HHS Public Access

Author manuscript

Dev Psychopathol. Author manuscript; available in PMC 2022 August 01.

Published in final edited form as:

Dev Psychopathol. 2022 February ; 34(1): 355–366. doi:10.1017/S0954579420000929.

Father Absence, Age at Menarche, and Genetic Confounding: A Replication and Extension using a Polygenic Score

Gabriel L. Schlomer[1], **Kristine Marceau**[2]

[1]Division of Educational Psychology and Methodology, University at Albany, SUNY

[2]Human Development and Family Studies, Purdue University

Abstract

Father absence has a small but robust association with earlier age at menarche (AAM), likely reflecting both genetic confounding and an environmental influence on life history strategy development. Studies that have attempted to disambiguate genetic versus environmental contributions to this association have shown conflicting findings, though genomic-based studies have begun to establish the role of gene-environment interplay in the father absence/AAM literature. The purpose of this study was to replicate and extend prior genomic work using the Avon Longitudinal Study of Parents and Children (ALSPAC), a prospective longitudinal cohort study (N = 2,685), by 1) testing if an AAM polygenic score (PGS) could account for the father absence/AAM association, 2) replicating GxE research on lin-28 homolog B (*LIN28B*) variation and father absence, and 3) testing the GxE hypothesis using the PGS. Results showed that the PGS could not explain the father absence/AAM association and there was no interaction between father absence and the PGS. Findings using *LIN28B* largely replicated prior work that showed *LIN28B* variants predicted later AAM in father present girls, but this AAM-delaying effect was absent or reversed in father absent girls. Findings are discussed in terms genetic confounding, the unique biological role of *LIN28B*, and using PGSs for GxE tests.

Keywords

ALSPAC; menarche; father absence; PGS; *LIN28B*

Introduction

Age at menarche (AAM) is a key indicator of the timing of the later phases of pubertal maturation in girls. In part because of its ease of measurement and predictive value, it is also the most common measure in research evaluating the correlates and causes of pubertal timing and has been linked to several reproductive, health, and psychosocial outcomes. One key predictor of earlier AAM is father absence. Whether the association between father absence and earlier AAM is causal remains unclear, as well as the potential mechanisms that can explain the association. To advance our understanding of these mechanisms and potential confounds linking father absence and earlier AAM, genetically informed research is needed. The goal of the present study is to investigate two forms of gene-environment

interplay that may be important for understanding how father absence and AAM are linked: gene-environment correlation (*r*GE) and gene-environment interaction (GxE). In doing so, this study attempts to replicate and extend previous research that has used a molecular genetic approach to investigate the father absence/AAM association.

Father Absence and Age at Menarche

Because AAM is related to a host of physical health (Gong, Wu, Vogtmann, Lin, & Wang, 2013; Prentice & Viner, 2013) and psychosocial factors (Mrug et al., 2014; Skoog, Bayram Ozdemir, & Stattin, 2016), etiological research on AAM variability has been the subject of research over a number of years (Ellis, 2004). Much of this research has focused on putative environmental causes of individual differences in AAM including stress exposures during development, socioeconomic status, and family structure (Richardson, La Guardia, & Klay, 2018). Emergent from this research is a corpus of studies that link biological father absence to earlier AAM, largely based on reasoning from life history theory (Ellis, Figueredo, Brumbach, & Schlomer, 2009). Derived from work by Draper and Harpending (1982), Belsky, Steinberg, & Draper (1991) hypothesized that psychosocial stressors in and around the family, such as father absence and other stressors, experienced early in life serve as cues that regulate life history strategy development, including AAM. According to this perspective, father absence that occurs early in development (the first 5–7 years) signals that paternal investment is unnecessary for successful child-rearing in the likely future reproductive environment. A faster life history strategy, characterized by less stable pair bonds and less discriminate sexual behaviors, would be more evolutionarily favorable in this environment. Unique to Belsky et al. (1991) relative to Draper and Harpending was the proposition that father absence would also influence physical maturation as part of an organized life history strategy (see J. Belsky, 2012). Parallel developments by Ellis and colleagues (Ellis, 2004; Ellis, Schlomer, Tilley, & Butler, 2012; James, Ellis, Schlomer, & Garber, 2012) futher posited a special role of father absence in regulating sexual development via neurobiological mechanisms separate from other family-level and ecological stressors. Although some studies using these perspectives have failed to detect the association between father absence and AAM (e.g., in countries where malnutrition – potentially a stronger predictor of AAM – is relatively high; Kyweluk, Georgiev, Borja, Gettler, & Kuzawa, 2018; Sohn, 2017), meta-analytic findings of more than 30 studies show a reliable, though small, assocaition between father absence and earlier AAM (Webster, Graber, Gesselman, Crosier, & Orozco Schember, 2014).

Genetic Confouding in the Father Absence/AAM Association

Extant research utilizing life history theory as the predictive framework for the association between father absence and AAM has generally presumed that father absence exerts its effect on AAM via environmental means. This assumption was challenged based on AAM heritability (Rowe, 2002) and suggests the father absence/AAM association could arise from a genetic confound reflecting passive *r*GE. Subsequently, several attempts have been made to disambiguate genetic versus environmental explanations, with contrasting results. In one of the earliest studies, Mendle and colleges (2006) used a children-of-twins design to test the father absence genetic confounding hypothesis, and found that for youth of identical twin mothers, having a step-uncle was just as predictive of AAM as having a

step-father (for which father absence is a requirement). These findings suggest that genetic transmission or a shared environmental confound could explain the association between father absence and AAM. Tither & Ellis (2008) used a differential sibling exposure design, which controls for genetic and shared environment confounds by examining within-family associations, and found that the sister with greater exposure to father absence showed earlier AAM compared to her sister with less exposure, especially among fathers with high psychopathology. This finding suggests that father absence (and/or associated stressors) exerts a non-shared environmental effect on daughter's AAM. Findings across studies that use family-based research designs such as these have been critical for moving research on father absence and AAM further and suggests that there is likely some familial confounding that could be genetic in origin, but that familial or genetic confounding is unlikely to fully explain the association (Barbaro, Boutwell, Barnes, & Shackelford, 2017; D'Onofrio & Lahey, 2010). However, thus far, family based designs have only been used to explore genetic and environmental mechanisms as alternative explanations for the father absence/AAM association, thus overlooking aspects of gene-environment interplay important for development (Rutter et al., 1997). In addition, family-based designs operationalize genetic and familial confounds latently and are not directly measured. That is, because these designs infer genetic contributions in-lieu of examining DNA, these designs cannot inform the biological mechanisms through which father absence may exert an effect on AAM. Examining the genomic influences - through specific biological pathways and/or genomic aggregates - may help to explain associations between father absence and AAM (measured *r*GE; Jaffee & Price, 2007) and clarify if and how they are modified by structural environmental cues (GxE) critical for life history hypotheses.

As costs of genotyping and whole-genome assays have decreased, a growing body of literature has begun to incorporate measured genotypes into research on parenting and developmental outcomes. An initial study on father absence and AAM found support for the androgen receptor (*AR*) genetic confounding hypothesis when examining the GGC variant (Comings, Muhleman, Johnson, & MacMurray, 2002). This finding was not replicated, however, in a pair of independent follow-up studies (Jorm, Christensen, Rodgers, Jacomb, & Easteal, 2004; Schlomer, Murray, Yates, Hair, & Vandenbergh, 2019). Notably, a better-characterized CAG variant within *AR* was related to earlier AAM independent of father absence and moderated the association between father absence and women's sexual behaviors (Schlomer et al., 2019). Although not specific to father absence, two additional primary and replication studies found that variation in the estrogen receptor-α gene (*ESR1*) moderated the effect of family relationships on AAM (Hartman, Widaman, & Belsky, 2015; Manuck, Craig, Flory, Halder, & Ferrell, 2011). Following this line of research, Schlomer & Cho (2017) tested the hypothesis that father absence would moderate the association between lin-28 homolog B (*LIN28B*) variants (rs364663 (A/T) and rs314273 (G/T)) and AAM based on findings from genome-wide association studies (GWAS) that reliably linked *LIN28B* variation with later AAM and functional studies that showed overexpression of *lin28* genes was related to an overall phenotype that resembled a slow life history strategy. Results of GxE models showed that for both SNPs, daughters with the T/T genotype showed later AAM, but only if they were from father present households. A genetic index analysis

that combined the two SNPs into a composite showed that having at least one T/T genotype was related to later AAM among father present women but not father absent women.

These initial studies that integrate genomic information into the father absence/AAM literature reflect an important next step in the field by directly testing the interplay between genes and environments. The research to date, however, also has limitations. Perhaps most salient is that this work has primarily relied on candidate gene methods. Although often informed by biological and/or empirical evidence, candidate GxE (cGxE) studies, which focus on only one or a few variants, have proven difficult to replicate (Dick et al., 2015) and, at the extreme, some have argued that all cGxE research should be considered suspect given concerns about multiple testing (Duncan, Pollastri, & Smoller, 2014). Stemming from this issue, replication attempts of cGxE research is of critical importance (Culverhouse et al., 2018).

An additional critique of the cGxE approach is that single-gene analyses do not reflect the biological reality that complex phenotypes, such as AAM, are underlain by many genes of small effect (Fisher, 1918; Lohmueller, Pearce, Pike, Lander, & Hirschhorn, 2003). As a result, genomic research in psychological science has moved toward utilizing polygenic scores (PGS) that aggregate hundreds to thousands of SNPs into an overall composite (D. W. Belsky & Israel, 2014). PGSs demonstrate large phenotypic main effect associations relative to candidate polymorphisms and are a promising avenue for GxE research (Conley, Laidley, Boardman, & Domingue, 2016). Indeed, recent research on the genetics of pubertal development showed that a PGS consisting of variants linked to AAM (Day et al., 2017) was related to indicators of pubertal timing and tempo in both boys and girls (Horvath, Knopik, & Marceau, 2019). Further, the father absence/AAM genetic confounding hypothesis was recently tested by Gaydosh and colleagues (2018) who found that a similar AAM PGS was not related to father absence and that both father absence and the PGS uniquely predicted AAM. These results suggest that the association between father absence and AAM was not confounded by the PGS. These findings have not been replicated, however, and research examining possible GxE between father absence and PGSs on AAM has not been conducted.

Present Study

The present study is designed to replicate and extend recent advances in the literature on father absence and AAM using genomic approaches by utilizing data from a large prospective longitudinal cohort study. First, we sought to replicate Gaydosh et al. (2018) by using an AAM PGS to test a possible measured *r*GE confound between father absence and AAM. Following prior findings, we hypothesized that SNPs related to pubertal timing would correlate with AAM, be uncorrelated with exposure to father absence, and that father absence and the PGS would predict AAM while controlling for the other. If supported, the results would show no evidence of measured *r*GE - in other words, any genetic confounding of the association would not be driven by puberty-related genetic variation captured by the PGS. Next, we sought to replicate Schlomer & Cho (2017) by testing GxE using a targeted, candidate gene approach. Following Schlomer & Cho, we hypothesized first that T alleles in *LIN28B* rs364663 and rs314273 SNPs would be associated with later AAM and second

that this association would be found for father present girls and not for father absent girls. Finally, we extended work in this area by testing GxE using father absence and the PGS. Consistent with the *LIN28B* hypotheses, we expected that the later AAM associated with the PGS would be found for father present girls only.

Methods

Participants

Data for this study were drawn from the Avon Longitudinal Study of Parents and Children (ALSPAC). ALSPAC is a population-based, longitudinal cohort study that initially enrolled 14,541 pregnant women in the United Kingdom whose expected date of delivery was between April 1st 1991 and December 31st 1992 (Boyd et al., 2013; Fraser et al., 2013). 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. Data collection began when mothers were 8 weeks pregnant and includes 68 data collection time points between birth and child age 18, including mother and child reports. An additional 913 children were enrolled over the course of the study. Thus, the total sample available after age 7 years included 15,454 pregnancies, resulting in 15,589 fetuses of which 14,901 were alive at 1 year of age. Written informed consent was obtained from all study participants. 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). Approval for this study was obtained from the University at Albany, SUNY Institutional Review Boards.

Specific to the current study, the ALSPAC sample included $N = 7,262$ cases who self-identified as female. Of the female sample, genome-wide data were available on $N = 4,355$ participants. By design, the ALSPAC genomic sample is entirely European ancestry. Participants were excluded from the current study if they were missing data on AAM or father absence ($N = 1,488$, see Measures below). Additional cases were excluded from analysis due to missingness on other variables due to non-response or quality controls ($N = 182$). The final analytic sample for the current study was $N = 2,685$ female participants[1]. Comparisons between the analytic sample and the larger ALSPAC female sample revealed few differences. For example, average AAM was somewhat later among the included sample ($M = 151.38$ vs. 150.31 months; $r = .035$), and rates of father absence (14.6% vs. 22.2%, kappa = .078) and financial difficulties ($M_{difference} = .250$ SD's, $r = .12$) were somewhat lower in the included sample than in the excluded sample, although effect sizes for these differences were small.

Measures.

Please note that the ALSPAC study website contains details of all the data that are available through a fully searchable data dictionary and variable search tool: http://www.bristol.ac.uk/alspac/researchers/our-data/.

[1]Analyses using FIML estimation to account for missing data produced highly consistent results (see below) and are available by request from the first author.

Age at Menarche.—Daughter's age at menarche was prospectively assessed up to 9 times during approximately-annual assessments between the ages of 8 and 17 years of age. At each time point, surveys addressed if menstruation had begun and if so at what age. Age at menarche responses from the first available wave were used in the current study to minimize recall bias (Culpin et al., 2014; Gaydosh et al., 2018). Cases that were ±3 SD's away from the mean were removed from the analysis (N = 13). Mean AAM in months was 151.382 (SD = 13.802) and ranged from 114 to 192.5.

Father Absence.—ALSPAC mothers completed household composition surveys when children were 1 year 7 months, 2 years 7 months, 3 years 9 months, and 7 years old (Culpin et al., 2014). Mothers were asked if the current live-in father figure was the biological father of the target child. Response options were "Yes", "No", "No live-in father figure", and "Don't know." For each assessment, "Yes" responses were coded as biological father present. Cases where the live-in father figure was not the target child's biological father (N's = 35 to 121 across waves) or there was no live-in father figure (N's = 103 to 175) were coded as biological father absent. "Don't know" responses (< 0.5%) were dropped from the analysis. Children were coded as 1 = father absent prior to age 7 if their mother reported father absence at any of the 4 assessments (N = 392, 14.6%). The remaining children were coded 0 = father present (N = 2,293, 85.4%).

Genotyping.—DNA were collected on approximately 11,000 children at age 7 using blood, cell line, and mouthwash samples and genotyped using the Illumina HumanHap550 platform (Boyd et al., 2013; Pembrey, 2004). Raw genotype data were subjected to quality control measures and participants were excluded based on gender mismatch, minimal or excessive heterozygosity, >3% missingness, cryptic relatedness (IBD > .10), and non-European ancestry. Genomic data were imputed using 1000 Genomes phase 1, version 3. Following quality controls, a total of N = 4,355 female participants with genome-wide data were available.

Polygenic Score.—The PGS used in the current study was derived from findings by Day et al. (2017), who conducted the largest GWAS to date on AAM, comprising a sample of over N = 300,000 women from the ReproGen Consortium, 23andME, and the UK Biobank. GWAS results revealed 389 non-redundant genome-wide significant ($p < .05 \times 10^{-8}$) signals that collectively explained approximately 7.0% of the variance in AAM (see Day et al., 2017).

Supplemental information provided by Day et al. (2017) were used to create the PGS in the ALPAC data. Of the 389 independent SNPs, N = 372 were available in ALSPAC. Following quality control procedures that account for excess missingness (> 1%), low minor allele frequency (< 1%), and Hardy-Weinberg violations ($p < 5\times10^{-7}$), 339 variants remained. The 339 variants were then weighted by their respective effect sizes (Dudbridge, 2013) provided by Day et al. and averaged, resulting in a normally distributed PGS, which was then mean centered via standardization within the analytic sample (M = 0.00, SD = 1.0).

LIN28B. Following Schlomer & Cho (2017), rs364663 and rs314273 were analyzed as tag-SNPs that reflect variation in 26 total SNPs within a *LIN28B* LD block. Genotype

frequencies for rs364663 were A/A = 846 (31.5%), A/T = 1,350 (50.3%), T/T = 489 (18.2%); and for rs314273 G/G = 1,185 (44.1%), G/T = 1,213 (45.2%), T/T = 287 (10.7%). These frequencies were similar to those reported in Schlomer & Cho (2017). Both SNPs were in Hardy-Weinberg Equilibrium (χ^2's (1) < 1.52, *ns*). SNPs were dummy coded (Aliev, Latendresse, Bacanu, Neale, & Dick, 2014) such that the reference group for each SNP was for no copies of the T allele (A/A and G/G for rs364663 and rs314273, respectively) given prior research that suggests the T allele is associated with later AAM (see Schlomer & Cho, 2017).

Covariates.—Two covariates were used in all analyses. Prior work on the ALSPAC genetic sample showed little population structure (Jones et al., 2016; Okbay et al., 2016). However, to help adjust the results for potential additional undetected population structure, principal coordinates analysis (PCA) was conducted on the genome-wide data using the --pca command in plink (Chang et al., 2015). Eigenvalues examined in a scree plot (Cattell, 1966) leveled off following the first PC (*PC1*), which was used as a covariate in all analysis (D'Agostino & Russell, 2005). In addition, a *financial difficulties* measure was created to help control for possible socioeconomic confounds, which prior research suggests is related to both father absence and AAM (Boynton-Jarrett & Harville, 2012). When children were 8, 33, and 85 months of age mothers responded to a series of questions about difficulty affording food, clothing, rent, heat, and things for the child (1 = very difficult, 4 = not difficult). Scale reliabilities within each wave were good (α = .885 – .892) and the three assessments showed mean correlation of r = .524. The three assessments were reversed coded so that higher values equal more financial difficulties and then averaged. The resulting scale ranged from 0 to 3 (M = .484, SD = .522). The financial difficulties variable was mean centered via standardization for use in regression models.

Analysis Plan

To formally test the genetic confounding hypothesis using the AAM PGS, a path model was conducted in Mplus (Muthen & Muthen, 1998–2015) using father absence to predict AAM and the PGS predicting both. PC1 and financial difficulties (not shown in Figure 1) were also included as exogenous predictors of father absence, AAM, and the PGS. Following this model, GxE analyses were conducted in a series of four hierarchical regression models conducted in R (*lm* function; R Core Team, 2018). All models included PC1 and financial difficulties as covariates in the first block, then the main effects were added in the second, and finally the interactions in the third block. The first three models were designed to test the hypothesis that the association between *LIN28B* SNPs and later age at menarche would be found for father present girls and not for father absent girls, consistent with models reported by Schlomer & Cho (2017). Model 1 included father absence, dummy coded rs364663, and the product terms between father absence and the dummy variables. Model 2 was an identical model including rs314273 instead of rs364663. Model 3 included a genetic index combining the two SNPs: following Schlomer & Cho (2017) the rs364663 and rs314273 SNPs were combined into a genetic index such that 0 = No T/T genotypes from either SNP (N = 2,196) and 1 = at least 1 T/T genotype (N = 489). The fourth model tested the hypothesis that the later AAM associated with the PGS would be found for father present

girls only and was identical to the prior GxE models except that the AAM PGS was included as the genetic indicator.

Results

Preliminary Analyses

Table 1 shows the correlations among variables used in the current study. Age at menarche was correlated with father absence, rs314273, the PGS, and financial difficulties in the expected directions. rs364663 and rs314273 were strongly correlated with each other, reflecting their common LD block within *LIN28B*. Both were also modestly correlated with the PGS and a small correlation was detected between rs314273 and PC1. Father absence was related to more financial difficulties and not correlated with the *LIN28B* SNPs nor the PGS, which suggests any detected GxE is unlikely to be an artifact of *r*GE (Bakermans-Kranenburg & van IJzendoorn, 2015).

Father Absence and PGS Genetic Confounding

Standardized results from the path model (Figure 1) showed that father absence ($\beta = -.052$, $p < .05$) and the PGS ($\beta = .293$, $p < .05$) were related to AAM in predicted directions. There was no association between father absence and the PGS ($\beta = -.021$, *ns*). These findings suggest that the PGS cannot account for the father absence/AAM association and that the PGS and father absence have unique associations with AAM. Constraining the father absence → AAM path equal to the PGS → AAM path resulted in significant model misfit ($\Delta\chi^2(1) = 244.719$, $p < .05$) indicating the association between the PGS and AAM was stronger than the father absence association.

Model 1: Father Absence and rs364663

Regression results showed that father absence was related to approximately 2.3 months earlier AAM ($b = -2.289$, $p < .05$) while controlling for rs364663, financial difficulties, and PC1. No main effects were found for the difference between daughters with the rs364663 A/A genotype versus A/T ($b = .191$, ns) or between daughters with A/A versus T/T ($b = .893$, *ns*; model $R^2 = .006$). Adding the product terms to the model ($^{\Delta}R^2 = .004$) showed an interaction between father absence and A/A vs. A/T genotypes ($b = 4.205$, $p < .05$) and no interaction between father absence and A/A vs. T/T ($b = -1.218$, *ns*). An additional analysis using A/T as the reference group showed an interaction between father absence and A/T vs. T/T genotype ($^{\Delta}R^2 = .004$; $b = -5.423$, $p < .05$; no main effect of A/T vs. T/T was detected, $b = .703$, *ns*).

Follow-up simple effects tests were conducted by re-centering model variables to obtain conditional effects (Frazier, Tix, & Barron, 2004; see Figure 2a). Qualifying the interaction between father absence and A/T vs. T/T was a trend for later AAM between girls with A/T vs. T/T genotypes who were father present ($b = 1.507$, $p = .06$). Specifically, AAM was approximately 1.5 months later among T/T father present girls compared to A/T father present girls. Among father absent girls, the T/T genotype was related to nearly 4 months *earlier* AAM compared to A/T girls ($b = -3.916$, $p < .05$; see Figure 2a). Reflecting the interaction between father absence and A/A vs. A/T, AAM was 3.8 months earlier among

father absent girls who had the A/A genotype compared to the A/T ($b = 3.790$, $p < .05$), whereas the A/A vs. A/T slope for father present girls was flat ($b = -.415$, *ns*). This overall pattern of results suggests that having the rs364663 T/T genotype may be related to *later* age at menarche among father present women, consistent with Schlomer & Cho (2017), and that having either the A/A or T/T genotype is related to *earlier* age at menarche among father absent women.

Model 2: Father Absence and rs314273

Like Model 1, father absence was related to approximately 2.3 months earlier AAM ($b = -2.311$, $p < .05$) when controlling for financial hardship, PC1, and rs314273 (model $R^2 = .008$). No AAM difference was found for women with G/G vs G/T genotypes ($b = .640$, *ns*). Girls with the T/T genotype showed 2 months later AAM compared to G/T girls, irrespective of father absence status ($b = 2.011$, $p < .05$). In addition, an interaction was detected between father absence and G/G vs. T/T genotypes ($b = -4.680$, $p < .05$) as well as between father absence and G/T vs. T/T ($b = -6.345$, $p < .05$); no interaction was found between father absence and G/G vs. G/T ($b = 1.667$, *ns*; $^{\Delta}R^2 = .003$). Qualifying these interactions (see Figure 2b), AAM was nearly 2.5 months later among father present girls with the T/T genotype compared to G/T ($b = 2.467$, $p < .05$) or G/G ($b = 2.892$, $p < .05$) genotypes. Among father absent girls, there was a trend toward *earlier* AAM, by approximately 4 months, for T/T genotype girls compared to G/T ($b = -3.878$, $p = .07$) and no difference was found for G/G vs T/T ($b = -1.786$, ns). Similar to Model 1, these results indicate that the rs314273 T/T genotype is related to later AAM among father present girls and may be related to earlier AAM among father absent girls.

Model 3: *LIN28B* Genetic Index

Model 3 was identical to Models 1 and 2 except that the *LIN28B* index was used, which combines rs364663 and rs314273 into an overall score (0 = no T/T genotypes, 1 = at least 1 T/T genotype). Results are depicted in Figure 3a. After adjusting for the covariates, there was a main effect of father absence ($b = -2.291$, $p < .05$) and no main effect of the *LIN28B* index ($b = .776$, *ns*; $R^2 = .006$); the interaction was statistically significant ($b = -3.701$, $p = .05$; $^{\Delta}R^2 = .001$). Age at menarche for girls with at least one rs364663 or rs314273 T/T genotype was approximately 5 months later if they were father present compared to father absent ($b = 5.273$, $p < .05$); this difference, approximately 1.5 months, was marginal among girls without a T/T genotype ($b = -1.572$, $p = .07$), consistent with findings reported in Schlomer & Cho (2017).

Expanding on this analysis, the *LIN28B* genetic index was re-coded such that 0 = No T/T genotypes ($N = 2{,}196$), 1 = At least 1 T/T genotype ($N = 202$) and 2 = Two T/T genotypes ($N = 287$). Dummy coded analyses showed an interaction between father absence and 0 vs. 2 T/T genotypes ($b = -5.474$, $p < .05$; $^{\Delta}R^2 = .002$). Results are depicted in Figure 3b. Among father present girls, AAM was nearly 3 months later for girls with 2 T/T genotypes compared to having 1 ($b = 2.980$, $p < .05$) or 0 ($b = 2.645$, $p < .05$). Among father absent girls, more copies of the T/T genotype was not related to AAM (2 vs. 1: $b = -1.600$, *ns*; 2 vs. 0: $b = -2.830$, *ns*).

Model 4. GxE Analysis with the PGS

Model 4 was identical to prior models except the PGS was used as the genetic variable. Results showed father absence continued to predict earlier AAM when controlling for financial difficulties, PC1, and the PGS ($b = -2.023$, $p < .05$) and the AAM PGS predicted later AAM ($b = 4.025$, $p < .05$), as expected; for every one standard deviation increase in the PGS, AAM was 4 months later. The interaction between father absence and the AAM PGS was not detected ($b = .058$, *ns*; $\Delta R^2 < .001$). To test for possible non-linear associations, squared and cubed versions of the PGS were evaluated. No main effects were found for the PGS beyond the linear association (squared: $b = -.062$, *ns*; cubed: $b = .014$, *ns*) nor did the squared or cubed PGS's interact with father absence (b's = .283 and .143, *ns*, respectively).

Sensitivity Analyses.

The dummy coded analyses reported above for rs364663 and rs314273 allow potential non-linear relationships between SNPs and AAM to be evaluated (Aliev et al., 2014). In GWAS, however, the SNP-phenotype association is presumed to be linear. As a follow-up test, analyses were conducted to examine the linear association between the *LIN28B* SNPs and AAM as well as their interactions with father absence. While controlling for covariates, results showed that rs314273 was related to later AAM by nearly 1 month per T-allele, irrespective of father absence ($b = .879$, $p < .05$). No linear association was found for rs364663 ($b = .411$, *ns*). Interactions for both SNPs with father absence were non-significant (rs364663: $b = .009$, *ns*; rs314273: $b = -1.186$, *ns*).

An additional consideration is whether the father absence/AAM association is driven by the presence of a stepfather or other non-biologically related adult male (e.g., the mother's live-in boyfriend), which can be a consequence of father absence (Ellis & Garber, 2000; Mendle et al., 2006). To test this possibility, an ANCOVA was conducted using a nominal family status variable with three levels: father present ($N = 2{,}293$), biological father absent ($N = 220$), and non-biological adult male present ($N = 172$). PC1 and financial difficulties were included as covariates. Results showed a null interaction between the *LIN28B* index and family status in the omnibus model ($F(2,2677) = 2.279$, *ns*). A Helmert contrast showed that AAM was later among father present girls ($M = 152.69$) compared to biological father absent/non-biological adult male present girls ($M = 149.32$; $t = 3.47$, $p < .05$), reflecting prior results. No difference was found for the contrast between biological father absent ($M = 148.72$) and non-biological adult male present girls ($M = 149.911$; $t = 0.67$, *ns*). This result suggests that the reported father absence findings are not driven by stepfather/non-biologically related adult male presence.

It is also possible that chronicity of father absence during early development may be driving the dichotomous father present/absent effect reported in this study. To evaluate this possibility, a father absence "dosage" variable was created that quantified the number of waves a participant reported being father absent (0 to 4) and was used in place of the dichotomous variable in the models described above. To ensure the accuracy of the dosage variable, analyses were limited to cases with complete father absence data at all 4 waves ($N = 2{,}075$). Results of main effects models showed that father absence dosage was not significantly related to age at menarche when covariates and the respective genetic variables

were included. These non-significant findings are likely attributable to smaller standardized effect sizes observed for the dosage (β's ~ −.033, *ns*) compared to the dichotomous (β's ~.060, p's < .01) father absence variable as well as the larger standard errors that resulted from the sample reduction. Given the lack of main effects of father absence dosage, moderation analyses were not further pursued.

In addition, because financial difficulties could be theoretically conceptualized as either a mechanism by which father absence could affect age at menarche (e.g., James et al., 2012) or a component of a larger environmental stress factor that also includes father absence (e.g., Belsky et al., 1991), using financial difficulties and a control variable could potentially remove important variance in AAM. To evaluate this possibility, models were conducted with financial difficulties uncontrolled. Results were highly similar; the largest change in the beta coefficients was .09 months and the pattern of significance was the same. We also examined the components of financial difficulties separately to provide further context for the role of financial difficulties. To do so, each item of the composite was separately averaged across the three assessment waves (child age 8 months, 32 months, and 85, months, see Methods section) resulting in 5 indicators of financial difficulties (i.e., difficulty affording: 1) food, 2) clothing, 3) heat, 4) rent, and 5) items for the child). These five indicators were included in analysis models described above in place of the overall financial difficulties composite. Across all models, none of the individual indictors were significantly related to AAM in the main effects analyses (b's < .80, *ns*), including the difficulty affording food indicator (b's < .40, *ns*). Using the individual items as covariates in place of the composite did not affect the pattern of results observed in the interaction models.

Last, the *LIN28B* findings reported above could be interpreted as evidence for differences in susceptibility to father absence. Differential susceptibility theory (J. Belsky & Pluess, 2009; Ellis, Boyce, Belsky, Bakermans-Kranenburg, & van IJzendoorn, 2011) posits that intrapersonal characteristics that would otherwise be conceptualized as risk factors within adverse environments may be related to more positive outcomes in more supportive environments. To test the hypothesis that *LIN28B* conveys differential sensitivity to father absence, a reparametrized regression model (Widaman et al., 2012) was conducted to test the differential susceptibility hypothesis for the interaction between the *LIN28B* index (coded 0/1) and father absence. The reparametrized regression model includes the cross-over point of the conditional slopes as an estimated model parameter, which permits 95% confidence intervals to be calculated. The cross-over parameter and its confidence interval can be used to infer the likely form of the interaction in the population. Disordinal interactions are consistent with differential susceptibility theory whereas ordinal interactions indicate diathesis stress or, alternatively, vantage sensitivity, which posits sensitivity to ostensibly positive environments, such as father presence (see Pluess & Belsky, 2013; Widaman et al., 2012). In the current analysis, the cross-over point reflects the estimated point where the conditional slopes for the association between the *LIN28B* index and AAM intersect (one slope for father present and one for father absent). Results of the reparametrized regression showed the cross-over point was estimated at −.516 with a 95% confidence interval that ranged from −1.336 to .304 (Figure 4). Because the cross-over is estimated beyond the range of the predictor (*LIN28B* index, 0/1) the interaction form is ordinal, which we interpret as consistent with vantage sensitivity given the later AAM associated with LIN28B was present

for father present girls and not for father absence girls. It is important to note, however, that because the upper 95% CI falls within the range of the *LIN28B* index, a disordinal interaction in the population cannot be ruled out. Notably, the ordinal interaction is also consistent with findings from Schlomer & Cho (2017).

Discussion

The purpose of this study was to replicate and extend recent research on the father absence/AAM association that used genomic approaches. The genetic confounding hypothesis was tested first following prior research that suggested a similar AAM PGS did not confound the association (Gaydosh et al., 2018). Contrary to the genetic confounding hypothesis, independent associations were found in the current study for both father absence and the PGS in a path model and the PGS was not correlated with father absence. In line with the prior study, these findings indicate that father absence is not confounded by the AAM PGS evaluated in this study. It is important to point out, however, that the PGS reflects a narrow genetic control given it explains approximately 7% of the additive genetic contribution to AAM (Day et al., 2017). Any remaining association between father absence and AAM after controlling for the PGS could still be accounted for by additional unmeasured additive genetic variance, epistasis, epigenetics, or a number of other biological and/or non-biological factors. Although the current and prior studies cannot resolve the ongoing debate about whether father absence constitutes an environmental pressure that regulates pubertal development posited by life history theory, these studies indicate that the genetic variance captured by the PGS cannot account for the father absence/AAM association. In addition, and perhaps more importantly, the current results are informative regarding the relative magnitude of the genetic and putative environmental effects captured by the PGS and father absence, respectively. As evidenced by the model misfit when the father absence → AAM and PGS → AAM paths were constrained equal, the PGS was much more strongly related to AAM relative to father absence. Although father absence *may* exert a small environmental effect on AAM, the additive genetic contribution captured by the PGS explained a much larger portion of the variance. An implication of this observation is that the small effect of father absence may be more likely to go undetected in genetically informed studies than the other way around, particularly given other factors known to influence the father absence/AAM association (e.g., SES, ethnicity). Thus, failure to detect an effect of father absence on AAM, particularly in genomic study, may not necessarily indicate a lack of association but may reflect the inability of a given sample to detect what amounts to a relatively small main effect of father absence, which will be more subject to study variances (e.g., sampling, measurement) compared to the relatively larger genomic associations. That is not to say that father absence or other family environmental effects on AAM should be dismissed as relatively unimportant. For example, half of menstrual cycles are anovulatory for approximately 1 year among girls whose menarche occurs prior to age 12. Among girls whose menarche occurs at age 13 or older, the anovulatory period lasts for 4.5 years. Small effect sizes can translate into substantial implications for reproductive outcomes (Ellis, 2004).

Further, it may come as a surprise find such a relatively large genetic-phenotype association given the effect size discrepancies often found between molecular genetic studies and

family-based genetic designs (e.g., twin studies; Manolio et al., 2009). Maximizing phenotypic variance explained is precisely the goal, however, of PGS methods (Wray et al., 2014). By scanning the entire genome via GWAS for statistically relevant signals and then aggregating them, PGSs are designed to explain substantially more variance in a phenotype than any one signal could on its own. This approach reflects an important and necessary advance in genomic research, particularly as it applies to behavioral and psychological science. Nonetheless, the systematic scan and aggregation of phenotypically relevant factors on only one side of the gene-environment equation could lead to underestimating the importance of the environment, broadly defined, when comparisons are made between PGSs and single environmental factors that on their own may explain only tiny amounts of variation. This imbalance - reflecting advances in genomic measurement unparalleled in environmental measurement - has not gone unnoticed in epidemiological research of disease etiology and calls have been made to measure the “exposome” in a manner inspired by genotyping arrays (Wild, 2005).

Relevant to the current study, father absence associations are commonly detected and are reliably small (Webster et al., 2014) and, arguments about the environmentality of father absence and related factors aside, maximal environmental prediction of AAM could be achieved in a similar manner as PGSs via systematic scans of the likely many environments of small effect that influence AAM. In this framework, father absence is only one of potentially many environments that could be aggregated into a poly-environment score that could maximize phenotypic prediction in a similar manner as PGSs. Indeed, this approach of aggregating environmental effects is theoretically consistent with Belsky et al.’s (1991) original proposition, which did not privilege father absence above other psychosocial stressors but posited that multiple stress-related factors would influence AAM. Environment-wide association studies that systematically evaluate the “exposome” in a manner conceptually similar to GWAS have been successfully applied in primary and replication research on Type 2 diabetes (Hall et al., 2014; Patel, Bhattacharya, & Butte, 2010) and offer a promising avenue for studying environmental influences on reproductive health (Buck Louis, Yeung, Sundaram, Laughon, & Zhang, 2013) and early life exposures on developmental outcomes (Vrijheid et al., 2014). Though beyond the scope of the current study, additional research should examine whether greater environmental prediction, and potential GxE, would be observed when aggregating sets of environmental factors hypothesized to influence AAM, such as father absence, financial difficulties, and others.

A second goal of this study was to replicate findings from Schlomer & Cho (2017). The initial evidence from Schlomer and Cho suggested that the later AAM linked to *LIN28B* SNPs rs364663 and rs314273 is specific to father present girls. The potential reliability of this initial evidence was undermined by its limitations, which included a relatively small sample of college students (N = 300), retrospective self-reports of AAM, and father absence operationalized as parent’s current marital status. In addition, because the study used a candidate gene approach to evaluate GxE, the prior probability that detected associations would be meaningful could be considered low (Dick et al., 2015; Ioannidis, 2005). The current study improves upon these limitations via a larger community-based sample (N = 2,685), contemporaneously measured AAM, and a father absence variable based on repeated measures of household composition drawn from mother reports. With these methodological

improvements in mind, the results of the current replication attempt were largely consistent with the prior work: The association between *LIN28B* SNPs and later age at menarche typically found in GWAS was specific to father present girls. Evidence for this pattern was found both in the individual SNP analyses and, most strongly, in the genetic index analyses, consistent with Schlomer & Cho. In addition, the ordinal interaction pattern found in this study, supporting a vantage sensitivity interpretation of the genetic index interaction, was also replicated. The larger sample size of the current study also afforded an additional sensitivity analysis that showed these findings were not driven by stepfather/non-biologically related adult male presence. Taken together, primary and replication studies find evidence to support a *LIN28B*-by-father absence GxE on variation in AAM such that *LIN28B* is linked to later AAM among father present girls and not among father absent girls.

The hypothesis that father absence would moderate the association between *LIN28B* and AAM in Schlomer & Cho (2017) was based on prior father absence research using life history theory, which posits early developmental experiences in and around the family provide evolutionarily relevant cues that regulate the development of life history strategies. Interestingly, father absence - hypothesized as one of these cues - only modified the effect of *LIN28B*, and not the PGS, on AAM. This finding suggests that father absence has a discriminating effect on the pathway(s) that *LIN28B* operates. *LIN28B* was chosen for evaluation based on its reliable association with AAM across genome-wide studies (Day et al., 2017), studies that showed *LIN28B* is highly evolutionarily conserved (Thornton & Gregory, 2012), non-human animal research that suggests over expression of a *lin28b* homologue results in a phenotype that resembles a slow life history strategy (i.e., larger growth and later reproduction; Zhu et al., 2010), and functional studies that link *LIN28B* with gene expression via the lin28/let-7 pathway that is key for metabolic regulation (Zhu et al., 2010). Nonetheless, the precise biological mechanism that would directly link father absence to *LIN28B* regulation via GxE is unknown.

Recent research suggests *LIN28B* has a more fundamental role in regulating developmental processes in humans and other animals beyond metabolic regulation (Leinonen, Chen, Tukiainen, Panula, & Widen, 2019). *LIN28* genes are found in all metazoan organisms, are highly pleiotropic, and are critical for stem cell differentiation and other developmental processes across disparate species (Moss & Tang, 2003; Pasquinelli et al., 2000). Functionally, lin28 proteins produced by *LIN28* genes bind to let-7 micro-RNA (miRNA) precursors preventing their biogenesis. miRNA's like let-7 regulate the expression of other genes post-transcriptionally by binding to messenger RNA (mRNA) and interfering with translation. Accumulating research suggests the lin28/let-7 axis plays a role in hypothalamic-pituitary- gonadal (HPG) axis regulation relevant for pubertal timing. More specifically, research on rodents has found that *lin28b* and its homolog *lin28a* are highly expressed in hypothalamic, pituitary, and gonadal tissues (see Corre et al., 2016). *lin28b* mRNA, indicating *lin28b* expression, is abundant in the hypothalamus of neonatal mice and declines with age, reaching a low point prior to puberty. At the same time, levels of let-7 are also low during the neonatal period and increase during development (Sangiao-Alvarellos et al., 2013). This research suggests that reduced expression of *lin28b* with age may lead to higher let-7 availability in the hypothalamus, which could repress the expression of other

genes that serve as a “break” for hypothalamic onset of puberty (Aylwin, Toro, Shirtcliff, & Lomniczi, 2019; Plant, 2001).

Research on the lin28/let-7 axis in humans is sparse, however, though *LIN28B* mRNA expression in human hypothalamic and pituitary tissue is related to the *LIN28B* rs7759938 C-allele, which is in high linkage disequilibrium with the rs314273 T-allele (R^2 = 1.0; ldlink.nci.nih.gov), a focal SNP in the current study (Leinonen et al., 2019). Thus, it is hypothesized, though not yet tested, that the T alleles of the *LIN28B* SNPs tested herein would be similarly associated with increased *LIN28B* expression during development and that early father absence exposure (and/or related factors) would be related to reduced expression. Research on early environmental exposures in and around the family, such as father absence and associated stressors, also points to hypothalamic regulation as a possible mechanism for earlier puberty via HPA axis regulation (Ellis, Essex, & Boyce, 2005; Essex, Klein, Cho, & Kalin, 2002; Mackrell et al., 2014). Few studies, however, have simultaneously tested the links between early stress, HPA regulation, and pubertal development (J. Belsky, Ruttle, Boyce, Armstrong, & Essex, 2015) and no studies have looked specifically at father absence and AAM. Additional research is needed on possible environmental regulation - by not only father absence but also broader indices of environmental factors key to life history strategy development - of the lin28/let-7 pathway in humans and its implications for hypothalamic onset of puberty.

Also notable are two additional interesting and somewhat unexpected patterns that emerged in the current study. First was the earlier AAM associated with the rs364663 A/A genotype among father absent girls. It is unclear what might be driving this finding but an admittedly post-hoc explanation could be heterozygote advantage or so-called “hybrid vigor” (Birchler, Yao, & Chudalayandi, 2006). Hybrid vigor refers to the evolutionary condition where a hybrid organism has higher fitness by virtue of heterozygosity compared to homozygotes of the same species. A classic example is malaria resistance found in humans heterozygous for sickle-cell anemia genetic risk (Tchuenche, 2007). Although purely speculative, perhaps heterozygosity at rs364663 might convey resistance to environmental factors that lead to earlier AAM, such as father absence. Additional research, including further replication of the current findings, is needed to explore this and other possible explanations.

A second interesting observation of these results is the possibility that *LIN28B* genotypes that show later AAM among father present girls may also be related to *earlier* AAM among father absent girls, in an ostensibly “for better and for worse” manner (J. Belsky, Bakermans-Kranenburg, & van IJzendoorn, 2007). The findings of the present study most consistently support the proposition that *LIN28B* genotypes are related to later AAM at menarche in father present households, however the overall pattern of results (as depicted in Figures 2 and 3) are at least suggestive that *LIN28B* variants may be related to earlier AAM in father absent households. Indeed, given that relatively smaller number of participants who reported being father absent, failure to statistically detect this pattern may be the result of relatively low power within this group rather than true lack of association. Additional research with even larger samples will be needed to further evaluate this possibility.

Last, the third goal of this study was to expand prior GxE research on father absence and AAM by testing moderation using an empirically derived PGS score. Main effect results were consistent with path model findings, though no GxE was detected. Based on these findings, a critical question is why GxE was detected for *LIN28B* and not for the PGS? One explanation is that the *LIN28B* GxE is a false positive. We view this interpretation as unlikely given there is now primary and replication evidence that suggests otherwise. In addition, the mechanistic and evolutionary research that link *LIN28B* and father absence as well as qualities unique to this research question, such as the strong theoretical foundation and measurement features (see Schlomer & Cho, 2017 for a discussion), suggest that the *LIN28B* interaction with father absence is reliable. It is possible that *LIN28B* and father absence influence AAM through a common biological pathway, such as HPG/HPA regulation, and that the PGS, which reflects a broader range of biological processes related to menarche, is a more general measure that is effective at detecting phenotypic main effects but perhaps less well-suited for GxE when specific pathways transact with specific environments.

Another possibility is that there may be non-linear associations between AAM and SNPs that make up the PGS. Polygenic scores assume additive linearity whereas the current data and the data from Schlomer & Cho (2017) as well as from others (Aliev et al., 2014) indicate that linear associations between SNPs and phenotypes may not always be the most precise model. To the degree that linearity assumptions are untrue, PGSs will underestimate phenotypic associations and additionally undermine moderation if the interaction is non-linear as well, as was the case with the *LIN28B* findings. Indeed, at least some of the missing heritability issue might be due to non-linear SNP associations. To improve PGS prediction, non-linearity would need to be addressed at the SNP level prior to aggregation, though there are no established methods for doing so. We attempted to evaluate PGS non-linearity using polynomial interactions in this study, but this approach does not address the potential issue at the SNP-level.

Last, it is currently on open question about whether we should expect PGSs to be moderated by environmental contexts at all. There has been a justified push for some time in psychological research to move away from candidate gene studies in favor of more polygenic approaches given PGSs have greater predictive power and reflect the biological reality that many genes of small effect underlay complex phenotypes. While PGSs are more powerful for detecting phenotypic main effects, there are some concerns about evaluating GxE. For instance, creating PGSs using methods that do not require stringent genome-wide significance thresholds, such as the Purcell method (International Schizophrenia Consortium, 2009), could lead to more error prone genetic variables given the goal of this approach is to find the signal-to-noise ratio that maximizes main effect prediction (Evans, Visscher, & Wray, 2009). Noisy genetic variables could lead to capitalizing on chance and spurious GxE, an issue that has been brought up many times in the candidate gene literature (Dick et al., 2015). Another issue of PGSs when it comes to GxE is that if markers within a score are moderated in different directions, an overall null effect could be detected (D. W. Belsky & Israel, 2014). It is possible that many GxE's using a PGS approach may simply not be found despite widespread acknowledgement that genes and environments work in tandem. Finally, by design PGSs only include SNPs that show a strong enough association

in a given sample size to meet a given significance threshold. This also means that the main effect of these SNPs are able to penetrate through considerable environmental variability represented by the hundreds of thousands or more participants that typically make up GWAS samples. Given these properties, it may be the case that we should not expect to find GxE weth SNP sets that are arguably unlikely to demonstrate conditional associations. From a differential susceptibility perspective, however, PGSs that capture developmental plasticity will be more likely to show environmental moderation. Moving forward, these issues will need to be addressed along with PGS replication research to avoid these studies leading to another body of literature that appears to be largely unreplicable.

Limitations and Conclusions

The findings of the current study should be evaluated in light of its limitations. Perhaps most salient is the homogenous sample, consisting entirely of European ancestry individuals. The generalizability of the current findings may be limited to this population and thus additional research with more diverse samples is needed. Nonetheless, the homogenous ancestry of the current sample could also be considered a strength since population stratification is unlikely to explain these results. Another possible limitation was that the PGS was constructed using weights from a meta-analytic GWAS (Day et al., 2017) – a discovery sample that included the present sample in it. Although ALSPAC only contributed 3% of the participants in the discovery GWAS, this may inflate the effect of the PGS reported here. Sensitivity analyses in a similar study with a different pubertal outcome (timing and tempo of secondary sex characteristics; Horvath et al., 2019) suggested that this bias was negligible.

The present study replicated and extended recent genomic research on the association of father absence and AAM. We confirmed that puberty-related genes, in terms of a PGS, do not confound father absence/AAM associations. To the extent that there is some genetic confounding of the association, other biological mechanisms must be involved. In addition, we replicated the effects of a specific gene, *LIN28B*, which is frequently implicated in the timing of pubertal development and its effect is modified by father absence in that having a T allele (presumably resulting in greater *LIN28B* expression hypothalamic and pituitary tissues) is more predictive of later puberty among father present girls, but this puberty-delaying effect is absent or possibly even reversed (e.g., related to earlier puberty) in father absent girls. However, this finding did not extend to the PGS, which has two possible implications: 1) the additive construction of the PGS from multiple SNPs is inappropriate to understanding the effects of father absence on AAM and 2) *LIN28B* has a unique mechanistic role in the timing of puberty that is particularly open to early environmental influences, consistent with life history theory. Human studies on the mechanisms of *LIN28B* for puberty are scarce, though there is some evidence on which to hypothesize that the effects found here could be mediated through increased *LIN28B* expression in the hypothalamus and pituitary gland. Thus, *LIN28B* and the lin28/let-7 pathway particularly as expressed in the hypothalamus and pituitary may be the best target for future research aiming to understand the precise biological mechanisms that would link father absence to AAM via GxE.

Acknowledgements:

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

Funding: The UK Medical Research Council and Wellcome (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and GLS and KM will serve as guarantors for the contents of this paper.

A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf) GWAS data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe.

This research was additionally supported by K01DA039288 to KM.

References

Aliev F, Latendresse SJ, Bacanu SA, Neale MC, & Dick DM (2014). Testing for measured gene-environment interaction: problems with the use of cross-product terms and a regression model reparameterization solution. Behavior Genetics, 44(2), 165–181. doi:10.1007/s10519-014-9642-1 [PubMed: 24531874]

Aylwin CF, Toro CA, Shirtcliff E, & Lomniczi A (2019). Emerging genetic and epigenetic mechanisms underlying pubertal maturation in adolescence. Journal of Research on Adolescence, 29(1), 54–79. doi:10.1111/jora.12385 [PubMed: 30869843]

Bakermans-Kranenburg MJ, & van IJzendoorn MH (2015). The hidden efficacy of interventions: genexenvironment experiments from a differential susceptibility perspective. Annual Review of Psychology, 66, 381–409. doi:10.1146/annurev-psych-010814-015407

Barbaro N, Boutwell BB, Barnes JC, & Shackelford TK (2017). Genetic confounding of the relationship between father absence and age at menarche. Evolution and Human Behavior, 38(3), 357–365. doi:10.1016/j.evolhumbehav.2016.11.007

Belsky DW, & Israel S (2014). Integrating genetics and social science: genetic risk scores. Biodemography and Social Biology, 60(2), 137–155. doi:10.1080/19485565.2014.946591 [PubMed: 25343363]

Belsky J (2012). The development of human reproductive strategies. Current Directions in Psychological Science, 21(5), 310–316. doi:10.1177/0963721412453588

Belsky J, Bakermans-Kranenburg MJ, & van IJzendoorn MH (2007). For better and for worse: Differential susceptibility to environmental influences. Current Directions in Psychological Science, 16(6), 300–304. doi:10.1111/j.1467-8721.2007.00525.x

Belsky J, & Pluess M (2009). Beyond diathesis stress: Differential susceptibility to envionmental influences. Psychological Bulletin, 135(6), 885–908. [PubMed: 19883141]

Belsky J, Ruttle PL, Boyce WT, Armstrong JM, & Essex MJ (2015). Early adversity, elevated stress physiology, accelerated sexual maturation, and poor health in females. Developmental Psychology, 51(6), 816–822. doi:10.1037/dev0000017 [PubMed: 25915592]

Belsky J, Steinberg L, & Draper P (1991). Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. Child Development, 62(4), 647–670. doi:10.1111/j.1467-8624.1991.tb01558.x [PubMed: 1935336]

Birchler JA, Yao H, & Chudalayandi S (2006). Unraveling the genetic basis of hybrid vigor. Proceedings of the National Academy of Sciences of the United States of America, 103(35), 12957–12958. doi:10.1073/pnas.0605627103 [PubMed: 16938847]

Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, … Davey Smith G (2013). Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. International Journal of Epidemiology, 42(1), 111–127. doi:10.1093/ije/dys064 [PubMed: 22507743]

Boynton-Jarrett R, & Harville EW (2012). A prospective study of childhood social hardships and age at menarche. Annals of Epidemiology, 22(10), 731–737. doi:10.1016/j.annepidem.2012.08.005 [PubMed: 22959664]

Buck Louis GM, Yeung E, Sundaram R, Laughon SK, & Zhang C (2013). The exposome--exciting opportunities for discoveries in reproductive and perinatal epidemiology. Paediatric and Perinatal Epidemiology, 27(3), 229–236. doi:10.1111/ppe.12040 [PubMed: 23574410]

Cattell RB (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1(2), 245–276. [PubMed: 26828106]

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, & Lee JJ (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience, 4, 7. doi:10.1186/s13742-015-0047-8 [PubMed: 25722852]

Comings DE, Muhleman D, Johnson JP, & MacMurray JP (2002). Parent-daughter transmission of the androgen receptor gene as an explaination of the effect of father absence on age of menarche. Child Development, 73(4), 1046–1051. [PubMed: 12146732]

Conley D, Laidley TM, Boardman JD, & Domingue BW (2016). Changing polygenic penetrance on phenotypes in the 20(th) century among adults in the US population. Scientific Reports, 6, 30348. doi:10.1038/srep30348 [PubMed: 27456657]

Consortium IS (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature, 460(7256), 748–752. doi:10.1038/nature08185 [PubMed: 19571811]

Corre C, Shinoda G, Zhu H, Cousminer DL, Crossman C, Bellissimo C, … Palmert MR (2016). Sex-specific regulation of weight and puberty by the Lin28/let-7 axis. Journal of Endocrinology, 228(3), 179–191. doi:10.1530/JOE-15-0360

Culpin I, Heron J, Araya R, Melotti R, Lewis G, & Joinson C (2014). Father absence and timing of menarche in adolescent girls from a UK cohort: the mediating role of maternal depression and major financial problems. Journal of Adolescence, 37(3), 291–301. doi:10.1016/j.adolescence.2014.02.003 [PubMed: 24636689]

Culverhouse RC, Saccone NL, Horton AC, Ma Y, Anstey KJ, Banaschewski T, … Bierut LJ (2018). Collaborative meta-analysis finds no evidence of a strong interaction between stress and 5-HTTLPR genotype contributing to the development of depression. Molecular Psychiatry, 23(1), 133–142. doi:10.1038/mp.2017.44 [PubMed: 28373689]

D'Agostino RB, & Russell HK (2005). Scree Test. In Armitage P & Colton T (Eds.), Encyclopedia of Biostatistics (2nd ed.). Chichester, England: John Wiley & Sons.

D'Onofrio BM, & Lahey BB (2010). Biosocial Influences on the Family: A Decade Review. Journal of Marriage and Family, 72(3), 762–782. doi:10.1111/j.1741-3737.2010.00729.x [PubMed: 24009400]

Day FR, Thompson DJ, Helgason H, Chasman DI, Finucane H, Sulem P, … Perry JRB (2017). Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. Nature Genetics, 49(6), 834–841. doi:10.1038/ng.3841 [PubMed: 28436984]

Dick DM, Agrawal A, Keller MC, Adkins A, Aliev F, Monroe S, … Sher KJ (2015). Candidate gene-environment interaction research: reflections and recommendations. Perspectives on Psychological Science, 10(1), 37–59. doi:10.1177/1745691614556682 [PubMed: 25620996]

Draper P, & Harpending H (1982). Father absence and reproductive strategy: An evolutionary perspective. Journal Of Anthropological Research, 38(3), 255–273.

Dudbridge F (2013). Power and predictive accuracy of polygenic risk scores. PLoS Genetics, 9(3), e1003348. doi:10.1371/journal.pgen.1003348 [PubMed: 23555274]

Duncan LE, Pollastri AR, & Smoller JW (2014). Mind the gap: why many geneticists and psychological scientists have discrepant views about gene-environment interaction (GxE) research. American Psychologist, 69(3), 249–268. doi:10.1037/a0036320

Ellis BJ (2004). Timing of pubertal maturation in girls: an integrated life history approach. Psychological Bulletin, 130(6), 920–958. doi:10.1037/0033-2909.130.6.920 [PubMed: 15535743]

Ellis BJ, Boyce WT, Belsky J, Bakermans-Kranenburg MJ, & van IJzendoorn MH (2011). Differential susceptibility to the environment: an evolutionary--neurodevelopmental theory. Development and Psychopathology, 23(1), 7–28. doi:10.1017/S0954579410000611 [PubMed: 21262036]

Ellis BJ, Essex MJ, & Boyce WT (2005). Biological sensitivity to context: II. Emperical explorations of an evolutionary-developmental theory. Development and Psychopathology, 17, 303–328. doi:10.10170S0954579405050157 [PubMed: 16761547]

Ellis BJ, Figueredo AJ, Brumbach BH, & Schlomer GL (2009). Fundamental dimensions of environmental risk : The Impact of harsh versus unpredictable environments on the evolution and development of life history strategies. Human Nature, 20(2), 204–268. doi:10.1007/s12110-009-9063-7 [PubMed: 25526958]

Ellis BJ, & Garber J (2000). Psychosocial antecedents of variation in girls' pubertal timing: Maternal depression, stepfather presence, and maritial and family stress. Child Development, 71(2), 485–501. [PubMed: 10834479]

Ellis BJ, Schlomer GL, Tilley EH, & Butler EA (2012). Impact of fathers on risky sexual behavior in daughters: a genetically and environmentally controlled sibling study. Development and Psychopathology, 24(1), 317–332. doi:10.1017/S095457941100085X [PubMed: 22293012]

Essex MJ, Klein MH, Cho E, & Kalin NH (2002). Maternal stress beginning in infancy may sensitize children to later stress exposure: Effects on cortisol and behavior. Biological Psychiatry, 52(8), 776–784. doi:Pii S0006-3223(02)01553-6 Doi 10.1016/S0006-3223(02)01553-6 [PubMed: 12372649]

Evans DM, Visscher PM, & Wray NR (2009). Harnessing the information contained within genome-wide association studies to improve individual prediction of complex disease risk. Human Molecular Genetics, 18(18), 3525–3531. doi:10.1093/hmg/ddp295 [PubMed: 19553258]

Fisher RA (1918). The correlation between relatives on the supposition of mendelian inheritance. Transactions of the Royal Society of Edinbergh, 53, 399–433.

Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, … Lawlor DA (2013). Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. International Journal of Epidemiology, 42(1), 97–110. doi:10.1093/ije/dys066 [PubMed: 22507742]

Frazier PA, Tix AP, & Barron KE (2004). Testing moderator and mediator effects in counseling psychology research. Journal of Counseling Psychology, 51(1), 115–134. doi:10.1037/0022-0167.51.1.115

Gaydosh L, Belsky DW, Domingue BW, Boardman JD, & Harris KM (2018). Father absence and accelerated reproductive development in non-hispanic white women in the United States. Demography, 55(4), 1245–1267. doi:10.1007/s13524-018-0696-1 [PubMed: 29978338]

Gong TT, Wu QJ, Vogtmann E, Lin B, & Wang YL (2013). Age at menarche and risk of ovarian cancer: a meta-analysis of epidemiological studies. International Journal of Cancer, 132(12), 2894–2900. doi:10.1002/ijc.27952 [PubMed: 23175139]

Hall MA, Dudek SM, Goodloe R, Crawford DC, Pendergrass SA, Peissig P, … Ritchie MD (2014). Environment-wide association study (EWAS) for type-2 diabetes in the Marshfield Personalized Medicine Research Project Biobank. Paper presented at the Pacific Symposium on Biocomputing.

Hartman S, Widaman KF, & Belsky J (2015). Genetic moderation of effects of maternal sensitivity on girl's age of menarche: Replication of the Manuck et al. study. Development and Psychopathology, 27(3), 747–756. doi:10.1017/S0954579414000856 [PubMed: 25195863]

Horvath G, Knopik VS, & Marceau K (2019). Polygenic influences on pubertal timing and tempo and depressive symptoms in boys and girls. Journal of Research on Adolescence. doi:10.1111/jora.12502

Ioannidis JP (2005). Why most published research findings are false. PLoS Medicine, 2(8), e124. doi:10.1371/journal.pmed.0020124 [PubMed: 16060722]

Jaffee SR, & Price TS (2007). Gene-environment correlations: a review of the evidence and implications for prevention of mental illness. Molecular Psychiatry, 12(5), 432–442. doi:10.1038/sj.mp.4001950 [PubMed: 17453060]

James J, Ellis BJ, Schlomer GL, & Garber J (2012). Sex-specific pathways to early puberty, sexual debut, and sexual risk taking: tests of an integrated evolutionary-developmental model. Developmental Psychology, 48(3), 687–702. doi:10.1037/a0026427 [PubMed: 22268605]

Jones HJ, Stergiakouli E, Tansey KE, Hubbard L, Heron J, Cannon M, … Zammit S (2016). Phenotypic manifestation of genetic risk for schizophrenia during adolescence in the general

population. JAMA Psychiatry, 73(3), 221–228. doi:10.1001/jamapsychiatry.2015.3058 [PubMed: 26818099]

Jorm AF, Christensen H, Rodgers B, Jacomb PA, & Easteal S (2004). Association of adverse childhood experiences, age of menarche, and adult reproductive behavior: does the androgen receptor gene play a role? American Journal of Medical Genetics. Part B: Neuropsychiatric Genetics, 125B(1), 105–111. doi:10.1002/ajmg.b.20114

Kyweluk MA, Georgiev AV, Borja JB, Gettler LT, & Kuzawa CW (2018). Menarcheal timing is accelerated by favorable nutrition but unrelated to developmental cues of mortality or familial instability in Cebu, Philippines. Evolution and Human Behavior, 39(1), 76–81. doi:10.1016/j.evolhumbehav.2017.10.002

Leinonen JT, Chen YC, Tukiainen T, Panula P, & Widen E (2019). Transient modification of lin28b expression - Permanent effects on zebrafish growth. Molecular and Cellular Endocrinology, 479, 61–70. doi:10.1016/j.mce.2018.09.001 [PubMed: 30196135]

Lohmueller KE, Pearce CL, Pike M, Lander ES, & Hirschhorn JN (2003). Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nature Genetics, 33(2), 177–182. doi:10.1038/ng1071 [PubMed: 12524541]

Mackrell SV, Sheikh HI, Kotelnikova Y, Kryski KR, Jordan PL, Singh SM, & Hayden EP (2014). Child temperament and parental depression predict cortisol reactivity to stress in middle childhood. Journal of Abnormal Psychology, 123(1), 106–116. doi:10.1037/a0035612 [PubMed: 24661163]

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, … Visscher PM (2009). Finding the missing heritability of complex diseases. Nature, 461(7265), 747–753. doi:10.1038/nature08494 [PubMed: 19812666]

Manuck SB, Craig AE, Flory JD, Halder I, & Ferrell RE (2011). Reported early family environment covaries with menarcheal age as a function of polymorphic variation in estrogen receptor-alpha. Development and Psychopathology, 23(1), 69–83. doi:10.1017/S0954579410000659 [PubMed: 21262040]

Mendle J, Turkheimer E, D'Onofrio BM, Lynch SK, Emery RE, Slutske WS, & Martin NG (2006). Family structure and age at menarche: a children-of-twins approach. Developmental Psychology, 42(3), 533–542. doi:10.1037/0012-1649.42.3.533 [PubMed: 16756443]

Moss EG, & Tang L (2003). Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. Developmental Biology, 258(2), 432–442. doi:10.1016/s0012-1606(03)00126-x [PubMed: 12798299]

Mrug S, Elliott MN, Davies S, Tortolero SR, Cuccaro P, & Schuster MA (2014). Early puberty, negative peer influence, and problem behaviors in adolescent girls. Pediatrics, 133(1), 7–14. doi:10.1542/peds.2013-0628 [PubMed: 24324002]

Muthen LK, & Muthen B (1998–2015). Mplus User's Guide. Seventh Edition. Los Angles, CA: Muthen & Muthen.

Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA, … Benjamin DJ (2016). Genome-wide association study identifies 74 loci associated with educational attainment. Nature, 533(7604), 539–542. doi:10.1038/nature17671 [PubMed: 27225129]

Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, … Ruvkun, G. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature, 408(6808), 86–89. doi:10.1038/35040556 [PubMed: 11081512]

Patel CJ, Bhattacharya J, & Butte AJ (2010). An environment-wide association study (EWAS) on type 2 diabetes mellitus. PloS One, 5(5), e10746. doi:10.1371/journal.pone.0010746 [PubMed: 20505766]

Pembrey M (2004). The Avon Longitudinal Study of Parents and Children (ALSPAC): A resource for genetic epidemiology. European Journal of Endrocrinology, 151, U125–U129. doi:10.1530/eje.0.151u125

Plant TM (2001). Neurobiological bases underlying the control of the onset of puberty in the rhesus monkey: a representative higher primate. Frontiers in Neuroendocrinology, 22(2), 107–139. doi:10.1006/frne.2001.0211 [PubMed: 11259134]

Pluess M, & Belsky J (2013). Vantage sensitivity: individual differences in response to positive experiences. Psychological Bulletin, 139(4), 901–916. doi:10.1037/a0030196 [PubMed: 23025924]

Prentice P, & Viner RM (2013). Pubertal timing and adult obesity and cardiometabolic risk in women and men: a systematic review and meta-analysis. International Journal of Obesity (2005), 37(8), 1036–1043. doi:10.1038/ijo.2012.177 [PubMed: 23164700]

Richardson GB, La Guardia AC, & Klay PM (2018). Determining the roles of father absence and age at menarche in female psychosocial acceleration. Evolution and Human Behavior, 39(4), 437–446. doi:10.1016/j.evolhumbehav.2018.03.009

Rowe DC (2002). On genetic variation in menarche and age at first sexual intercourse: A critique of the Belsky-Draper hypothesis. Evolution and Human Behavior, 23(5), 365–372. doi:10.1016/S1090-5138(02)00102-2

Rutter M, Dunn J, Plomin R, Simonoff E, Pickles A, Maughan B, … Eaves L (1997). Integrating nature and nurture: implications of person-environment correlations and interactions for developmental psychopathology. Development and Psychopathology, 9(2), 335–364. doi:10.1017/s0954579497002083 [PubMed: 9201448]

Sangiao-Alvarellos S, Manfredi-Lozano M, Ruiz-Pino F, Navarro VM, Sanchez-Garrido MA, Leon S, … Tena-Sempere M (2013). Changes in hypothalamic expression of the Lin28/let-7 system and related microRNAs during postnatal maturation and after experimental manipulations of puberty. Endocrinology, 154(2), 942–955. doi:10.1210/en.2012-2006 [PubMed: 23291449]

Schlomer GL, & Cho H-J (2017). Genetic and environmental contributions to age at menarche: Interactive effects of father absence and LIN28B. Evolution and Human Behavior, 38(6), 761–769. doi:10.1016/j.evolhumbehav.2017.06.002

Schlomer GL, Murray J, Yates B, Hair K, & Vandenbergh DJ (2019). Father absence, age at menarche, and sexual behaviors in women: Evaluating the genetic confounding hypothesis using the androgen receptor gene. Evolutionary Behavioral Sciences, 13(3), 205–222. doi:10.1037/ebs0000137

Skoog T, Bayram Ozdemir S, & Stattin H (2016). Understanding the link between pubertal timing in girls and the development of depressive symptoms: The role of sexual harassment. Journal of Youth and Adolescence, 45(2), 316–327. doi:10.1007/s10964-015-0292-2 [PubMed: 25971216]

Sohn K (2017). The null relation between father absence and earlier menarche. Human Nature, 28(4), 407–422. doi:10.1007/s12110-017-9299-6 [PubMed: 28799087]

Tchuenche JM (2007). Theoretical population dynamics model of a genetically transmitted disease: sickle-cell anaemia. Bulletin of Mathematical Biology, 69(2), 699–730. doi:10.1007/s11538-006-9148-y [PubMed: 16933112]

Thornton JE, & Gregory RI (2012). How does Lin28 let-7 control development and disease? Trends in Cell Biology, 22(9), 474–482. doi:10.1016/j.tcb.2012.06.001 [PubMed: 22784697]

Tither JM, & Ellis BJ (2008). Impact of fathers on daughters' age at menarche: a genetically and environmentally controlled sibling study. Developmental Psychology, 44(5), 1409–1420. doi:10.1037/a0013065 [PubMed: 18793072]

Vrijheid M, Slama R, Robinson O, Chatzi L, Coen M, van den Hazel P, … Nieuwenhuijsen MJ (2014). The human early-life exposome (HELIX): project rationale and design. Environmental Health Perspectives, 122(6), 535–544. doi:10.1289/ehp.1307204 [PubMed: 24610234]

Webster GD, Graber JA, Gesselman AN, Crosier BS, & Orozco Schember T (2014). Life history theory of father absence and menarche: A meta-analysis. Evolutionary Psychology, 12(2), 273–294. [PubMed: 25299880]

Widaman KF, Helm JL, Castro-Schilo L, Pluess M, Stallings MC, & Belsky J (2012). Distinguishing ordinal and disordinal interactions. Psychological Methods, 17(4), 615–622. doi:10.1037/a0030003 [PubMed: 22984788]

Wild CP (2005). Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiology, Biomarkers & Prevention, 14(8), 1847–1850. doi:10.1158/1055-9965.EPI-05-0456

Wray NR, Lee SH, Mehta D, Vinkhuyzen AA, Dudbridge F, & Middeldorp CM (2014). Research review: Polygenic methods and their application to psychiatric traits. Journal of Child Psychology and Psychiatry and Allied Disciplines, 55(10), 1068–1087. doi:10.1111/jcpp.12295

Zhu H, Shah S, Shyh-Chang N, Shinoda G, Einhorn WS, Viswanathan SR, … Daley GQ (2010). Lin28a transgenic mice manifest size and puberty phenotypes identified in human genetic association studies. Nature Genetics, 42(7), 626–630. doi:10.1038/ng.593 [PubMed: 20512147]

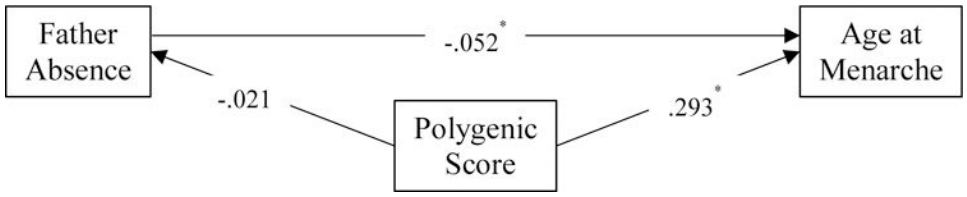

Figure 1.
Testing the father absence genetic confounding hypothesis using an AAM PGS

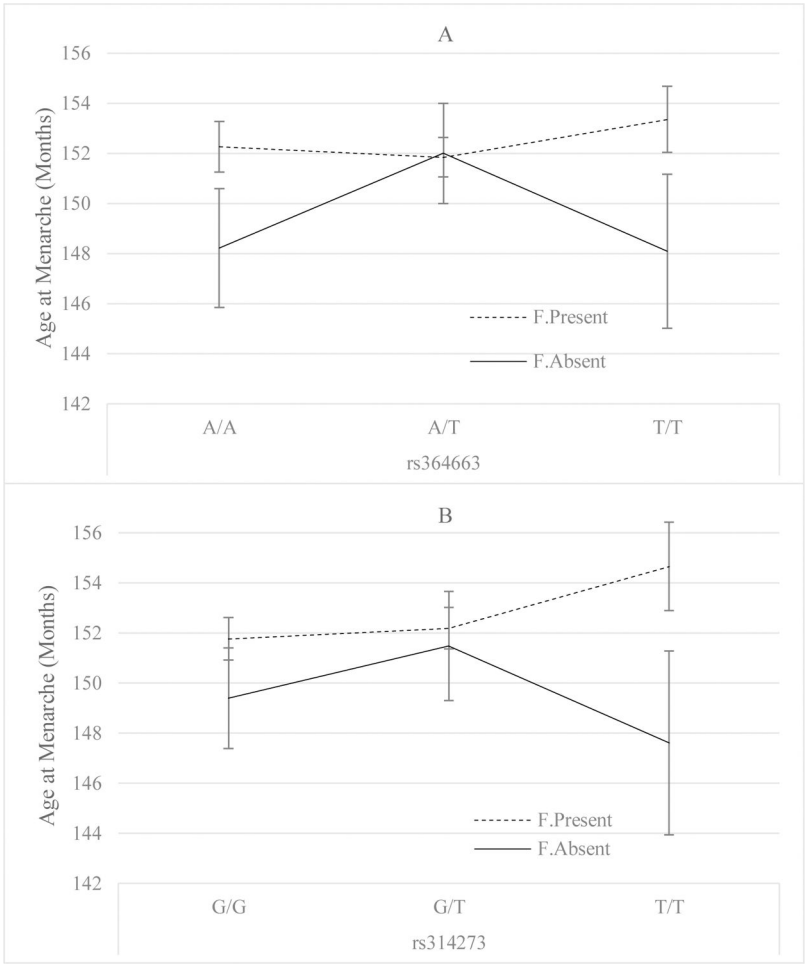

Figure 2.
Interactions between father absence and dummy coded rs364663 (A) and rs314273 (B)

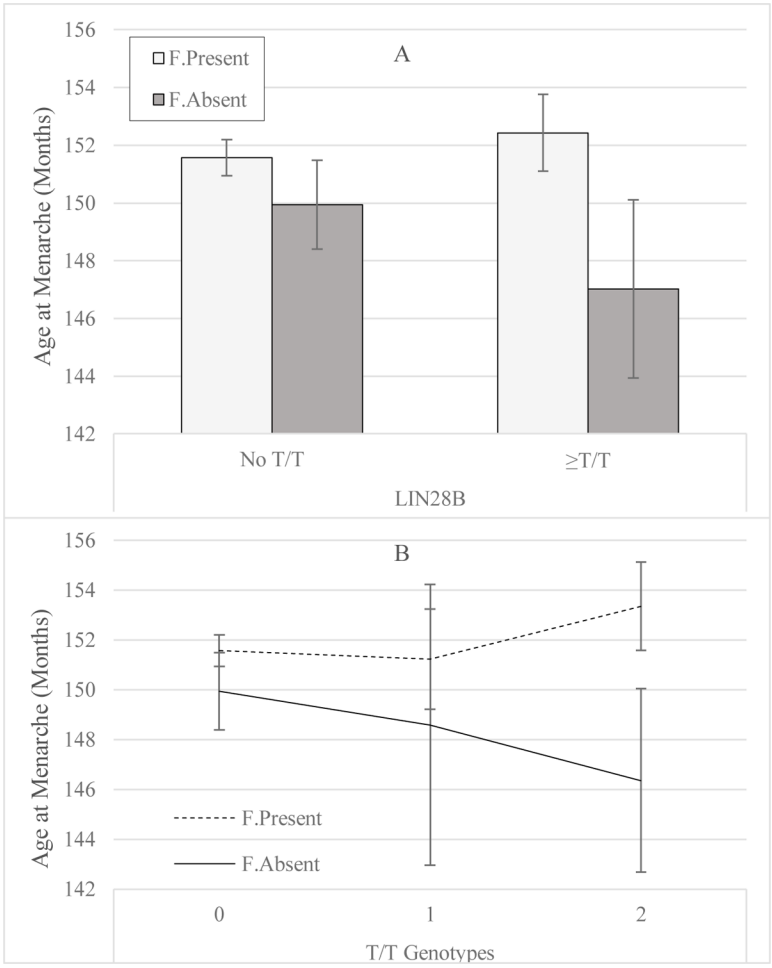

Figure 3.

Father absence-by-*LIN28B* genetic index interactions. (A) at least 1 T/T genotype between rs364663 and rs314273. (B) number of T/T genotypes

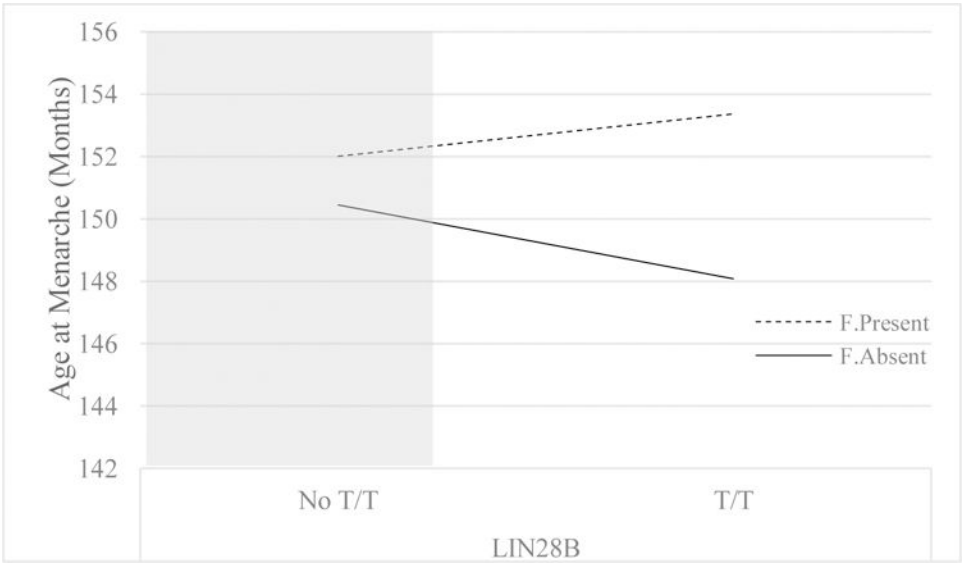

Figure 4.

Test of the interaction form (ordinal vs. disordinal). Shaded area reflects the 95% confidence interval for the slope cross-over point (−1.336 to .304).

Table 1.

Correlation matrix of analysis variables

	Measure	1	2	3	4	5	6	7
1.	Age at Menarche	--						
2.	Father Absence	−.067*	--					
3.	rs364663	.020	−.001	--				
4.	rs314273	.042*	.003	.801*	--			
5.	Polygenic Score	.295*	−.025	.140*	.179*	--		
6.	Financial Difficulties	−.050*	.223*	.013	.002*	−.017	--	
7.	PC1	−.012	.021	−.023	−.050*	−.006	.023	--

Note: rs364663 and rs314273 are linearly coded (0/1/2). PC1 refers to the first principal coordinate.

* $p < .05$; $N = 2{,}685$