Father Absence and Accelerated Reproductive Development in Non-Hispanic White Women in the United States

Lauren Gaydosh[1], **Daniel W. Belsky**[2,3], **Benjamin W. Domingue**[4], **Jason D. Boardman**[5], and **Kathleen Mullan Harris**[6]

[1]Vanderbilt University, Center for Medicine, Health, and Society, 300 Calhoun Hall, PMB #351665, 2301 Vanderbilt Place, Nashville, TN 27235, USA

[2]Duke University School of Medicine Center for Population Health Science

[3]Duke University Population Research Institute

[4]Stanford University, Graduate School of Education

[5]Department of Sociology University of Colorado at Boulder, and Institute of Behavioral Science

[6]Department of Sociology, University of North Carolina at Chapel Hill,Carolina Population Center

Abstract

Girls who experience father absence in childhood also experience accelerated reproductive development in comparison with peers with present fathers. One hypothesis advanced to explain this empirical pattern is genetic confounding, wherein gene-environment correlation (rGE) causes a spurious relationship between father absence and reproductive timing. We test this hypothesis by constructing polygenic scores for age at menarche and first birth using recently available genome-wide association study results and molecular genetic data on a sample of non-Hispanic white females from the National Longitudinal Study of Adolescent to Adult Health. We find that young women's accelerated menarche polygenic scores are unrelated to their exposure to father absence. In contrast, polygenic scores for earlier age at first birth tend to be higher in young women raised in homes with absent fathers. Nevertheless, father absence and the polygenic scores independently and additively predict reproductive timing. We find no evidence in support of the rGE hypothesis for accelerated menarche and only limited evidence in support of the rGE hypothesis for earlier age at first birth.

Introduction

Early pubertal timing is associated with risk for adverse behavioral and health outcomes across the life course (Karapanou and Papadimitriou 2010; Mendle et al. 2009; Tamakoshi et al. 2011). Pubertal timing is influenced by both genetic and environmental factors (Ellis 2004; Ellis et al. 2011; Polderman et al. 2015). One environmental factor consistently associated with earlier pubertal timing is exposure to childhood adversity (Belsky et al. 1991; Ellis 2004)—in particular, absence of the biological father from the home (Webster et

Lauren Gaydosh (corresponding author) lauren.m.gaydosh@vanderbilt.edu.

al. 2014). Two hypotheses are commonly advanced to explain the association between absence of the biological father and early puberty. One hypothesis, derived from evolutionary life history theory (Ellis et al. 2009; Del Giudice et al. 2015; Stearns 1992; Trivers 1972), is that father absence triggers accelerated maturation in the developing organism (Belsky et al. 1991). An alternative hypothesis is that the association is an artifact of gene-environment correlation (rGE) (Barbaro et al. 2017; Rowe 2002).

We report findings from a molecular genetic test of this rGE hypothesis. Using results from a recent genome-wide association study (GWAS) of age at menarche (Day et al. 2017), we quantify genetic influences on pubertal timing in a sample of female adolescents from the National Longitudinal Study of Adolescent to Adult Health (Add Health). We test whether the genetics discovered in GWAS of menarcheal timing (1) predict timing of puberty in Add Health girls, (2) correlate with girls' exposure to father absence, and (3) explain any association between father absence and early puberty. We find no evidence to support the rGE hypothesis with respect to the genetics associated with menarcheal timing. We also consider genetics discovered in GWAS of age at first birth (Barban et al. 2016)—a related reproductive phenotype—and find limited evidence in support of the rGE hypothesis.

Background

Early Puberty and Life Course Development

Pubertal timing is of central interest to demographers because it signals the start of reproductive capacity. For girls, age at menarche is associated with a cascade of behaviors and events in the transition to adulthood that relate to fertility and health outcomes. Earlier menarche is associated with earlier sexual debut, first pregnancy and birth, and marriage (Kiernan 1977; Sandler et al. 1984; Udry 2008). These associations replicate across cultural contexts (Udry and Cliquet 1982). Earlier pubertal timing is also associated with greater risk-taking behavior more generally (Igra and Irwin Jr. 1996; Patton et al. 2004). Age at menarche is negatively correlated with poor health outcomes, including type 2 diabetes (He et al. 2010), cardiovascular disease risk (Feng et al. 2008; Lakshman et al. 2009; Remsberg et al. 2005), breast cancer (Stoll et al. 1994), and mortality (Charalampopoulos et al. 2014; Tamakoshi et al. 2011).

Father Absence and Accelerated Reproductive Development

Rooted in life course and evolutionary frameworks, a robust literature has documented associations between early childhood environment and reproductive profiles across adolescence and into adulthood (see Ellis 2004 for a thorough review). In particular, children exposed to psychosocial stressors during sensitive periods of childhood exhibit accelerated reproductive trajectories, such as early puberty, early sexual debut, and early childbearing (Belsky et al. 1991; Browning et al. 2004; Foster et al. 2008; Graber et al. 1995; Wu and Martinson 1993).

One specific childhood stressor that has received considerable attention is father absence. Girls who grow up in households from which their biological father is absent are consistently observed to experience menarche at younger ages compared with peers with

present fathers (Alvergne et al. 2008; Bogaert 2008; Culpin et al. 2014; Hoier 2003; Quinlan 2003). Father absence is also associated with subsequent reproductive timing: girls with absent fathers have younger age of first sex, pregnancy, and birth (Anderson 2015; Ellis et al. 2003; Kiernan and Hobcraft 1997; Mendle et al. 2009; Moffitt et al. 1992).

Research in anthropology and human development hypothesizes that associations between early-life stressors, including father absence, and accelerated reproductive development reflect a causal process (Belsky et al. 1991; Chisholm et al. 1993; Draper and Harpending 1982; Ellis 2004; Ellis and Garber 2000; Ellis et al. 1999). In this causal process, father absence—either specifically or as one of a several features of early adversity—signals to the developing child that resources may be scarce or unpredictable. This signal, in turn, elicits a biological response in the form of increased allocation of current resources toward achieving reproductive maturity as soon as possible (Belsky et al. 1991; Chisholm et al. 1993; Ellis 2004). Consistent with this hypothesis, father absence and other measures of early adversity are associated with accelerated pubertal development (Foster et al. 2008; Mendle et al. 2009, 2015). We focus here on father absence because it is central to hypotheses of environmental causation and because it is among the best-replicated environmental correlates of accelerated pubertal development in nongenetically sensitive**[AU: Addressing your query here….No hyphen is needed for compounds involving an "-ly" adverb.]** designs (Webster et al. 2014).

Gene-Environment Correlation as an Explanation for Association Between Father Absence and Accelerated Reproductive Development

In contrast to anthropological and human development hypotheses of environmental causation, behavioral genetics researchers have proposed a genetic hypothesis for why girls who grow up in father-absent households tend to experience accelerated reproductive development (Hardy et al. 1998; Rowe 2002). This genetic hypothesis is built on three sets of research findings. First, reproductive timing is influenced by genetic factors (Barban et al. 2016; Burt et al. 2006; Campbell and Udry 1995; Elks et al. 2010; He et al. 2009; Tither and Ellis 2008; Towne et al. 2005). A recent meta-analysis by Polderman and colleagues (2015) estimated heritability for reproductive traits at 0.32, and a study of age at menarche specifically estimated heritability at 0.49 (Towne et al. 2005). In other words, roughly one-half of the variation in age at menarche has been attributed to genetic variation. Second, girls who mature earlier also experience earlier sexual debut (Moore et al. 2014), union formation (Kiernan 1977), and childbearing (Rowe 2002; Sandler et al. 1984; Udry 2008). Third, unions formed at earlier ages and with early childbearing are more likely to be unstable (Booth and Edwards 1985; Bumpass and Sweet 1972).

The hypothesis derived from this evidence posits that the genetic factors that accelerate pubertal timing place women at risk for early childbearing within unstable unions, which in turn results in their daughters inheriting genetics associated with early puberty and all the consequences thereof, including having an absent father (Hardy et al. 1998; Mendle et al. 2006; Rowe 2000, 2002). Stated differently, the mother's genotype is associated with accelerated puberty and the resulting potential for early partnership formation and instability, and the daughter may inherit the genetics of early maturation and be exposed to

father absence. This hypothesis conceptualizes the association between father absence and daughters' early puberty as spurious, confounded by a gene-environment correlation (rGE). If this is true, we would expect to observe an association between the genetics of accelerated reproductive development and father absence, and an attenuation in the association between father absence and reproductive timing when controlling for known genetic factors associated with reproductive development.

To date, empirical tests of genetic confounding are limited to family-based genetic studies (Mendle et al. 2006; Ryan 2015; Tither and Ellis 2008). These studies have compared phenotypic similarities between pairs of relatives with different degrees of genetic relatedness, such as monozygotic and dizygotic twins. They have used these comparisons to partition phenotypic variance into genetic and environmental components and to further isolate environmental variance shared between siblings in a family (Plomin et al. 2013). Behavioral genetic studies have found no evidence of shared environmental influence on pubertal timing, generally interpreting this as evidence against environmental causation hypotheses because father absence is an environmental exposure shared by siblings (Mendle et al. 2006; Rowe 2000; Ryan 2015; Tither and Ellis 2008); this evidence is consistent with genetic confounding. The advent of GWAS and the discovery of molecular genetic influences on reproductive development now afford an opportunity to put the genetic hypothesis to a molecular test. For a more complete discussion of the advantages and disadvantages of molecular genetic tests vis-à-vis behavioral genetics than is possible here, see Conley and Fletcher (2017).

Genome-Wide Association Studies of Human Reproductive Development

Advances in genome science are now yielding molecular detail about genetic influences on human traits and behaviors. GWAS are large-scale data-mining expeditions of the human genome. They correlate millions of genetic variants, called single-nucleotide polymorphisms (SNPs), with phenotypes in samples of tens or even hundreds of thousands of individuals. GWAS on age at menarche (Day et al. 2017; Day et al. 2015; Perry et al. 2014) and age at first birth (Barban et al. 2016) provide an opportunity to conduct a molecular test of the rGE hypothesis advanced to explain associations between father absence and timing of puberty. GWAS results can be used to parameterize predictive algorithms, called *polygenic scores*, which can be applied to genomic data from independent samples (Belsky and Israel 2014; Dudbridge 2013). Using this approach, we conduct the first molecular genetic test of the hypothesis of rGE between genetic influences on reproductive development and father absence. We evaluate genetic confounding of the associations between father absence and accelerated reproductive development, considering both the genetics associated with menarche and age at first birth.

Data and Methods

Data

The National Longitudinal Study of Adolescent to Adult Health (Add Health) is an ongoing, nationally representative longitudinal study of the social, behavioral, and biological linkages in health and developmental trajectories from early adolescence into adulthood. The cohort

was drawn from a probability sample of 132 middle and high schools and is representative of U.S. adolescents in grades 7–12 in 1994–1995. Since the start of the project, participants have been interviewed in home at four data collection waves (numbered I–IV), most recently in 2008. Data include measures of family structure, reproductive development, and genome-wide molecular genetic data (Harris 2010; Harris et al. 2013). Given the complications inherent in working with genetic data in diverse samples (Martin et al. 2017; Wojcik et al. 2017), we restrict our primary analytic sample to 2,681 unrelated non-Hispanic white women, and supplementary analyses of sisters to 411 genetically confirmed non-Hispanic white sisters. Add Health genetic data are available on roughly 800 non-Hispanic black and 700 Hispanic females for future research.

Reproductive Timing—Age at menarche was recorded from interviews conducted at Waves I–III. We use data from the interview wave at which a woman first reported having had her first menstruation. Age at menarche is measured in completed years. Age at first sex was recorded from interviews conducted at Waves I–IV. We use data from the interview wave at which a woman first reported having had vaginal intercourse. Age at first sex was recorded in month units at Add Health Waves I and II and in year units at Waves III and IV. We exclude from analysis data from 13 women who reported an age at first sex less than age 10. Age at first birth was recorded from interviews conducted at Waves I–IV. We use data from the interview wave at which a woman first reported having had her first live birth in the pregnancy/birth history portion of the interview. Age at first birth is measured in months. We exclude from analysis data from two women who reported first birth before age 10.

Age at menarche, first sex, and first birth are positively correlated (see Table 1). Women who develop later in adolescence tend to have a later age at sexual debut, followed by a later age at first birth. Age at first sex and age at first birth are most strongly correlated ($r = .37$).

In Fig. 1, we present Kaplan-Meier estimates for the cumulative proportion of women experiencing each reproductive event by age. The mean age at which women experience a given event is calculated as the area under the Kaplan-Meier survival function. All non-Hispanic white women experience menarche by age 20, with a mean age of 12.2. Nearly all (96 %) experience sexual debut by age 32, with a mean age for those who have sex of 17.2. Finally, 64 % experience a first birth by age 32, and the mean age for those who have a first birth is 26.2.

These estimates are generally consistent with other nationally representative estimates of the timing of reproductive events. Our estimate for mean age at menarche is 0.3 years earlier than estimates for non-Hispanic white girls in the National Health and Nutrition Study (NHANES), which may from differences in methodology: cross-sectional design and proxy report (for girls aged 8–11) in NHANES, compared with longitudinal and self-report in Add Health (Anderson and Must 2005). Our estimate for mean age at first sex is consistent with the 2002 national estimate of 17.4 for all reproductive-age women from the National Survey of Family Growth (Chandra et al. 2005). Finally, our estimate for mean age at first birth falls within national estimates for mean age at first birth among non-Hispanic white women, ranging from 25.9 in 2000 to 27.0 in 2014 (Mathews and Hamilton 2016).

Father Absence—We measure father absence as an indicator for whether the child's biological father was ever nonresident from birth to age 7. We construct this measure drawing from four sources of data: (1) parent's self-reported union history and current union status; (2) youth's reports of their current family structure; (3) youth's reports of the duration of life spent with current household members; and (4) whether and for how long a youth lived with a nonresident biological parent (Gaydosh and Harris forthcoming). Almost one-fifth (18 %) of girls were born into a family structure where the biological father was nonresident (not shown). By age 7, 31 % of girls experience biological father nonresidence for a period of at least one year (see Table 1).

Genotyping—At the Wave IV interview in 2008, saliva and capillary whole blood were collected from respondents. Of the 15,701 individuals interviewed, 15,159 consented to genotyping, and 12,254 agreed to genetic data archiving. After quality-control procedures (Highland et al. 2018; McQueen et al. 2015), genotype data were available for 9,974 individuals. Data were generated for 7,917 individuals from the Illumina HumanOmni1-Quad BeadChip, which includes 1,140,419 markers, and for 2,057 individuals from the Illumina HumanOmni2.5-Quad BeadChip, which includes 2,369,541 markers. We conduct analysis on the 609,130 SNPs genotyped on both chips. Additional information on the selection of the final SNP marker set can be found in the quality control report (Highland et al. 2018).

Polygenic Scores (PGS)—We calculate polygenic scores for Add Health participants based on published GWAS results for age at menarche from the ReproGen Consortium (Day et al. 2017) and for age at first birth from the Social Science Genetic Association Consortium and Sociogenome (Barban et al. 2016). This allows us to control for the known genetic factors associated with two reproductive phenotypes, both of which are associated with father absence. The GWAS were conducted on independent discovery samples to estimate associations between SNPs and the phenotypes of interest (age at menarche and first birth). SNPs are a common type of genetic variation where there is a difference in a single base pair in the DNA sequence; for each SNP, individuals may have no, one, or two copies of a reference allele (Bush et al. 2012).

Scores are calculated following the method described by Dudbridge (2013). Briefly, SNPs in the genotyped sample are matched to published GWAS results. For each of these SNPs ($N =$ 577,997 menarche, $N =$ 527,258 age at first birth), a loading is calculated as the number of phenotype-associated alleles (0, 1, or 2) multiplied by the effect size estimated in the original GWAS. Loadings are then summed across the SNP set to calculate the polygenic score.

The use of polygenic scores in population research is complicated by *population stratification*, or the nonrandom patterning of alleles across global populations (Cardon and Palmer 2003). Population stratification is a potential confounder in genetic association studies (Hamer and Sirota 2000) and, by extension, polygenic score analysis (Belsky and Israel 2014; Martin et al. 2017). GWAS and polygenic scoring rely on a subset of genetic markers to act as proxies for unmeasured genetic variation. This approach is effective because much of the genome is in *linkage disequilibrium*: that is, certain groups of

genotypes tend to be inherited together and are thus correlated. Measuring one in a set of such genotypes can effectively provide information about the multiple unmeasured genotypes nearby. However, because patterns of linkage disequilibrium reflect genetic inheritance, they vary across populations of different ancestry. A result of ancestry-associated differences in patterns of linkage disequilibrium is that a given genotype may contain different information about the genome when measured in one population versus another (Carlson et al. 2013; Shifman et al. 2003; Wojcik et al. 2017). In the context of polygenic scoring, applying GWAS results derived in one population to compute a polygenic score in a different population introduces measurement error, which is one reason why polygenic scores derived from GWAS of European ancestry populations have reduced effect sizes in African ancestry populations (Belsky et al. 2013; Domingue et al. 2014, 2015).

Importantly, reduction in the predictive accuracy of polygenic scores may also reflect confounding by environmental factors that are differentially distributed across populations. Large differences in environmental factors, such as neighborhood social disorganization or exposure to discrimination, might suppress the influence of genetic factors in minority populations compared with non-Hispanic and white populations (Boardman et al. 2012, 2017). We believe that these are critical questions; however, given differences in allele frequency and linkage disequilibrium patterns across groups, the purpose of this study— testing rGE—is best served by restricting the analytic sample to increase the comparability to the population studied in the original GWAS. We therefore restrict our analysis to a genetically identified sample of unrelated women. Inclusion in this sample is based on the respondent's first two genome-wide principal components falling within a given distance of the centroid of the first two principal components computed on the self-identified non-Hispanic white respondents (for greater detail, see the online supplement for Domingue et al. 2018).

Sample restrictions based on self-reported race/ethnicity may not completely protect against population stratification–related confounding (Campbell et al. 2005). To address residual population stratification within self-reported non-Hispanic whites in Add Health, we follow established practice and adjust analyses for principal components estimated from the genome-wide SNP data (Price et al. 2010). Principal component analysis is a statistical technique used to capture variation in a large set of variables using a reduced set of factors. We estimate principal components among non-Hispanic white respondents using the genome-wide SNP data according to the method described by Price and colleagues (2006), using the PLINK command *pca*. We then residualize polygenic scores for the first 10 principal components: that is, we regress polygenic scores on the 10 principal-component scores and compute residual values from the predictions (Conley et al. 2016a). Residualized polygenic scores are standardized (mean = 0, standard deviation = 1) for analysis.

We scale the polygenic score for age at menarche so that higher values correspond to genetic prediction of earlier age at menarche. We hereafter refer to this measure as the *accelerated menarche polygenic score*. Similarly, we scale the polygenic score for age at first birth so that higher values correspond to genetic prediction of earlier age at first birth. We hereafter refer to this measure as the *earlier first-birth polygenic score*. Accelerated menarche and

earlier first-birth polygenic scores are computed for $n = 2,681$ unrelated non-Hispanic white women.

Methods

We use survival methods to analyze timing of reproductive events. We employ the Kaplan-Meier method of estimating the survival function to describe the cumulative proportion of females experiencing an event by a given age. This method accommodates censoring, in which a woman has not yet experienced an event by the end of follow-up, and provides estimates of mean age at event as the area under the curve. We use the Mantel-Haenszel log-rank test to compare survival curves for different risk groups as a way to test for the association of covariates.

We use a nonparametric proportional hazard model approach to test associations of genetic and environmental risks with accelerated reproductive development. This approach yields coefficients with a relative hazard interpretation, similar to the more familiar Cox model. The nonparametric proportional hazard approach can accommodate ties in event times, which are common in discretely measured data, such as age at menarche or age at first birth.

We estimate a generalized linear model with binomial error structure and a complementary log-log link function, including probability weights to adjust for the sampling design of Add Health (Chen and Chantala 2014; Highland et al. 2018). The model is expressed as

$$\log(-\log(1 - \lambda(t_j|x_i))) = \alpha_j + x_i'\beta,$$

where $\lambda(t_j|x_i)$ is the hazard of individual i at time t_j, and α_j is the complementary log-log transformation of the baseline hazard $\lambda_0(t_j)$. This method accommodates continuous time measured in discrete intervals (completed years) and does not make any assumptions about the shape of the baseline hazard or the hazard in the interval (for background on survival analysis, see Box-Steffensmeier and Jones 2004; Kalbfleisch and Prentice 2002). Individuals contribute exposure intervals only with complete information: for example, a respondent who reports 12 as age at menarche could have experienced the event any time between her 12th and 13th birthday. However, because we cannot know when exactly in the interval the event occurred, she contributes exposure only to the preceding intervals (i.e., up to age 12). All analyses were conducted using Stata version 14 (StataCorp 2015).

Results

Father Absence and Reproductive Timing

Table 2 presents hazard ratios for the relationship between father absence and reproductive development. Childhood experience of father absence—from birth to age 7—is significantly associated with earlier timing of all three reproductive events. Net of age, the risk of menarche is 24 % higher, on average, for girls who experienced father absence in childhood compared with peers who lived continuously with their fathers from birth to age 7. Exposure

to father absence by age 7 is also associated with earlier age at first sex and first birth, with larger effect sizes (hazard ratio (HR) = 1.49 and 1.52, respectively).

We illustrate the magnitude of the associations between father absence and reproductive timing in Fig. 2. Across all reproductive events, father absence is associated with a significant acceleration in timing. Girls with absent fathers experience menarche two months earlier than girls with present fathers (at ages 12.1 and 12.3, respectively). Similarly, girls with absent fathers experience sexual debut almost a full year earlier compared with those with present fathers (at ages 16.2 and 17.6, respectively). Finally, girls with absent fathers go on to have their first birth nearly two years earlier, on average, compared with those with present fathers (at ages 25.0 and 26.8, respectively).

Accelerated Menarche Polygenic Score

Girls with larger accelerated menarche polygenic scores experience menarche at younger ages, on average, compared with peers with lower polygenic scores (HR = 1.33; see Table 2). We find no association between the accelerated menarche polygenic score and age at first sex or age at first birth.

Figure 3 presents the cumulative proportion of women experiencing each reproductive event by age separately for high and low polygenic score groups. High score (solid) is defined as greater than or equal to 1 standard deviation above the mean, and low score (hollow) is defined as less than or equal to 1 standard deviation below the mean. Girls in the high polygenic score group experience menarche almost a full year earlier, on average, compared with those in the low polygenic score group (at ages 11.7 and 12.8, respectively).

Testing Gene-Environment Correlation

We test for gene-environment correlation in the relationship between father absence and menarcheal timing by computing correlations between girls' accelerated menarche polygenic scores and exposure to father absence. The accelerated menarche polygenic score is not associated with father absence; we find no correlation between the accelerated menarche polygenic score and father absence ($r = .0001$; see Table 1). In Table 3, we compare the mean standardized polygenic scores by father absence, finding that girls with absent fathers have accelerated menarche polygenic scores similar to those of girls with present fathers.

We formally test for confounding by rGE in a hazard model regressing father absence on reproductive timing, controlling for the accelerated menarche polygenic score (see Table 2). Father absence remains significantly associated with age at menarche, first sex, and first birth after controlling for **[AU: I edited the sentence to avoid a modifier problem. That is, sentence as written attributes the action of "controlling for" to "Father absence." If you don't like "after we control for" here, can you suggest a rewrite that avoids the modifier problem?]** known genetic influences on menarcheal timing. For age at menarche in particular, father absence and genetic risk are independent and additive predictors. We find no evidence that the genetics of menarche confound the relationship between father absence and pubertal timing.

Earlier First-Birth Polygenic Score

We conduct the same analyses as outlined earlier using the polygenic score for age at first birth. On average, women with higher earlier first-birth polygenic scores experience accelerated timing of all three reproductive events (see Table 4). Similar to results for the accelerated menarche polygenic score, the effect size for the standardized earlier first-birth polygenic score is largest for the phenotype matching the original GWAS (for age at menarche, HR = 1.07; for age at first sex, HR = 1.20; for age at first birth, HR = 1.28).

Figure 4 presents the cumulative proportion of women experiencing each reproductive event by age separately for high and low earlier first-birth polygenic score groups. High score (solid) is defined as greater than or equal to 1 standard deviation above the mean, and low score (hollow) is defined as less than or equal to 1 standard deviation below the mean. Girls in the high earlier first-birth polygenic score group experience menarche approximately two months earlier than their peers in the low polygenic score group (at ages 12.1 and 12.3, respectively). The gap in reproductive timing widens for subsequent events: girls in the high earlier first-birth polygenic score group experience first sex more than two years earlier (at age 16.1 vs. 18.3) and first birth almost three years earlier (at age 24.8 vs. 27.7) compared with girls in the low earlier first-birth polygenic score group.

Finally, we test the rGE hypothesis using the earlier first-birth polygenic score. The earlier first-birth polygenic score and father absence are correlated at $r = .08$ (Table 1); women who experience father absence have significantly higher earlier first-birth polygenic scores (Table 3). Some evidence suggests that the distribution of earlier first-birth polygenic scores varies by father absence, as shown in the right panel of Fig. 5. However, we test whether the association between father absence and reproductive timing is confounded by the earlier first-birth polygenic score (Table 4). When the earlier first-birth polygenic score is added to the hazard models, effect sizes for associations between father absence and reproductive timing phenotypes are slightly reduced (by 4 % for age at menarche and first sex, and by 8 % for age at first birth), but they remain statistically significant.

Robustness

Ultimately, the only way to exclude confounding to genetic analysis from population stratification is to analyze genetic differences between relatives who share the same ancestry (Conley et al. 2015). We conduct our primary analysis using a subsample of unrelated individuals. Add Health includes a sibling sample, and we use this sample to repeat polygenic score analyses within a design that excludes potential confounding by population stratification. We test genetic associations using sister fixed-effects models, restricting again to individuals who identify as non-Hispanic white ($n = 409$). We report results in the appendix (Table 5). Among sisters, the carrier of the higher accelerated menarche polygenic score experienced marginally earlier menarche compared with her lower-scored sister. The effect size is approximately the same as found in the analysis of unrelated individuals, arguing strongly against confounding by population stratification. In contrast, the effect size for the earlier first-birth polygenic score is substantially reduced in the sibling fixed-effects analysis and is no longer statistically significant. This reduction in effect size could signal some confounding by population stratification. Alternatively, it could imply other processes

causing sisters to have correlated reproductive development outcomes. However, the sample size for sisters in the analysis of first birth is greatly reduced given that many women have not experienced a first birth by the time of last interview. Low discordance on father absence prohibits an analysis of father absence in the sister subsample.

Results are robust to different measures of family environment. When we restrict the measure of father absence to nonresidence at the time of birth, results are almost identical. Stepfather coresidence is similarly predictive of earlier reproductive timing and is not confounded by genetic risk (see Table 6 in the appendix).

Limitations

This study has limitations. Ages at reproductive events are measured in years, which is imprecise, especially for timing of menarche. Nevertheless, the polygenic score for accelerated menarche and age at menarche are correlated, indicating that the differences observed are substantive and not the result of measurement error. Data on age at first birth are right-censored, with follow-up ending between ages 24 and 32 (Add Health Wave IV). Although the modeling strategy accommodates this type of data, future analyses should consider women with completed reproduction. Father coresidence is reported retrospectively. However, a strength of our measure is that it draws information from both primary caregiver and child reports, and it has been validated elsewhere (Gaydosh and Harris forthcoming). Polygenic scores based on current GWAS results are incomplete measures of genetic influences on reproductive development. The GWAS that we use to develop polygenic scores included hundreds of thousands of individuals. Nevertheless, the majority of genetic influence on reproductive development inferred based on family studies remains unexplained. As GWAS sample sizes increase, polygenic score performance is expected to improve (Okbay et al. 2016), which may change our ability to detect genetic confounding of associations between father absence and reproductive timing. Future GWAS of additional phenotypes, such as family structure, may also enable further testing of possible genetic correlation (Barbaro et al. 2017).

Residual population stratification may confound our analysis of rGE. Our sibling comparison analysis rules out confounding by population stratification in the case of genetic associations with reproductive phenotypes. However, sibling comparison analysis is not feasible for tests of genetic association with father absence. We include statistical adjustment for principal components estimated from the genetic data to address this limitation, the standard in the field (Price et al. 2006). Nevertheless, studies with larger samples of siblings could strengthen confidence that our null findings for rGE were not confounded by residual population stratification.

Add Health represents a single cohort in the United States. Genetic influences on human traits and behaviors change over time and across space (Conley et al. 2016b; Demerath et al. 2013; Liu and Guo 2015; Tropf et al. 2017; Walter et al. 2016). Replication of our findings in samples from different birth cohorts and from different countries will clarify the extent to which findings from recent U.S. birth cohorts generalize.

Finally, our findings are restricted to non-Hispanic white women. As we discuss earlier, this is motivated by the limitations of the current GWAS results, which are conducted in European ancestry populations. This is a pressing limitation to current empirical investigations of genetic and social factors influencing biological and behavioral outcomes, and GWAS in non-European ancestry populations should be a priority. It is our hope that new data, methods, or integrative frameworks become available in the near future that better enable researchers to use genome-wide data from all individuals.

Discussion

We find no evidence of genetic confounding in the relationship between father absence and accelerated menarcheal timing. Women with higher accelerated menarche polygenic scores experience menarche and first sex earlier. Family structure risks for early puberty and sexual debut are independent of accelerated menarche polygenic scores. The rGE hypothesis explaining the link between father absence and earlier puberty is not supported in our sample. Father absence and GWAS–discovered genetic risk for pubertal timing are independent and additive predictors of adolescent development. If common genetic factors do link family environments and pubertal timing, they have not yet been uncovered in GWAS of menarche.

The findings for age at first birth are similar although less conclusive. Some evidence shows that girls with absent fathers have, on average, greater GWAS–discovered genetic risk for younger age at first birth compared with girls with present fathers. Although the magnitude of the associations between father absence and reproductive timing declines after we control for the earlier first-birth polygenic score, the father absence association remains significant. Despite possible genetic confounding, known genetic factors associated with age at first birth cannot fully account for the acceleration in reproductive timing associated with father absence. Examining individual reproductive timing and subsequent partnership formation and stability is a fruitful area for future research on this topic.

Our findings do not support the use of molecular genetic testing in order to make specific predictions about outcomes for individuals. However, molecular genetic data can provide useful information about population parameters, contributing to our understanding of variation in reproductive timing. Polygenic scores for reproductive timing capture otherwise unobserved heterogeneity, allowing for more precise estimation of environmental effects, as well as investigations of interactions between genes and environments. We provide here an example of an approach proposed by Manski (2011) and others (Belsky and Israel 2014; Benjamin et al. 2012) that applies molecular genetic discoveries to test effects of environmental variables. Future studies may use the polygenic scores that we studied as a measure of genetic liability to early puberty.

The lack of evidence for rGE between age at menarche and father absence suggests that something else may underlie the consistent observation that girls who live apart from their fathers mature earlier than those who live together. Our findings are consistent with existing anthropological and developmental theories positing that early exposure to adversity accelerates physical development, perhaps through environmental signaling of resource

scarcity, and behavioral or physiological responses to stress (Belsky et al. 2007; Kyweluk et al. 2018; McEwen 2012). Future research should investigate how father-absent family environments may influence the biological antecedents to puberty. Such research may identify targets for social interventions to modify the reproductive trajectories on which family environments set developing girls. The modest rGE between age at first birth and father absence suggests new directions for social science research. An ever-expanding array of genetically informed social science data resources afford new opportunities to investigate how genetics associated with early reproduction become correlated with family structure across generations and what social environmental factors may modify this relationship.

Acknowledgments

Appendix

Table 5

Unrelated non-Hispanic white females and sister comparison

	Age at Menarche		Age at First Sex		Age at First Birth	
	OLS	FE	OLS	FE	OLS	FE
Menarche Polygenic Score	-0.39 ***	-0.23 †	-0.15 **	-0.25	-1.55	3.70
	(0.03)	(0.13)	(0.05)	(0.24)	(1.10)	(5.58)
First-Birth Polygenic Score	-0.11 ***	0.13	-0.51 ***	-0.25	-6.52 ***	-6.02
	(0.03)	(0.13)	(0.05)	(0.24)	(1.13)	(5.88)
Model	OLS	FE	OLS	FE	OLS	FE
N	2,681	409	2,582	409	1,517	249

Notes: Results from separate linear regression models for reproductive timing for each polygenic score and reproductive event separately for unrelated white females and sister fixed-effects. Standard errors are shown in parentheses. All models include appropriate sample weights and correct for clustered sampling design. OLS = ordinary least squares; FE = fixed effects.

† $p < .10$

** $p < .01$

*** $p < .001$

Table 6

Alternative environmental exposures, hazard ratios for reproductive events: Unrelated non-Hispanic white females

	Menarche (outcome and polygenic score)			First Birth (outcome and polygenic score)		
Polygenic Score	1.33***	1.32***	1.33***	1.29***	1.27***	1.27***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Father Absent at Birth	1.21***		1.19***	1.50***		1.47***
	(0.06)		(0.06)	(0.09)		(0.09)
Stepfather by Age 7		1.28***	1.37***		1.52***	1.46***
		(0.09)	(0.10)		(0.13)	(0.13)
Age	One-year intervals, 7–19			One-year intervals, 15–31		
Events	2,681			1,517		
N	2,681			2,681		

Notes: Results from complementary log-log hazard models for reproductive timing. Yearly age intervals for the range specified. Standard errors are shown in parentheses. All models include appropriate sample weights and correct for clustered sampling design.

*
  $p < .05$

**
  $p < .01$

***
  $p < .001$

References

Alvergne A, Faurie C, & Raymond M (2008). Developmental plasticity of human reproductive development: Effects of early family environment in modern-day France. Physiology & Behavior, 95, 625–632. [PubMed: 18822309]

Anderson KG (2015). Father absence, childhood stress, and reproductive maturation in South Africa. Human Nature, 26, 401–425. [PubMed: 26471378]

Anderson SE, & Must A (2005). Interpreting the continued decline in the average age at menarche: Results from two nationally representative surveys of US girls studied 10 years apart. Journal of Pediatrics, 147, 753–760. [PubMed: 16356426]

Barban N, Jansen R, de Vlaming R, Vaez A, Mandemakers JJ, Tropf FC, … Mills MC (2016). Genome-wide analysis identifies 12 loci influencing human reproductive behavior. Nature: Genetics, 48, 1462–1472. [PubMed: 27798627]

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, 357–365.

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

Belsky DW, Moffitt TE, Sugden K, Williams B, Houts R, McCarthy J, & Caspi A (2013). Development and evaluation of a genetic risk score for obesity. Biodemography and Social Biology, 59, 85–100. [PubMed: 23701538]

Belsky J, Steinberg LD, & Draper P (1991). Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. Child Development, 62, 647–670. [PubMed: 1935336]

Belsky J, Steinberg LD, Houts RM, Friedman SL, DeHart G, Cauffman E, … NICHD Early Child Care Research Network. (2007). Family rearing antecedents of pubertal timing. Child Development, 78, 1302–1321. [PubMed: 17650140]

Benjamin DJ, Cesarini D, Chabris CF, Glaeser EL, Laibson DI, Guðnason V, … Lichtenstein P (2012). The promises and pitfalls of genoeconomics. Annual Review of Economics, 4, 627–662.

Boardman JD, Barnes L, Wilson R, & Evans D (2012). Social disorder, APOE-E4 genotype, and change in cognitive function among older adults living in Chicago. Social Science & Medicine, 74, 1584–1590. [PubMed: 22465377]

Boardman JD, Wedow R, Goode JA, Domingue BW, & Harris KM (2018). A social explanation for observed differences in genetic associations for years of completed education among black and white adults: The role of discrimination. Paper presented at the annual meetings of the Population Association of America, Denver, CO.Bogaert AF (2008). Menarche and father absence in a national probability sample. Journal of Biosocial Science, 40, 623–636. [PubMed: 17761007]

Booth A, & Edwards J (1985). Age at marriage and marital instability. Journal of Marriage and the Family, 47, 67–75.

Box-Steffensmeier JM, & Jones BS (2004). Event history modeling: A guide for social scientists. New York, NY: Cambridge University Press.

Browning C, Leventhal T, & Brooks-Gunn J (2004). Neighborhood context and racial differences in early adolescent sexual activity. Demography, 41, 697–720. [PubMed: 15622950]

Bumpass LL, & Sweet J (1972). Differentials in marital instability: 1970. American Sociological Review, 37, 754–766.

Burt SA, McGue M, DeMarte JA, Krueger RF, & Iacono WG (2006). Timing of menarche and the origins of conduct disorder. Archives of General Psychiatry, 63, 890–896. [PubMed: 16894065]

Bush WS, Moore JH, Li J, McDonnell S, & Rabe K (2012). Chapter 11: Genome-wide association studies. PLoS Computational Biology, 8(12), e1002822. doi:10.1371/journal.pcbi.1002822 [PubMed: 23300413]

Campbell B, & Udry J (1995). Stress and age at menarche of mothers and daughters. Journal of Biosocial Science, 27, 127–134. [PubMed: 7738076]

Campbell CD, Ogburn EL, Lunetta KL, Lyon HN, Freedman ML, Groop LC, … Hirschhorn JN (2005). Demonstrating stratification in a European American population. Nature: Genetics, 37, 868–872. [PubMed: 16041375]

Cardon L, & Palmer L (2003). Population stratification and spurious allelic association. Lancet, 361, 598–604. [PubMed: 12598158]

Carlson CS, Matise TC, North KE, Haiman CA, Fesinmeyer MD, Buyske S, … (2013). Generalization and dilution of association results from European GWAS in populations of non-European ancestry: The PAGE Study. PLoS Biology, 11(9), e1001661 10.1371/journal.pbio.1001661 [PubMed: 24068893]

Chandra A, Martinez G, Mosher W, Abma J, & Jones J (2005). Fertility, family planning, and reproductive health of U.S. women: Data from the 2002 National Survey of Family Growth (Vital and Health Statistics Report, Series 23, No. 25). Washington, DC: National Center for Health Statistics.

Charalampopoulos D, McLoughlin A, Elks CE, & Ong KK (2014). Age at menarche and risks of all-cause and cardiovascular death: A systematic review and meta-analysis. American Journal of Epidemiology, 180, 29–40. [PubMed: 24920784]

Chen P, & Chantala K (2014). Guidelines for analyzing Add Health data. Chapel Hill: Carolina Population Center, University of North Carolina at Chapel Hill.

Chisholm JS, Ellison PT, Evans J, Lee PC, Lieberman LS, Pavlik Z, … Worthman CM (1993). Death, hope, and sex: Life-history theory and the development of reproductive strategies. Current Anthropology, 34, 1–24.

Conley D, Domingue BW, Cesarini D, Dawes C, Rietveld CA, & Boardman JD (2015). Is the effect of parental education on offspring biased or moderated by genotype? Sociological Science, 2, 82–105. [PubMed: 29051911]

Conley D, & Fletcher J (2017). The genome factor. Princeton, NJ: Princeton University Press.

Conley D, Laidley T, Belsky DW, Fletcher JM, Boardman JD, & Domingue BW (2016a). Assortative mating and differential fertility by phenotype and genotype across the 20th century. Proceedings of the National Academy of Sciences of the United States of America, 113, 6647–6652. [PubMed: 27247411]

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

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, 291–301. [PubMed: 24636689]

Day FR, Perry JRB, & Ong KK (2015). Genetic regulation of puberty timing in humans. Neuroendocrinology, 102, 247–255. [PubMed: 25968239]

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, 834–841. [PubMed: 28436984]

Del Giudice M, Gangestad SW, & Kaplan HS (2015). Life history theory and evolutionary psychology In Buss DM (Ed.), The handbook of evolutionary psychology (2nd ed., Vol. 1, pp. 88–114). Hoboken, NJ: John Wiley & Sons.

Demerath EW, Choh AC, Johnson W, Curran JE, Lee M, Bellis C, … Towne B (2013). The positive association of obesity variants with adulthood adiposity strengthens over an 80-year period: A gene-by-birth year interaction. Human Heredity, 75(2–4), 175–185. [PubMed: 24081233]

Domingue BW, Belsky DW, Conley D, Harris KM, & Boardman JD (2015). Polygenic influence on educational attainment: New evidence from the National Longitudinal Study of Adolescent to Adult Health. AERA Open, 1(3),1–13. https://doi:10.1177/2332858415599972 [PubMed: 28164148]

Domingue BW, Belsky DW, Harris KM, Smolen A, McQueen MB, & Boardman JD (2014). Polygenic risk predicts obesity in both white and black young adults. PLoS ONE, 9(7), e101596 https://doi: 10.1371/journal.pone.0101596 [PubMed: 24992585]

Domingue BW, Beksly DW, Fletcher JM, Conley D, Boardman JB, & Harris KM (2018). The social genome of friends and schoolmates in the National Longitudinal Study of Adolescent to Adult Health. Proceedings of the National Academy of Sciences. Advance online publication. 10.1073/pnas.1711803115

Draper P, & Harpending H (1982). Father absence and reproductive strategy: An evolutionary perspective. Journal of Anthropological Research, 31, 255–273.

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

Elks CE, Perry JRB, Sulem P, Chasman DI, Franceschini N, He C, … Murray A (2010). Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. Nature: Genetics, 42, 1077–1085. [PubMed: 21102462]

Ellis BJ (2004). Timing of pubertal maturation in girls: An integrated life history approach. Psychological Bulletin, 130, 920–958. [PubMed: 15535743]

Ellis BJ, Bates JE, Dodge KA, Fergusson DM, Horwood LJ, Pettit GS, & Woodward L (2003). Does father absence place daughters at special risk for early sexual activity and teenage pregnancy? Child Development, 74, 801–821. [PubMed: 12795391]

Ellis BJ, Figueredo AJ, Brumbach BH, & Schlomer GL (2009). Fundamental dimensions of environmental risk. Human Nature, 20, 204–268. [PubMed: 25526958] Ellis BJ, & Garber J (2000). Psychosocial antecedents of variation in girls' pubertal timing: Maternal depression, stepfather presence, and marital and family stress. Child Development, 71, 485–501. [PubMed: 10834479]

Ellis BJ, McFadyen-Ketchum S, Dodge KA, Pettit GS, & Bates JE (1999). Quality of early family relationships and individual differences in the timing of pubertal maturation in girls: A longitudinal test of an evolutionary model. Journal of Personality and Social Psychology, 77, 387–401. [PubMed: 10474213]

Ellis BJ, Shirtcliff EA, Boyce WT, Deardorff J, & Essex MJ (2011). Quality of early family relationships and the timing and tempo of puberty: Effects depend on biological sensitivity to context. Development and Psychopathology, 23, 85–99. [PubMed: 21262041]

Feng Y, Hong X, Wilker E, Li Z, Zhang W, Jin D, … Xu X (2008). Effects of age at menarche, reproductive years, and menopause on metabolic risk factors for cardiovascular diseases. Atherosclerosis, 196, 590–597. [PubMed: 17675039]

Foster H, Hagan J, & Brooks-Gunn J (2008). Growing up fast: Stress exposure and subjective "weathering" in emerging adulthood. Journal of Health and Social Behavior, 49, 162–177. [PubMed: 18649500] Graber JA, Brooks-Gunn J, & Warren MP (1995). The antecedents of menarcheal age: Heredity, family environment, and stressful life events. Child Development, 66, 346–359. [PubMed: 7750370]

Gaydosh L, & Harris KM (Forthcoming). Childhood family instability and young adult health. Journal of Health and Social Behavior.

Hamer D, & Sirota L (2000). Beware the chopsticks gene. Molecular Psychiatry, 5, 11–13. [PubMed: 10673763]

Hardy JB, Astone NM, Brooks-Gunn J, Shapiro S, & Miller TL (1998). Like mother, like child: Intergenerational patterns of age at first birth and associations with childhood and adolescent characteristics and adult outcomes in the second generation. Developmental Psychology, 34, 1220–1232. [PubMed: 9823507]

Harris KM (2010). An integrative approach to health. Demography, 47, 1–22. [PubMed: 20355681]

Harris KM, Halpern CT, Hussey J, Whitsel EA, Killeya-Jones L, Tabor J, … Smolen A (2013). Social, behavioral, and genetic linkages from adolescence into adulthood. American Journal of Public Health, 103(S1), S25–S32. [PubMed: 23927505]

He C, Kraft P, Chen C, Buring JE, Paré G, Hankinson SE, … Chasman DI (2009). Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. Nature: Genetics, 41, 724–728. [PubMed: 19448621]

He C, Zhang C, Hunter DJ, Hankinson SE, Buck Louis GM, Hediger ML, & Hu FB (2010). Age at menarche and risk of Type 2 diabetes: Results from 2 large prospective cohort studies. American Journal of Epidemiology, 171, 334–344. [PubMed: 20026580]

Highland H, Avery C, Duan Q, Li Y, & Harris KM (2018). Quality control analysis of Add Health GWAS data. User guide. Retrieved from https://www.cpc.unc.edu/projects/addhealth/documentation/guides/AH_GWAS_QC.pdf

Hoier S (2003). Father absence and age at menarche. Human Nature, 14, 209–233. [PubMed: 26190208]

Igra V, & Irwin C Jr. (1996). Theories of adolescent risk-taking behavior In DiClemente RJ, Hansen WB, & Ponton LE (Eds.), Handbook of adolescent health risk behavior (pp. 35–51). New York, NY: Springer Science+Business Media.

Kalbfleisch JD, & Prentice RL (2002). The statistical analysis of failure time data. Hoboken, NJ: John Wiley & Sons.

Karapanou O, & Papadimitriou A (2010). Determinants of menarche. Reproductive Biology and Endocrinology, 8, 115 10.1186/1477-7827-8-115 [PubMed: 20920296]

Kiernan KE (1977). Age at puberty in relation to age at marriage and parenthood: A national longitudinal study. Annals of Human Biology, 4, 301–308. [PubMed: 931356]

Kiernan KE, & Hobcraft J (1997). Parental divorce during childhood: Age at first intercourse, partnership and parenthood. Population Studies, 51, 41–55.

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, 76–81.

Lakshman R, Forouhi NG, Sharp SJ, Luben R, Bingham SA, Khaw K-T, … Ong KK (2009). Early age at menarche associated with cardiovascular disease and mortality. Journal of Clinical Endocrinology & Metabolism, 94, 4953–4960. [PubMed: 19880785]

Liu H, & Guo G (2015). Lifetime socioeconomic status, historical context, and genetic inheritance in shaping body mass in middle and late adulthood. American Sociological Review, 80, 705–737. [PubMed: 27231400]

Manski CF (2011). Genes, eyeglasses, and social policy. Journal of Economic Perspectives, 25(4), 83–93.

Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, … Kenny EE (2017). Human demographic history impacts genetic risk prediction across diverse populations. American Journal of Human Genetics, 100, 635–649. [PubMed: 28366442]

Mathews TJ, & Hamilton BE (2016). Mean age of mothers is on the rise: United States 2000–2014 (NCHS data brief). Hyattsville, MD: National Center for Health Statistics.

McEwen BS (2012). Brain on stress: How the social environment gets under the skin. Proceedings of the National Academy of Sciences, 109(Suppl. 2), 17180–17185.

McQueen MB, Boardman JD, Domingue BW, Smolen A, Tabor JW, Killeya-Jones L, … Mullen Harris K. (2015). The National Longitudinal Study of Adolescent to Adult Health (Add Health) sibling pairs genome-wide data. Behavior Genetics, 45, 12–23. [PubMed: 25378290]

Mendle J, Harden KP, Turkheimer E, Van Hulle CA, D'Onofrio BM, Brooks-Gunn J, … Lahey BB (2009). Associations between father absence and age of first sexual intercourse. Child Development, 80, 1463–1480. [PubMed: 19765012]

Mendle J, Ryan RM, & McKone KM (2015). Early childhood maltreatment and pubertal development: Replication in a population-based sample. Journal of Research on Adolescence, 26, 595–602. [PubMed: 28581653]

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, 533–542. [PubMed: 16756443]

Moffitt TE, Caspi A, Belsky J, & Silva PA (1992). Childhood experience and the onset of menarche: A test of a sociobiological model. Child Development, 63, 47–58. [PubMed: 1551329]

Moore SR, Harden KP, & Mendle J (2014). Pubertal timing and adolescent sexual behavior in girls. Developmental Psychology, 50, 1734–1745. [PubMed: 24588522]

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. [PubMed: 27225129]

Patton GC, McMorris BJ, Toumbourou JW, Hemphill SA, Donath S, & Catalano RF (2004). Puberty and the onset of substance use and abuse. Pediatrics, 114, 300–306.

Perry JR, Day F, Elks CE, Sulem P, & Thompson DJ, Ferreira T, … Ong KK (2014). Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. Nature, 514(7520), 92–97. [PubMed: 25231870]

Plomin R, DeFries JC, Knopik VS, & Neiderheiser JM (2013). Behavioral genetics. Basingstoke, UK: Palgrave Macmillan.

Polderman TJC, Benyamin B, de Leeuw CA, Sullivan PF, van Bochoven A, Visscher PM, & Posthuma D (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nature: Genetics, 47, 702–709. [PubMed: 25985137]

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, & Reich D (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nature: Genetics, 38, 904–909. [PubMed: 16862161]

Price AL, Zaitlen NA, Reich D, & Patterson N (2010). New approaches to population stratification in genome-wide association studies. Nature Reviews: Genetics, 11, 459–463.

Quinlan RJ (2003). Father absence, parental care, and female reproductive development. Evolution and Human Behavior, 24, 376–390.

Remsberg KE, Demerath EW, Schubert CM, Chumlea WC, Sun SS, & Siervogel RM (2005). Early menarche and the development of cardiovascular disease risk factors in adolescent girls: The Fels Longitudinal Study. Journal of Clinical Endocrinology & Metabolism, 90, 2718–2724. [PubMed: 15728207]

Rowe DC (2000). Environmental and genetic influences on pubertal development: Evolutionary life history traits? In Rodgers JL, Rowe DC, & Miller WB (Eds.), Genetic influences on human fertility and sexuality (pp. 147–168). Boston, MA: Kluwer Academic Publishers.

Rowe DC (2002). On genetic variation in menarche and age at first sexual intercourse: A critique of the Belsky-Draper hypotheses. Evolution and Human Behavior, 23, 365–372.

Ryan RM (2015). Nonresident fatherhood and adolescent sexual behavior: A comparison of siblings approach. Developmental Psychology, 51, 211–223. [PubMed: 25621757]

Sandler DP, Wilcox AJ, & Horney LF (1984). Age at menarche and subsequent reproductive events. American Journal of Epidemiology, 119, 765–774. [PubMed: 6720673]

Shifman S, Kuypers J, Kokoris M, Yakir B, & Darvasi A (2003). Linkage disequilibrium patterns of the human genome across populations. Human Molecular Genetics, 12, 771–776. [PubMed: 12651872]

StataCorp. (2015). Stata Statistical Software (Release 14) [Software]. College Station, TX: StataCorp LP.

Stearns SC (1992). The evolution of life histories. London, UK: Oxford University Press.

Stoll BA, Vatten LJ, & Kvinnsland S (1994). Does early physical maturity influence breast cancer risk? Acta Oncologica, 33, 171–176. [PubMed: 8204271]

Tamakoshi K, Yatsuya H, Tamakoshi A, & JACC Study Group. (2011). Early age at menarche associated with increased all-cause mortality. European Journal of Epidemiology, 26, 771–778. [PubMed: 22006230]

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

Towne B, Czerwinski SA, Demerath EW, Blangero J, Roche AF, & Siervogel RM (2005). Heritability of age at menarche in girls from the Fels Longitudinal Study. American Journal of Physical Anthropology, 128, 210–219. [PubMed: 15779076]

Trivers RL (1972). Parental investment and sexual selection In Campbell B (Ed.), Sexual selection & the descent of man 1871–1971 (pp. 136–179). New York, NY: Aldine de Gruyter.

Tropf FC, Lee SH, Verweij RM, Stulp G, van der Most PJ, de Vlaming R, … Mills MC (2017). Hidden heritability due to heterogeneity across seven populations. Nature: Human Behaviour, 1, 757–765.

Udry JR (2008). Age at menarche, at first intercourse, and at first pregnancy. Journal of Biosocial Science, 11, 433–441.

Udry JR, & Cliquet RL (1982). A cross-cultural examination of the relationship between ages at menarche, marriage, and first birth. Demography, 19, 53–63. [PubMed: 7067870]

Walter S, Mejía-Guevara I, Estrada K, Liu SY, & Glymour MM (2016). Association of a genetic risk score with body mass index across different birth cohorts. JAMA, 316, 63–69. [PubMed: 27380344]

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

Wojcik G, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, … Carlson CS (2017). Genetic diversity turns a new PAGE in our understanding of complex traits (Unpublished manuscript). Retrieved from 10.1101/188094

Wu LL, & Martinson BC (1993). Family structure and the risk of a premarital birth. American Sociological Review, 58, 210–232.
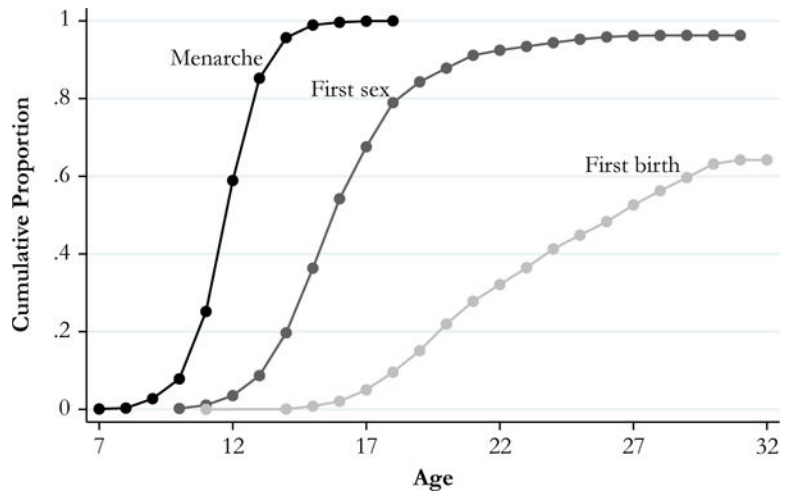
Fig. 1.
Reproductive timing: Unrelated non-Hispanic white females ($n = 2,681$). Mean time to event: menarche = 12.2; first sex = 17.2; and first birth = 26.2
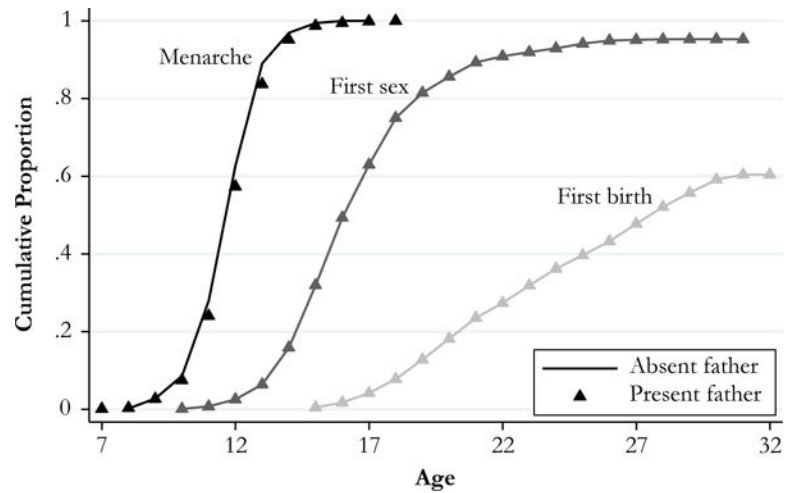
Fig. 2.
Reproductive timing by father absence: Unrelated non-Hispanic white females ($n = 2,631$).
Mantel-Haenszel log-rank test indicates that the differences between reproductive timing for females with absent and present fathers are statistically significant for all reproductive events ($p < .001$). Mean time to event: menarche, father absent = 12.1, father present = 12.3; first sex, father absent = 16.2, father present = 17.6; and first birth, father absent = 25.0, father present = 26.8
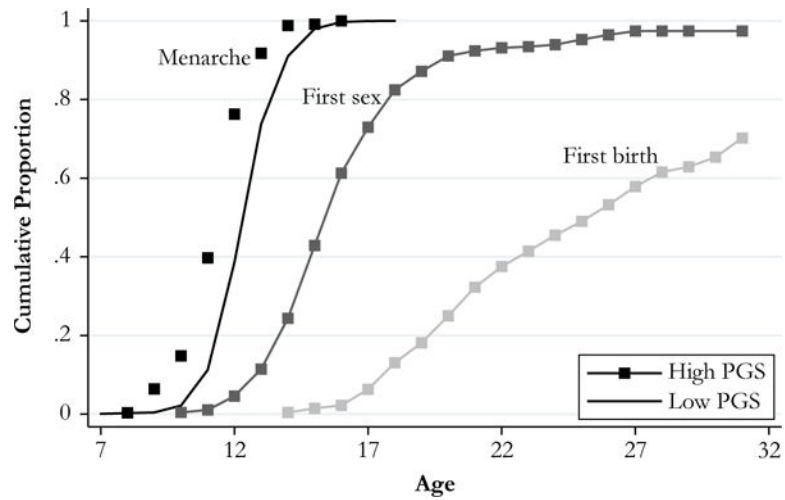
Fig. 3.
Reproductive timing by accelerated menarche polygenic score (PGS): Unrelated non-Hispanic white females ($n = 2,681$). High PGS 1 standard deviation above mean ($n = 441$); low PGS 1 standard deviation below mean ($n = 408$). Mantel-Haenszel log-rank test indicates that the difference between high- and low-PGS groups is statistically significant for menarche ($p < .001$) and first sex ($p < .05$). Mean time to event: menarche, high = 11.7, low = 12.8; first sex, high = 16.9, low = 17.2; and first birth, high = 25.4, low = 25.9
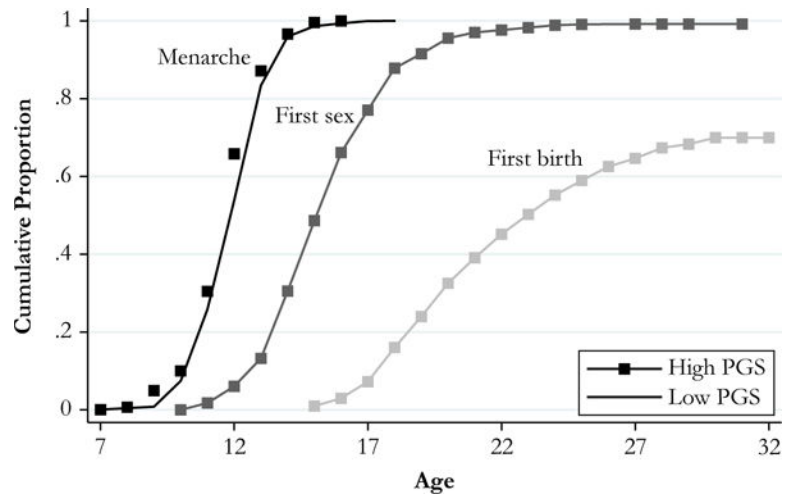
Fig. 4.
Reproductive timing by polygenic score (PGS) for earlier age at first birth: Unrelated non-Hispanic white females ($n = 2,681$). High PGS ≥ 1 standard deviation above mean ($n = 432$); low PGS ≤ 1 standard deviation below mean ($n = 418$). Mantel-Haenszel log-rank test indicates that the difference between high- and low-PGS groups is statistically significant for menarche ($p < .05$), first sex ($p < .001$), and first birth ($p < .001$). Mean time to event: menarche, high = 12.1, low = 12.3; first sex, high = 16.1, low = 18.3; and first birth, high = 24.8, low = 27.7
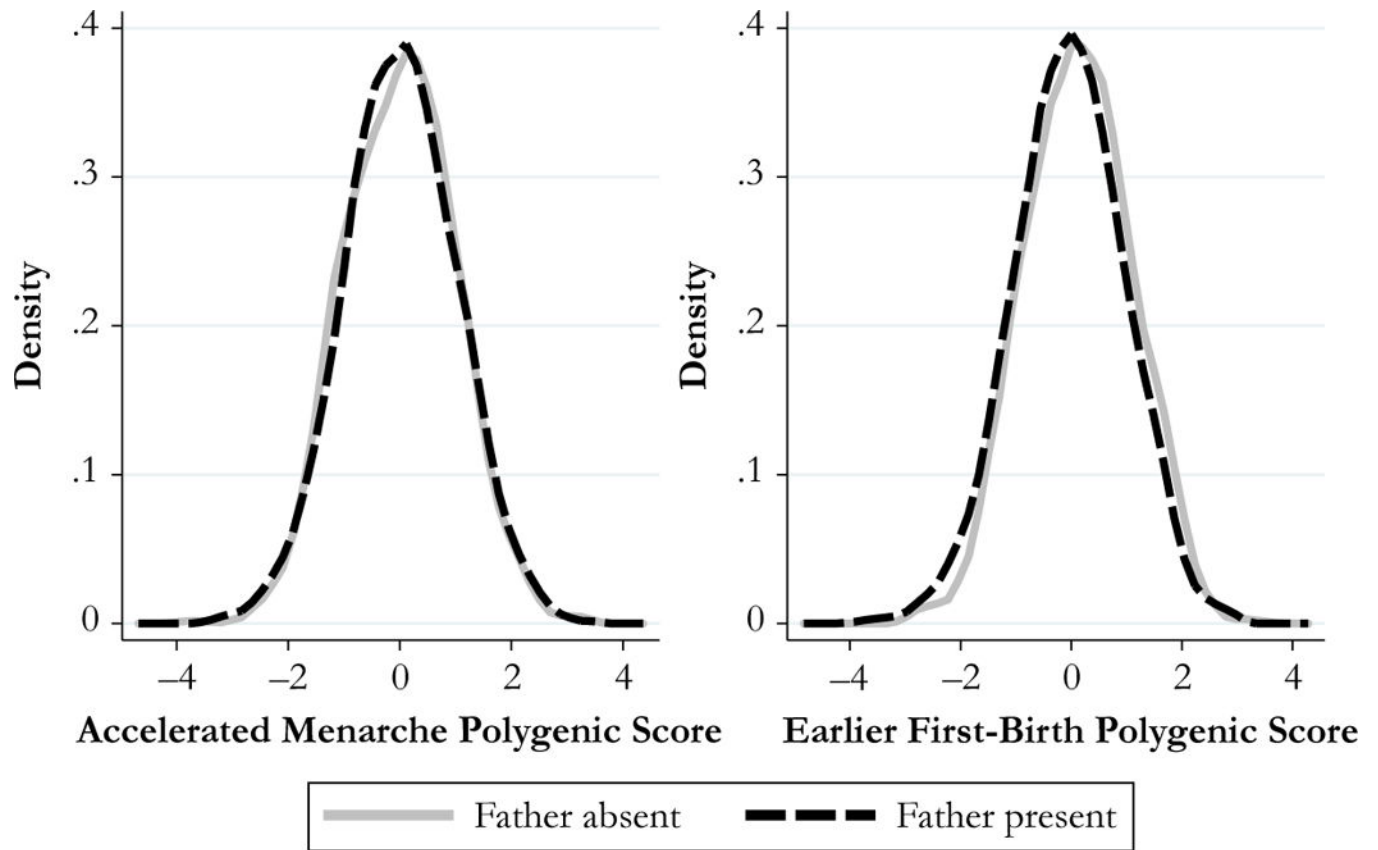
Fig. 5.
Polygenic score (PGS) distribution by father absence: Unrelated non-Hispanic white females (*n* = 2,681). The distribution of polygenic scores by father absence tests for rGE. The distribution of the accelerated menarche polygenic score is similar for girls in both groups. The distribution of the earlier first-birth polygenic score is shifted slightly to the right for girls who experience father absence

Table 1

Correlation matrix, unrelated non-Hispanic white females ($n = 2{,}681$)

	Mean (SD)/%	Age at Menarche	Age at First Sex	Age at First Birth	Menarche Polygenic Scores	Age at First Birth Polygenic Scores	Absent by Age 7
Age at Menarche	12.2 (1.4)						
Age at First Sex	16.7 (2.7)	.15 ***					
Age at First Birth	22.2 (3.4)	.12 ***	.37 ***				
Menarche Polygenic Scores	0 (1)	−.29 ***	−.06 **	−.04			
Age at First Birth Polygenic Scores	0 (1)	−.08 ***	−.19 ***	−.15 ***	.09 ***		
Absent by Age 7 (%)	31	−.08 ***	−.18 ***	−.12 ***	.00	.08 ***	

*
$p < .05$

**
$p < .01$

***
$p < .001$

Table 2

Hazard ratios for the risk of reproductive events and polygenic score (PGS) for age at menarche: Unrelated non-Hispanic white females

	Menarche		First Sex		First Birth	
Father Absent	1.24 ***	1.25 ***	1.49 ***	1.48 ***	1.52 ***	1.51 ***
	(0.05)	(0.05)	(0.06)	(0.06)	(0.08)	(0.08)
Menarche PGS	1.33 ***	1.33 ***	1.03	1.03	1.03	1.02
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)
Age	One-year intervals, 7–17		One-year intervals, 15–28		One-year intervals, 15–31	
Events	2,681		2,582		1,517	
N	2,681		2,681		2,681	

Notes: Results from complementary log-log hazard models for reproductive timing demonstrate independent and additive associations of father absence and the menarche polygenic score. Models control for yearly age intervals for the range specified. Standard errors are shown in parentheses. All models include appropriate sample weights and correct for clustered sampling design.

***
    p < .001

Table 3

Mean polygenic score (PGS) by father coresidence

	Menarche PGS	Age at First Birth PGS[***]
Father Absent ($n = 822$)	0.01 (0.03)	0.13 (0.03)
Father Present ($n = 1,859$)	0.02 (0.02)	−0.04 (0.02)

Note: Standard deviations are shown in parentheses.

[***]
$p < .001$ (*t* test for difference in means)

Table 4

Hazard ratios for the risk of reproductive events and polygenic score (PGS) for age at first birth: Unrelated non-Hispanic white females

	Menarche		First Sex		First Birth	
Father Absent	1.24 ***	1.23 ***	1.49 ***	1.47 ***	1.52 ***	1.47 ***
	(0.05)	(0.05)	(0.06)	(0.06)	(0.08)	(0.08)
Age at First Birth PGS	1.07 ***	1.07 ***	1.20 ***	1.20 ***	1.28 ***	1.26 ***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Age	One-year intervals 7–17		One-year intervals 15–28		One-year intervals 15–31	
Events	2,681		2,582		1,517	
N	2,681		2,681		2,681	

Notes: Results from complementary log-log hazard models for reproductive timing demonstrate independent and additive associations of father absence and the polygenic score for age at first birth. Models control for yearly age intervals for the range specified. Standard errors are shown in parentheses. All models include appropriate sample weights and correct for clustered sampling design.

***
 p < .001