, height at age 10 years, BMI at age 10 years, SES at age 10 years, low birth weight status, maternal smoking during pregnancy, atopy status at age 10 years, cell types, and smoking at age 18 years.

Pubertal events(sex)	Lung function at age 26	CpGs [@]	Gene	Region	Total effect ^{\$} (IOWBC)		Direct effect ^{\$} (IOWBC)		Indirect effect ^{\$} (IOWBC)		Total effect (ALSAPC)		Direct effect (ALSAPC)		Indirect effect (ALSAPC)	
					Est [#]	p-val [#]	Est	p-val	Est	P-val	Est	p-val	Est	p-val	Est	P-val
Menarche (female)	FEV1	cg25601319	<i>ARID1A</i>	TSS1500	0.22	1.93×10 ⁻³	0.18	0.01	0.04	0.03	0.23	4.16×10 ⁻⁴	0.23	4.07×10 ⁻⁴	-2.25×10 ⁻⁴	0.91
		cg12614633	<i>USP48</i>	TSS1500	0.23	1.37×10 ⁻³	0.18	0.01	0.05	0.03	0.23	3.34×10 ⁻⁴	0.23	3.23×10 ⁻⁴	-8.50×10 ⁻⁴	0.79
		cg05060880		intergenic	0.21	3.72×10 ⁻³	0.16	0.02	0.05	0.04	0.23	3.44×10 ⁻⁴	0.23	3.00×10 ⁻⁴	-1.79×10 ⁻³	0.70
Body hair growth (female)	FVC	cg00498211	<i>SLC1A4</i>	Body;5'UTR	0.17	0.04	0.12	0.16	0.05	0.01	0.14	0.02	0.15	0.02	-2.90×10 ⁻³	0.60
		cg24432048	<i>LOC100133612</i>	Body	0.14	0.10	0.10	0.25	0.04	0.05	0.14	0.02	0.15	0.02	-2.71×10 ⁻³	0.55
		cg04988476	<i>FGF12</i>	5'UTR	0.14	0.09	0.10	0.22	0.04	0.05	0.14	0.02	0.14	0.02	1.55×10 ⁻³	0.64
Body hair growth (female)	FEV1	cg12614633	<i>USP48</i>	TSS1500	0.13	0.12	0.06	0.48	0.07	0.01	0.19	2.34×10 ⁻³	0.19	2.32×10 ⁻³	-8.80×10 ⁻⁵	0.93
		cg25601319	<i>ARID1A</i>	TSS1500	0.12	0.15	0.08	0.36	0.04	0.03	0.19	2.52×10 ⁻³	0.19	2.50×10 ⁻³	5.74×10 ⁻⁶	1.00
		cg04988476	<i>FGF12</i>	5'UTR	0.13	0.11	0.09	0.28	0.04	0.04	0.19	2.23×10 ⁻³	0.19	2.52×10 ⁻³	2.39×10 ⁻³	0.54
		cg07143296	<i>PCNXL3</i>	TSS200	0.13	0.12	0.10	0.25	0.04	0.05	0.18	2.97×10 ⁻³	0.18	3.52×10 ⁻³	3.86×10 ⁻³	0.52
Body hair growth [□] (male)	FVC	cg06998210	<i>CA7</i>	Body;TSS1500	0.13	0.13	0.09	0.32	0.05	0.05	0.005	0.94	4.31×10 ⁻³	0.95	1.03×10 ⁻³	0.79
Growth spurt* (male)	FVC	cg01907761	<i>SPINK5L3</i>	TSS1500	0.20	0.03	0.15	0.10	0.05	0.05						
Voice deepening (male)	FVC	cg21300403	<i>BOLL</i>	Body;TSS200	0.15	0.08	0.10	0.21	0.04	0.05	-0.01	0.86	-0.01	0.90	-2.77×10 ⁻³	0.58
Facial hair* (male)	FVC	cg14834881	<i>FAM190A</i>	Body	0.15	0.08	0.10	0.26	0.05	0.03						

□: Age at body hair growth is identified by age at attainment of Tanner stage>2 in ALSAPC cohort. #: Est: Regression coefficient estimates, and the unit presented in litres(L). It represents the expected lung function change for one year increase in pubertal age. *: Growth spurt and facial hair in males are not available in ALSPAC. \$: Total effect is the sum of direct and indirect effects of age of pubertal events on lung function at age 26 years. Indirect effect is the effect of age of pubertal events on lung function via height growth and its related CpGs. @: CpGs name in bold font had direction-consistent mediation effects in IOWBC and ALSPAC.

Table S3A Association of age of pubertal events and lung function at age 26 through DNAm at CpGs survived forward screening at age 18 years in IOWBC without cell types in model. Potential confounders adjusted were listed in Figure 2 in the main text, including: asthma status at age 10 years, height at age 10 years, BMI at age 10 years, SES at age 10 years, low birth weight status, maternal smoking during pregnancy, atopy status at age 10 years, and smoking at age 18 years.

Pubertal events(sex)	Lung function at age 26	CpGs®	Gene	Region	Total effect ^s (IOWBC)		Direct effect ^s (IOWBC)		Indirect effect ^s (IOWBC)		Total effect (ALSAPC)		Direct effect (ALSAPC)		Indirect effect (ALSAPC)	
					Est [#]	p-val [#]	Est	p-val	Est	P-val	Est	p-val	Est	p-val	Est	P-val
Menarche (female)	FVC	cg15711028	<i>EIF4EBP2</i>	Body	0.19	0.01	0.15	0.03	0.04	0.05	0.21	4.41×10 ⁻⁴	0.21	5.64×10 ⁻⁴	3.96×10 ⁻³	0.45
		cg08680129	<i>PRX</i>	TSS1500	0.18	0.01	0.14	0.05	0.04	0.04	0.22	3.47×10 ⁻⁴	0.21	7.74×10 ⁻⁴	1.31×10 ⁻²	0.18
		cg05325369	<i>FBXO46</i>	5'UTR	0.17	0.02	0.14	0.06	0.04	0.05	0.21	5.31×10 ⁻⁴	0.21	5.10×10 ⁻⁴	-4.60×10 ⁻⁴	0.81
		cg09131512	<i>IER3</i>	TSS1500	0.17	0.02	0.21	3.10×10 ⁻³	-0.04	0.04	0.22	2.93×10 ⁻⁴	0.21	4.07×10 ⁻⁴	6.78×10 ⁻³	0.39
		cg12033187		intergenic	0.17	0.02	0.13	0.08	0.04	0.05	0.21	8.26×10 ⁻³	0.21	7.82×10 ⁻³	-4.60×10 ⁻⁴	0.86
		cg01315063	<i>C17orf96</i>	1stExon	0.19	0.01	0.14	0.06	0.06	0.02	0.21	4.70×10 ⁻⁴	0.21	5.58×10 ⁻⁴	2.65×10 ⁻³	0.54
Menarche (female)	FEV1	cg03341697	<i>FUT4</i>	5'UTR;1stExon	0.24	8.18×10 ⁻⁴	0.20	0.01	0.04	0.04	0.23	2.34×10 ⁻⁴	0.23	2.17×10 ⁻⁴	-3.60×10 ⁻⁴	0.85
		cg17394530		intergenic	0.24	7.69×10 ⁻⁴	0.28	5.60×10 ⁻³	-0.04	0.04	0.23	2.27×10 ⁻⁴	0.23	1.93×10 ⁻⁴	-2.16×10 ⁻³	0.60
		cg08894381		intergenic	0.24	9.4×10 ⁻⁴	0.28	1.06×10 ⁻⁴	-0.04	0.04	0.23	2.53×10 ⁻⁴	0.23	2.65×10 ⁻⁴	4.27×10 ⁻⁴	0.80
		cg08680129	<i>PRX</i>	TSS1500	0.24	7.42×10 ⁻⁴	0.20	5.67×10 ⁻³	0.04	0.03	0.23	1.82×10 ⁻⁴	0.22	3.41×10 ⁻⁴	9.51×10 ⁻³	0.29

		cg09131512	<i>IER3</i>	TSS1500	0.23	1.67×10 ⁻³	0.27	1.16×10 ⁻⁴	-0.05	0.03	0.24	1.53×10 ⁻⁴	0.23	1.91×10 ⁻⁴	4.36×10 ⁻³	0.45
		cg01315063	<i>C17orf96</i>	1stExon	0.25	5.13×10 ⁻⁴	0.20	0.01	0.04	0.05	0.23	2.06×10 ⁻⁴	0.23	2.67×10 ⁻⁴	4.12×10 ⁻³	0.45
		cg10219771		intergenic	0.24	5.62×10 ⁻⁴	0.20	0.01	0.05	0.03	0.23	2.33×10 ⁻⁴	0.23	2.46×10 ⁻⁴	1.88×10 ⁻⁴	0.94
Body hair growth (female)	FEV1	cg05414613	<i>SLC1A7</i>	Body	0.15	0.08	0.19	0.02	-0.04	0.05	0.19	2.35×10 ⁻³	0.19	2.33×10 ⁻³	6.30×10 ⁻⁵	0.98
		cg16332936	<i>TBX1</i>	Body	0.14	0.09	0.10	0.25	0.04	0.04	0.18	3.15×10 ⁻³	0.18	4.37×10 ⁻³	6.39×10 ⁻³	0.38
		cg04827020	<i>THBS1</i>	Body	0.18	0.03	0.12	0.14	0.05	0.03	0.18	3.40×10 ⁻³	0.17	6.26×10 ⁻³	1.09×10 ⁻²	0.21
		cg23328390	<i>C6orf134</i>	TSS200	0.16	0.06	0.20	0.02	-0.04	0.05	0.19	2.62×10 ⁻³	0.18	2.79×10 ⁻³	8.02×10 ⁻⁴	0.77
Body hair growth [□] (male)	FVC	cg19979646	<i>SUGT1</i>	TSS200	0.11	0.21	0.18	0.04	-0.07	0.02	-1.73×10 ⁻³	0.98	-0.01	0.94	3.81×10 ⁻³	0.64
		cg13636404	<i>ZNF655</i>	TSS1500	0.14	0.12	0.07	0.41	0.06	0.02	5.49×10 ⁻⁴	0.99	3.04×10 ⁻⁴	1.00	2.45×10 ⁻⁴	0.95
		cg04130455		intergenic	0.15	0.08	0.23	0.01	-0.07	0.01	3.40×10 ⁻³	0.96	0.01	0.93	-2.52×10 ⁻³	0.76
		cg05679686	<i>CORO1C</i>	Body	0.15	0.10	0.10	0.26	0.05	0.04	1.48×10 ⁻²	0.83	0.02	0.76	-6.63×10 ⁻³	0.43
		cg21276535	<i>MAP2K1</i>	Body	0.16	0.05	0.11	0.20	0.06	0.02	2.81×10 ⁻³	0.97	9.05×10 ⁻⁵	1.00	2.72×10 ⁻³	0.64
		cg09616494	<i>DTHD1</i>	Body	0.12	0.16	0.08	0.41	0.05	0.05	1.85×10 ⁻³	0.98	3.49×10 ⁻³	0.96	-1.64×10 ⁻³	0.66
		cg27626110	<i>SFXN1</i>	5'UTR	0.11	0.22	0.17	0.05	-0.06	0.03	1.74×10 ⁻⁴	1.00	-4.20×10 ⁻⁴	1.00	5.92×10 ⁻⁴	0.86
Growth spurt* (male)	FVC			5'UTR;TSS1500												
		cg18278528	<i>SYNPO</i>	500	0.18	0.06	0.12	0.20	0.06	0.03						
		cg16965936	<i>PRDM15</i>	Body	0.19	0.04	0.24	0.01	-0.05	0.05						
Voice deepening (male)	FVC	cg24011637	<i>FOSL2</i>	TSS1500	0.18	0.05	0.12	0.20	0.06	0.04						
		cg06405341	<i>AXIN2</i>	Body	0.17	0.04	0.23	0.01	-0.05	0.04	-0.01	0.84	-0.01	0.83	7.10×10 ⁻⁴	0.85
Facial hair* (male)	FVC	cg01482375	<i>LATS1</i>	5'UTR	0.17	0.04	0.12	0.16	0.05	0.04	-0.02	0.78	-0.03	0.67	1.06×10 ⁻²	0.26
		cg01464985	<i>ZNF512</i>	TSS1500	0.16	0.05	0.22	0.01	-0.05	0.04						
		cg26065534	<i>ATP8A2</i>	Body	0.15	0.10	0.10	0.26	0.04	0.05						
		cg01808969	<i>FOXK1</i>	Body	0.16	0.07	0.21	0.02	-0.05	0.04						
		cg17959850	<i>EMB</i>	Body	0.15	0.08	0.09	0.29	0.06	0.03						

□: Age at body hair growth is identified by age at attainment of Tanner stage>2 in ALSAPC cohort. #: Est: Regression coefficient estimates, and the unit presented in litres(L). It represents the expected lung function change for one year increase in pubertal age. *: Growth spurt and facial hair in males are not available in ALSPAC. \$: Total effect is the sum of direct and indirect effects of age of pubertal events on lung function at age 26 years. Indirect effect is the effect of age of pubertal events on lung function via height growth and its related CpGs. @: CpGs name in bold font had direction-consistent mediation effects in IOWBC and ALSPAC.

Table S3B Association of age of pubertal events and lung function at age 26 through DNAm at CpGs survived backward screening at age 18 years in IOWBC without cell types in model. Potential confounders adjusted were listed in Figure 2 in the main text, including: asthma status at age 10 years, height at age 10 years, BMI at age 10 years, SES at age 10 years, low birth weight status, maternal smoking during pregnancy, atopy status at age 10 years, and smoking at age 18 years.

Pubertal events(sex)	Lung function at age 26	CpGs [@]	Gene	Region	Total effect ^{\$} (IOWBC)		Direct effect ^{\$} (IOWBC)		Indirect effect ^{\$} (IOWBC)		Total effect (ALSAPC)		Direct effect (ALSAPC)		Indirect effect (ALSAPC)	
					Est [#]	p-val [#]	Est	p-val	Est	P-val	Est	p-val	Est	p-val	Est	P-val
Menarche (female)	FEV1	cg25601319	<i>ARID1A</i>	TSS1500	0.22	2.46×10 ⁻³	0.18	0.01	0.04	0.04	0.23	2.23×10 ⁻⁴	0.23	2.04×10 ⁻⁴	1.31×10 ⁻³	0.68
		cg12614633	<i>USP48</i>	TSS1500	0.23	1.26×10 ⁻³	0.18	0.01	0.05	0.03	0.23	1.94×10 ⁻⁴	0.23	1.84×10 ⁻⁴	-1.08×10 ⁻³	0.80
		cg05060880		intergenic	0.21	3.66×10 ⁻³	0.16	0.02	0.05	0.04	0.23	1.90×10 ⁻⁴	0.23	1.74×10 ⁻⁴	-7.60×10 ⁻⁴	0.88
Body hair growth (female)	FVC	cg00498211	<i>SLC1A4</i>	Body;5'UTR	0.16	0.05	0.13	0.12	0.04	0.06	0.14	0.02	0.15	0.01	-5.75×10 ⁻³	0.38
		cg24432048	<i>LOC100133612</i>	Body	0.15	0.08	0.11	0.21	0.04	0.04	0.14	0.02	0.15	0.02	-2.14×10 ⁻³	0.67
		cg04988476	<i>FGF12</i>	5'UTR	0.15	0.06	0.11	0.18	0.04	0.04	0.14	0.02	0.14	0.02	1.04×10 ⁻³	0.69
Body hair growth (female)	FEV1	cg12614633	<i>USP48</i>	TSS1500	0.13	0.11	0.07	0.43	0.07	0.01	0.19	2.38×10 ⁻³	0.19	2.37×10 ⁻³	-1.03×10 ⁻⁴	0.92
		cg25601319	<i>ARID1A</i>	TSS1500	0.13	0.14	0.08	0.33	0.04	0.04	0.19	2.46×10 ⁻³	0.19	2.54×10 ⁻³	8.19×10 ⁻⁴	0.79
		cg04988476	<i>FGF12</i>	5'UTR	0.15	0.08	0.10	0.23	0.04	0.03	0.19	2.27×10 ⁻³	0.19	2.44×10 ⁻³	1.61×10 ⁻⁴	0.63
		cg07143296	<i>PCNXL3</i>	TSS200	0.14	0.10	0.10	0.22	0.03	0.05	0.18	2.87×10 ⁻³	0.18	3.52×10 ⁻³	4.22×10 ⁻³	0.50
Body hair growth [□] (male)	FVC	cg06998210	<i>CA7</i>	Body;TSS1500	0.13	0.15	0.08	0.34	0.04	0.06	0.005	0.94	4.50×10 ⁻³	0.95	5.96×10 ⁻⁴	0.87
Growth spurt* (male)	FVC	cg01907761	<i>SPINK5L3</i>	TSS1500	0.19	0.03	0.15	0.10	0.04	0.07						
Voice deepening (male)	FVC	cg21300403	<i>BOLL</i>	Body;TSS200	0.15	0.07	0.11	0.21	0.05	0.04	-0.01	0.85	-0.01	0.87	-2.39×10 ⁻³	0.63
Facial hair* (male)	FVC	cg14834881	<i>FAM190A</i>	Body	0.14	0.11	0.09	0.28	0.05	0.05						

□: Age at body hair growth is identified by age at attainment of Tanner stage>2 in ALSAPC cohort. #: Est: Regression coefficient estimates, and the unit presented in litres(L). It represents the expected lung function change for one year increase in pubertal age. *: Growth spurt and facial hair in males are not available in ALSPAC. \$: Total effect is the sum of direct and indirect effects of age of pubertal events on lung function at age 26 years. Indirect effect is the effect of age of pubertal events on lung function via height growth and its related CpGs. @: CpGs name in bold font had direction-consistent mediation effects in IOWBC and ALSPAC.

References

- 1 Asher MI, Keil U, Anderson HR, et al. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. *Eur Respir J* 1995; 8: 483-491.
- 2 Ogbuanu IU, Karmaus W, Arshad SH, et al. Effect of breastfeeding duration on lung function at age 10 years: a prospective birth cohort study. *Thorax* 2009; 64: 62-66.
- 3 Kurukulaaratchy RJ, Zhang H, Patil V, et al. Identifying the heterogeneity of young adult rhinitis through cluster analysis in the Isle of Wight birth cohort. *The Journal of allergy and clinical immunology* 2015; 135: 143-150.
- 4 Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic acids research* 1988; 16: 1215.
- 5 Kuan PF, Wang S, Zhou X, et al. A statistical framework for Illumina DNA methylation arrays. *Bioinformatics (Oxford, England)* 2010; 26: 2849-2855.
- 6 Lehne B, Drong AW, Loh M, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome biology* 2015; 16: 37.
- 7 Johnson WE, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; 8: 118-127.
- 8 Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics (Oxford, England)* 2014; 30: 1363-1369.
- 9 Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC bioinformatics* 2012; 13: 86.