Published in final edited form as:

Am J Obstet Gynecol. 2014 May ; 210(5): 398–405. doi:10.1016/j.ajog.2013.10.001.

The Contribution of Genetic and Environmental Factors to the Duration of Pregnancy

Timothy P. YORK, PhD^{1,2,3}, Lindon J. EAVES, PhD^{1,2,4}, Michael C. NEALE, PhD^{1,2,4}, and Jerome F. STRAUS III, MD, PhD^{2,3}

¹Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University School of Medicine, Richmond, Virginia, 23298, United States of America

²Department of Human and Molecular Genetics, Virginia Commonwealth University School of Medicine, Richmond, Virginia, 23298, United States of America

³Department of Obstetrics and Gynecology, Virginia Commonwealth University School of Medicine, Richmond, Virginia, 23298, United States of America

⁴Department of Psychiatry, Virginia Commonwealth University School of Medicine, Richmond, Virginia, 23298, United States of America

Abstract

This review describes how improvements in biometrical-genetic studies of twin kinships, half-sibships and cousinships have now demonstrated a sizeable fetal genetic and maternal genetic contribution to the spontaneous onset of labor. This is an important development since previous literature for the most part only reports an influence of the maternal genome. Current estimates of the percent of variation attributable to fetal genetic factors range from 11% to 35% while the range for the maternal genetic contribution is 13-20%. These same studies demonstrate an even larger influence of environmental sources over and above the influence of genetic sources and previously identified environmental risk factors. With these estimates in hand, a major goal for research on pregnancy duration is to identify specific allelic variation and environmental risk to account for this estimated genetic and environmental variation. A review of the current literature can serve as a guide for future research efforts.

Keywords

duration of pregnancy; gestational age at birth; preterm birth; gene; environment; genetic epidemiology

Address for correspondence: Timothy P. York, Ph.D. Virginia Institute for Psychiatric and Behavioral Genetics Department of Human and Molecular Genetics & OB/GYN Virginia Commonwealth University Box 980003, Richmond, Virginia, 23298-0003
tpyork@vcu.edu Phone: 804.828.8757 Fax: 804.827.0018.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclosure statement: The authors report no conflict of interest.

“Duration of pregnancy and intrauterine growth are dark areas in human biology”¹

Births less than 37 completed weeks of gestation are preterm, and account for the majority of perinatal mortality and morbidity.^{2, 3} Preterm birth has been shown to be associated with adverse outcomes early in life through adulthood.⁴⁻¹¹ In humans spontaneous preterm labor in infants without congenital abnormalities has been described as a culmination of a series of physiologic and anatomic changes in both the mother and fetus.¹²⁻¹⁵ Initial evidence for the identity of these factors was derived from reports that preterm birth is correlated among successive births of the same mother, and also between other familial relationships.^{1, 16} The etiological factors that explain these correlation patterns could be genetic factors shared among relatives, environmental factors shared within a family, or both.

Perhaps the most fascinating etiological aspect of the timing of birth is the consideration of both the fetal and maternal genomes, yet the interplay between the *in utero* environmental effect of maternal genotype and the developmental effect of fetal genotype presents challenges to the understanding of pathophysiologic mechanisms. On the other hand, environmental sources of variation can be partitioned into familial sources, i.e., those that influence all births of the same mother (e.g., socio-economic status), and those unique to individual pregnancies (e.g., infectious disease). Biometrical genetic theory and conceptual models developed over the past 100 years provide the framework to test competing hypotheses regarding genetic and environmental contributions to the onset of labor. Considering the magnitude of the public health impact of preterm birth, surprisingly few of these studies have been conducted to quantify these sources. Although these studies have provided compelling evidence for genetic influence on individual differences in the duration of pregnancy, many basic questions remain as to the nature of these sources, which are fundamental to continued research efforts. In this review we assess current genetic epidemiological findings in this area to address the following questions regarding the spontaneous onset of labor: 1) does the fetal and/or maternal genome contribute to differences in the duration of pregnancy?; 2) to what extent do environmental factors contribute differences in the duration of pregnancy and are these more/less important than genetic factors?; and 3) is there heterogeneity in the effect of these factors across self-identified racial/ethnic groups which could account for known differences in preterm birth rates?

Genetic Epidemiology and the Duration of Pregnancy

Birth outcomes research provides a fascinating model of inquiry, potentially involving direct effects of and interaction between the fetal and maternal genomes. Yet, before specific factors underlying the duration of pregnancy can be examined, standard practice in genetic epidemiology is to employ a top-down gene finding approach to first estimate the overall extent to which genetic factors influence the phenotype of interest -- whether these are of fetal origin, maternal origin, or both (Box 1). The classical twin design has been widely applied to estimate the contribution of genetic and environmental factors ranging from anthropomorphic traits to psychiatric outcomes.¹⁷ For studies of birth outcomes, these models have included extended family structures to avoid the biases inherent in twin births that restrict generalizability to singleton births. Including the offspring of twins, siblings and

half-siblings, allows for a wider array of hypotheses to be tested, and has been shown to increase statistical power.¹⁸ The information necessary to separate genetic from environmental effects is contained in the covariances of relationships that vary in their biological relatedness (see York *et. al.*¹⁹ for a listing of common expected covariances). As an example, cousins related through brothers on average share one-eighth of their genes and, therefore, a rough estimate of fetal genetic influences (since the mothers of these cousins are unrelated) can be derived by calculating the correlation between these collateral relatives. While inspection of familial correlations provides a useful summary of the magnitude of genetic and environmental influence, these parameters are typically estimated across several biological relationships simultaneously using structural equation modeling techniques.²⁰

Throughout this review we adopt the convention that preterm birth is a clinically meaningful threshold imposed on the continuum of gestational age. Other informative thresholds have been proposed, including very early preterm birth (births <32 weeks) and post-term birth (births >42 weeks). Unless otherwise demonstrated the choice of threshold does not necessarily imply anything fundamental about the *underlying genetic mechanisms*; it merely reflects a current judgment about when medical intervention is likely to be beneficial and cost-effective. Decisions about “where to draw the line” change as medical research obtains better information about the long-term consequences of particular trait levels. As understanding advances, so do the standards of assessment and the criteria for clinical intervention.

Genetic Contributions to the Duration of Pregnancy

Genetic variance is a function of allele frequencies and their effect sizes, both of which can vary across populations. Similarly, the influence of the environmental factors may differ between populations. Therefore, it should be cautioned that estimates of heritability from one population do not necessarily predict heritability in other populations. Since heritability is a ratio of variances (Box 1), and although the amount of genetic variance could be constant across populations, the total variance could differ, due to differences in the environmental variances, and result in population-specific heritability estimates. For example, human height has been shown to be highly heritable with estimates ranging from 87% to 93% across several Western countries.²¹ In poorer countries, the proportion due to environmental sources is larger resulting in lower heritabilities.²²

Heritability estimates for duration of pregnancy in European and European American samples support the influence of both fetal and maternal genomes (Table 1). The contribution of fetal genetic factors range from 11% to 35% while the range for the maternal genetic contribution is 13-20%. The Norwegian²³ and Swedish¹⁹ studies listed are consistent in their genetic proportions, while the US sample²⁴ of individuals self-described as European American report a much larger proportion of fetal genetic variance than the European samples. The interpretation of genetic proportions across studies is limited since the total variance of gestational age can differ between samples and would require the unstandardized parameter estimates to be reported on the same scale for each study. Nevertheless, these studies importantly demonstrate: 1) the consistency of both fetal and maternal genetic effects in explaining inter-individual differences in gestational age at birth,

and 2) that these estimates are as large or larger than environmental influences shared among births.

The two remaining studies listed in Table 1 partition the variance of gestational age at birth by dichotomizing the data at clinically meaningful thresholds. The first study²⁵ classifies births as preterm if born before 37 completed weeks and the second study²⁶ classifies births that have occurred after 41 completed weeks, or post-term. The results from these two studies differ considerably in the proportion of fetal genetic and familial environmental influence, which suggests that these factors may vary in their effects across the range of gestational age. While the attempt here is to diagnose different contributions at particular clinical definitions, it is not clear that moving the threshold provides additional information about genetic and environmental mechanisms under the assumption of a single underlying continuum of risk. This approach would generally lack the modeling framework to test whether variance components differ significantly at either end of the gestational age distribution. We have shown empirically, and by simulation studies, that imposing thresholds in the tails of the gestational age distribution results in genetic and environmental parameter estimates that vary widely due to a marked decrease in their precision, especially for more extreme thresholds.¹⁹

In epidemiological studies, there is a high price associated with imposing a threshold on an otherwise continuous phenotype such as gestational age at birth, as it provides no gain in information, and usually results in dramatically lower statistical power.^{27, 28}

Dichotomization classifies distinct groups of individuals on the phenotype of interest, arguing that there exists a natural boundary between these groups. While conceptually attractive, this position can be difficult to justify empirically, because it eliminates variation within groups. An alternative viewpoint suggests that individuals, rather than belonging to discrete groups, could be a “mixture” of groups. For instance, it is possible that the overall distribution of gestational age at birth is comprised of two underlying risk distributions, each describing a different model of genetic risk. This might involve an increase in genetic risk with increasing gestational age or a multiplicative model of risk. The point of this discussion is not to support one mechanism or the other, but rather that these models should not be assumed, but rather empirically tested.²⁹

Four other studies have provided support for genetic influences on the timing of birth, but are not included in Table 1 because they lack the research design or analytical framework to separate fetal from maternal genetic sources.³⁰⁻³³ Although births from a number of family structures are surveyed in these studies, the comparisons among relationships are often confounded by multiple genetic sources. For example, the relative risk of preterm birth between a woman and her maternal half-sister has been used to provide support for a maternal genetic effect. Yet, this relationship includes not only maternal genetic effects (i.e., $1/4 V_M$) but also the effect of the fetal genome (i.e., $1/16 V_F$). Furthermore, it is difficult to draw strong conclusions from such analyses since across most relationships there will be greater power to detect a maternal genetic effect versus a fetal genetic effect (i.e., since there is an inverse correlation between the size of the expected effect and the precision of the estimate).³⁴ These studies generally support the influence of the maternal genome, but do not provide strong evidence against the influence of the fetal genome.

Evidence for Fetal and Maternal Genetic Contributions to the Duration of Pregnancy

Large-scale genetic epidemiological studies presented in Table 1 have been provisionally supported by results from candidate gene association studies; the latter provide evidence that allelic variation from both fetal and maternal sources contribute to differences in the timing of birth (see review of 18 studies by Crider *et. al.*).³⁵ Yet, it is difficult to interpret this literature since most candidate gene studies contain a mix of small sample sizes, inconsistent phenotype definitions, lack of consideration for multiple testing issues, and the infrequent use of replication studies.^{36, 37} Additionally, the correlation between fetal and maternal genotypes makes interpretation of association findings difficult to ascribe to either source unless both the mother and the infant are genotyped.³⁸ Recently, studies have addressed a number of these shortcomings to provide some evidence of fetal genetic contributions to spontaneous^{39, 40} and indicated^{41, 42} preterm birth. Yet it remains to be seen whether replicable associations can be identified. A recent systematic study reviewed 189 polymorphisms in 89 genes reported as being associated with preterm birth.⁴³ The authors were able to perform meta-analyses on 36 gene variants and demonstrated that five gene variants had nominal significance but all displayed weak epidemiological credibility. Currently no robust genetic association findings have been identified that can account for current fetal or maternal heritability estimates.⁴⁴ Considering the history of gene finding for complex traits, it is very unlikely that piecemeal candidate gene studies is an efficient approach to account for the genetic variability expected from heritability estimates.⁴⁵

Uncovering the Genetic Architecture of the Duration of Pregnancy

As stated previously, geneticists often first employ a top-down approach to the task of gene identification for complex phenotypes by using twin and family methods to estimate the extent that genes are expected to contribute to phenotypic variance (Box 1). Once reasonably established, a bottom-up approach is then used to identify the causal genetic variants that comprise this variation, recently using whole-genome exploratory methods, such as genome-wide association studies (GWAS).⁴⁶ GWAS platforms utilize microarray technology to simultaneously assay approximately one million single nucleotide polymorphisms (SNPs) across the genome.⁴⁷ The selection of genetic variants to measure is based largely on the common-disease/common-variant hypothesis, which assumes that most genetic variation in complex phenotypes will be from relatively common allelic variation.⁴⁸ The results of the HapMap project provide evidence that this common variation can be reasonably measured using 500,000 to 1 million SNPs.^{47, 49} A major challenge involved with utilizing this hypothesis-free approach is the daunting task of screening hundreds of thousands of individual loci for effect sizes that are expected to account for a relatively small fraction of total phenotypic variation (1.1-1.5 fold increase in disease risk) while controlling for type I error. Very large sample sizes are often needed.

Since 2005, there have been 1,350 published GWAS of human disease and phenotypic traits⁵⁰, yet only one for preterm birth, which was based on 884 maternal cases.⁵¹ Design of future GWAS for the timing of birth should consider that the brief history of this approach has demonstrated that large sample sizes are essential for success.⁴⁵ While large samples are

needed to detect variants of small effect, expectations should be that these studies will initially identify only a subset of the loci, which explain only a fraction of the heritability.^{52, 53} For instance, a recent study found that 10-15% of variation in human height could be explained by 180 genetic variants in a combined sample of 133,653 individuals and a replication sample of 50,074 individuals.⁵⁴ A total of 32 variants accounted for only 1.45% phenotypic variation in body mass index (BMI) in a sample of 123,865 individuals, which was confirmed in a follow-up sample of 125,931 additional individuals.⁵⁵ It seems likely that gestational age at birth has a similar genetic architecture, because variants of large effect would be selected against due to reduced viability at the extremes of the distribution (i.e., stabilizing selection).

Although GWAS has provided initial evidence for thousands of genetic variants associated with a number of different phenotypes, it has yet to be shown whether this platform can appreciably explain initial estimates of heritability⁵⁶, and furthermore provide predictive value for disease risk.⁴⁷ Yet, the main advantage of this hypothesis-free approach is that it is not restricted by our limited understanding of the biological underpinnings of parturition as required for candidate gene studies.⁴⁵ Finding even a small number of robustly associated variants could greatly further our knowledge of biological factors causally related to the onset of labor. The application of this technology should be guided by consistent phenotyping practices (see Pennell et. al., 2007)⁵⁷ and the acquisition of large sample sizes could be achieved through consortium efforts (i.e., the height and BMI examples previously mentioned produced robust results from a meta-analysis of 46 studies). As most geneticists now agree that common variation will not capture all aspects of the genetic architecture of complex phenotypes, technological advances in next-generation sequencing will allow for the systematic identification of rare moderate-risk loci.^{48, 58}

Environmental Control of the Duration of Pregnancy

Studies to date have quantified numerous maternal risk factors for preterm birth. These have included previous preterm birth, race/ethnicity, primipara status, access to or use of prenatal care, marital status, maternal under/overweight, smoking during pregnancy, maternal age, educational achievement and socio-economic status, among others. Although the estimates listed in Table 1 adjust for a number of these risk factors, these results clearly support the continued contribution of familial and pregnancy-specific environmental variance to gestational age at birth above and beyond this measured risk. The pregnancy-specific environmental variance is consistently the largest, and accounts for approximately one-half the total variance across all studies listed. Although this term usually, in part, comprises errors of measurement, the sheer magnitude of this effect after common risk factors have been accounted for suggests that many more environmental risk factors have yet to be identified. These sources of variation have been shown to change minimally between adjusted and unadjusted values^{19, 24}, and corroborate other studies that have documented the poor performance of socioeconomic models to predict racial differences in birth outcomes.⁵⁹⁻⁶¹

It is important to note that estimates of environmental variance are separate from those of the additive contribution of genes. This point is particularly salient when considering that

the large racial disparity in preterm birth rates between African Americans and European Americans, approximately 18% and 11.5% respectively, remains largely unexplained by measured risk factors.⁵⁹⁻⁶¹ Thus, the failure to explain this difference based on socioeconomic models does not automatically imply that the disparity can be explained by allelic differences that exists between racial categories.

Our investigations in Virginia show that a sizeable proportion of environmental variance remains in births from either African American or European American populations,²⁴ even after accounting for both genetic and measured environmental risk factors. This study reports that the total variance in gestational age at birth is nearly twice as high in African Americans compared to European Americans and that this difference is largely explained by environmental sources, which are 3.1 times greater in African Americans (Figure 1). More specifically, the majority of this difference is accounted for by environmental factors that changed between pregnancies, as opposed to environmental sources that influenced all pregnancies of the same mother. These results are supported by multiple lines of evidence that demonstrate the primacy of social inequities for racial health disparities including preterm birth rates.⁶² Taken together, these results suggest that significant efforts are needed to identify novel socio-demographic and other environmental influences that contribute to the timing of birth (e.g., new vaginal microorganisms). The evidence that those factors are less stable across births of the same mother may provide clues to the marked increase in the African American preterm birth rate.

The influence of environmental heterogeneity known to exist between self-identified racial groups on birth outcomes could, in part, be mediated by epigenetic changes in tissues of the fetus and mother.⁶³ There is substantial evidence supporting the influence of environmental exposures during development driving epigenetic plasticity.⁶⁴⁻⁷¹ Epigenetics refers to functionally relevant modifications of chromatin (i.e., DNA and associated proteins) that do not involve changes in nucleotide sequence and lead to modifications that can be transmitted to daughter cells. The two epigenetic mechanisms frequently studied are chromatin condensation and DNA methylation; these work together to influence the transcription of genes.⁷² Previous work has found that genetic and epigenetic mechanisms combine to control matrix metalloproteinase-1 (MMP-1) expression in fetal amnion tissue, and that hypomethylation of the MMP-1 promoter was associated with preterm premature rupture of membranes in African American women.⁷³ Considering that recent findings provide preliminary evidence for racial differences in DNA methylation between African American and Caucasian newborns⁷⁴ alongside significant associations of fetal DNA methylation with spontaneous preterm birth⁷⁵ further positions DNA methylation as a possible mechanism to account for racial differences in the duration of pregnancy.

Conclusions

The primary goal of genetic studies of gestational age at birth is to identify biological pathways that account for inter-individual differences.⁷⁶ This information would provide a better understanding of parturition, and could be exploited to develop predictive models and therapies to reduce the incidence of preterm births. The recent large-scale genetic epidemiological studies reviewed provide evidence that both fetal and maternal genomes

likely contain alleles that confer risk and protection to preterm birth and DNA from both sources should be included in future gene-mapping studies. This inclusion will also provide the opportunity to assess the contribution of fetal-maternal genetic epistatic interactions (where the expression of one gene depends upon the presence of modifier genes), as well as to resolve questions regarding genotype incompatibility.⁷⁷ Additional studies are needed to estimate heritability in populations of non-European ancestry. Currently, there are relatively few robust genetic association findings and these account for very little of the fetal or maternal heritability estimates. Considering the history of gene finding for complex traits, it is doubtful that piecemeal candidate gene studies will provide an efficient means to capture this variability.⁴⁵ GWAS, however, has the advantage of providing a hypothesis-free approach to discover novel etiological factors that challenge our current understanding of mechanisms that initiate labor.

While evidence exists across studies that both fetal and maternal genetic factors explain a modest proportion of individual differences in the duration of pregnancy, there is no obligatory requirement that these same genetic factors should operate to explain racial differences. Our studies indicate the majority of this difference is due to the effect of greater environmental heterogeneity in African Americans which generates larger differences among successive births to the same mother, in contrast to environmental influences that would create stability across sibling births.²⁴ Yet, genetic studies of admixed populations, such as African Americans, can profit by incorporating admixture association signals into GWAS as the risk for preterm birth has been shown to differ depending on ancestry.^{42, 78-81}

In summary, investment in GWAS as a proven technology to identify common allelic variation should be considered in large fetal and maternal samples. This exploratory platform can provide valuable insight to the genetic architecture of gestational age and a frame of reference for the continued use of high-throughput sequencing technologies. Measured genotypes in fetal and maternal samples may also facilitate future causal analyses of the developmental consequences of preterm birth.^{82, 83} Indeed, a current limitation in biomedical research is the disconnect between studies of biological process and those of environmental exposure.⁸⁴ The covariates usually included in studies of birth outcomes leave a significant amount of environmental variance unexplained, and these non-genetic sources are proportionally as large or larger than genetic sources. A greater investment of resources is needed to identify environmental exposures that account for the increased environmental heterogeneity across births, particularly within African American families. Future research may also profit from considering whether epigenetic signals play a role in the molecular chain of events linking environmental exposures to differential gene expression.

Acknowledgments

This work was supported by National Institutes of Health grants P60MD002256 (T.P.Y, J.F.S), R01HD034612 (J.F.S.) and HD073555 (M.C.N.).

Box 1**Estimation of heritability for studies of gestational age at birth**

Variation in phenotypic scores, in this case gestational age at birth (V_{GA}), can plausibly be partitioned into the contribution of both genetic (V_G) and environmental (V_E) factors and estimated in a general model as:

$$V_{GA} = V_G + V_E \quad (1)$$

Thus, the *observed* phenotypic variation can be attributed to the sum of *unobserved* genetic and environmental factors. This is the starting point for a top-down approach to gene finding in that it provides answers the basic questions of whether genes are expected to contribute to phenotypic variation and the expected magnitude of this effect. It is the ultimate goal of geneticists to account for these estimates by the aggregate variation from measured alleles. The genetic component can be further decomposed into both fetal genetic (V_F) and maternal genetic (V_M) components as:

$$V_G = V_F + V_M \quad (2)$$

In a similar way, the environmental component can be decomposed into that explained by environmental factors common to all pregnancies of the same mother (V_C), and that unique to each pregnancy (V_U), as:

$$V_E = V_C + V_U \quad (3)$$

The V_C estimate comprises all non-genetic factors that make phenotypes of siblings and half-siblings similar. V_U on the other hand encompasses those environmental factors that do not contribute to phenotypic similarity. V_C would include features of the common environment not influenced by maternal genetic sources shared among all pregnancies of the same mother that, for instance, could include aspects of the uterine environment. This assumes there is no covariance between the effects of the fetal and maternal genotypes (i.e., “passive genotype-environment correlation”), which can be relaxed in extensions of this model.⁸⁵ Thus, the total variance of gestational age at birth can be summarized as:

$$V_{GA} = V_F + V_M + V_C + V_U \quad (4)$$

Under this variance decomposition errors of measurement would be included in the V_U term. Heritability is thus defined as a ratio of variances and is the proportion of genetic variance in the total trait variation. The proportion of total variance in gestational age attributed to fetal genetic sources is estimated as, V_F / V_{GA} . Typically, model assumptions include: random mating; genetic effects are additive and constant over pregnancies; genetic and environmental factors do not interact; the influence of fetal and maternal genetic factors are the same for male and female fetuses and; environmental effects are pregnancy specific apart from the effects of maternal genotype, environmental effects shared across pregnancies and other aspects of the parental phenotype such as

cultural inheritance.²⁴ These assumptions can be relaxed in extensions of this general model.

References

1. Magnus P, Bakketeig LS, Skjaerven R. Correlations of birth weight and gestational age across generations. *AnnHumBiol.* 1993; 20:231–38.
2. Goldenberg RL. The management of preterm labor. *Obstet Gynecol.* 2002; 100:1020–37. [PubMed: 12423870]
3. Lawn JE, Wilczynska-Ketende K, Cousens SN. Estimating the causes of 4 million neonatal deaths in the year 2000. *Int J Epidemiol.* 2006; 35:706–18. [PubMed: 16556647]
4. Narang I. Review series: What goes around, comes around: childhood influences on later lung health? Long-term follow-up of infants with lung disease of prematurity. *Chron Respir Dis.* 2010; 7:259–69. [PubMed: 21084550]
5. Crump C, Sundquist K, Winkleby MA, Sundquist J. Preterm birth and risk of epilepsy in Swedish adults. *Neurology.* 2011; 77:1376–82. [PubMed: 21968843]
6. Wehkalampi K, Hovi P, Dunkel L, et al. Advanced pubertal growth spurt in subjects born preterm: the Helsinki study of very low birth weight adults. *The Journal of clinical endocrinology and metabolism.* 2011; 96:525–33. [PubMed: 21147886]
7. Santos IS, Matijasevich A, Domingues MR, Barros AJ, Victora CG, Barros FC. Late preterm birth is a risk factor for growth faltering in early childhood: a cohort study. *BMC Pediatr.* 2009; 9:71. [PubMed: 19917121]
8. Crump C, Winkleby MA, Sundquist K, Sundquist J. Risk of hypertension among young adults who were born preterm: a Swedish national study of 636,000 births. *Am J Epidemiol.* 2011; 173:797–803. [PubMed: 21320866]
9. Manzardo AM, Madarasz WV, Penick EC, et al. Effects of premature birth on the risk for alcoholism appear to be greater in males than females. *J Stud Alcohol Drugs.* 2011; 72:390–8. [PubMed: 21513675]
10. Swamy GK, Ostbye T, Skjaerven R. Association of preterm birth with long-term survival, reproduction, and next-generation preterm birth. *JAMA : the journal of the American Medical Association.* 2008; 299:1429–36. [PubMed: 18364485]
11. Skilton MR, Viikari JS, Juonala M, et al. Fetal growth and preterm birth influence cardiovascular risk factors and arterial health in young adults: the Cardiovascular Risk in Young Finns Study. *Arterioscler Thromb Vasc Biol.* 2011; 31:2975–81. [PubMed: 21940950]
12. Lye, SJ.; Tsui, P.; Dong, X., et al. Myometrial programming: a new concept underlying the regulation of myometrial function during pregnancy. Informa UK Ltd; Oxon: 2007.
13. Chaudhari BP, Plunkett J, Ratajczak CK, Shen TT, DeFranco EA, Muglia LJ. The genetics of birth timing: insights into a fundamental component of human development. *Clin Genet.* 2008; 74:493–501. [PubMed: 19037974]
14. Mendelson CR. Minireview: fetal-maternal hormonal signaling in pregnancy and labor. *Mol Endocrinol.* 2009; 23:947–54. [PubMed: 19282364]
15. Smith R, Paul J, Maiti K, Tolosa J, Madsen G. Recent advances in understanding the endocrinology of human birth. *Trends Endocrinol Metab.* 2012; 23:516–23. [PubMed: 22682480]
16. Clausson B, Lichtenstein P, Cnattingius S. Genetic influence on birthweight and gestational length determined by studies in offspring of twins. *BJOG.* 2000; 107:375–81. [PubMed: 10740335]
17. van Dongen J, Slagboom PE, Draisma HH, Martin NG, Boomsma DI. The continuing value of twin studies in the omics era. *Nat Rev Genet.* 2012; 13:640–53. [PubMed: 22847273]
18. Posthuma D, Boomsma DI. A note on the statistical power in extended twin designs. *Behav Genet.* 2000; 30:147–58. [PubMed: 10979605]
19. York TP, Eaves LJ, Lichtenstein P, et al. Fetal and maternal genes' influence on gestational age in a quantitative genetic analysis of 244,000 Swedish births. *Am J Epidemiol.* 2013; 178:543–50. [PubMed: 23568591]

20. Neale, MC.; Cardon, LR. *Methodology for Genetic Studies of Twins and Families*. Kluwer Academic Publishers; Dordrecht: 1992.
21. Silventoinen K, Sammalisto S, Perola M, et al. Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin research : the official journal of the International Society for Twin Studies*. 2003; 6:399–408. [PubMed: 14624724]
22. Silventoinen K. Determinants of variation in adult body height. *J Biosoc Sci*. 2003; 35:263–85. [PubMed: 12664962]
23. Lunde A, Melve KK, Gjessing HK, Skjaerven R, Irgens LM. Genetic and environmental influences on birth weight, birth length, head circumference, and gestational age by use of population-based parent-offspring data. *Am J Epidemiol*. 2007; 165:734–41. [PubMed: 17311798]
24. York TP, Strauss JF 3rd, Neale MC, Eaves LJ. Racial differences in genetic and environmental risk to preterm birth. *PLoS ONE*. 2010; 5:e12391. [PubMed: 20811627]
25. Svensson AC, Sandin S, Cnattingius S, et al. Maternal effects for preterm birth: a genetic epidemiologic study of 630,000 families. *Am J Epidemiol*. 2009; 170:1365–72. [PubMed: 19854802]
26. Oberg AS, Frisell T, Svensson AC, Iliadou AN. Maternal and fetal genetic contributions to postterm birth: familial clustering in a population-based sample of 475,429 Swedish births. *Am J Epidemiol*. 2013; 177:531–7. [PubMed: 23425630]
27. Neale MC, Eaves LJ, Kendler KS. The power of the classical twin study to resolve variation in threshold traits. *Behav Genet*. 1994; 24:239–58. [PubMed: 7945154]
28. Ragland DR. Dichotomizing continuous outcome variables: dependence of the magnitude of association and statistical power on the cutpoint. *Epidemiology*. 1992; 3:434–40. [PubMed: 1391136]
29. MacCallum RC, Zhang S, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. *Psychol Methods*. 2002; 7:19–40. [PubMed: 11928888]
30. Boyd HA, Poulsen G, Wohlfahrt J, Murray JC, Feenstra B, Melbye M. Maternal contributions to preterm delivery. *Am J Epidemiol*. 2009; 170:1358–64. [PubMed: 19854807]
31. Wilcox AJ, Skjaerven R, Lie RT. Familial patterns of preterm delivery: maternal and fetal contributions. *Am J Epidemiol*. 2008; 167:474–9. [PubMed: 18048376]
32. Treloar SA, Macones GA, Mitchell LE, Martin NG. Genetic influences on premature parturition in an Australian twin sample. *TwinRes*. 2000; 3:80–82.
33. Plunkett J, Feitosa MF, Trusgnich M, et al. Mother's genome or maternally-inherited genes acting in the fetus influence gestational age in familial preterm birth. *Hum Hered*. 2009; 68:209–19. [PubMed: 19521103]
34. Visscher PM, Hill WG, Wray NR. Heritability in the genomics era--concepts and misconceptions. *Nat Rev Genet*. 2008; 9:255–66. [PubMed: 18319743]
35. Crider KS, Whitehead N, Buus RM. Genetic variation associated with preterm birth: a HuGE review. *Genetics in medicine : official journal of the American College of Medical Genetics*. 2005; 7:593–604. [PubMed: 16301860]
36. Anum EA, Springel EH, Shriver MD, Strauss JF 3rd. Genetic Contributions to Disparities in Preterm Birth. *Pediatr Res*. 2009; 65:1–9. [PubMed: 18787421]
37. Plunkett J, Muglia LJ. Genetic contributions to preterm birth: implications from epidemiological and genetic association studies. *Ann Med*. 2008; 40:167–95. [PubMed: 18382883]
38. Hill LD, York TP, Kusanovic JP, et al. Epistasis between COMT and MTHFR in maternal-fetal dyads increases risk for preeclampsia. *PLoS ONE*. 2011; 6:e16681. [PubMed: 21304959]
39. Haataja R, Karjalainen MK, Luukkonen A, et al. Mapping a new spontaneous preterm birth susceptibility gene, IGF1R, using linkage, haplotype sharing, and association analysis. *PLoS Genet*. 2011; 7:e1001293. [PubMed: 21304894]
40. Karjalainen MK, Huusko JM, Ulvila J, et al. A potential novel spontaneous preterm birth gene, AR, identified by linkage and association analysis of X chromosomal markers. *PLoS ONE*. 2012; 7:e51378. [PubMed: 23227263]
41. Mann PC, Cooper ME, Ryckman KK, et al. Polymorphisms in the fetal progesterone receptor and a calcium-activated potassium channel isoform are associated with preterm birth in an Argentinian

- population. *Journal of perinatology : official journal of the California Perinatal Association*. 2013; 33:336–40. [PubMed: 23018797]
42. Wang H, Parry S, Macones G, et al. A functional SNP in the promoter of the SERPINH1 gene increases risk of preterm premature rupture of membranes in African Americans. *Proc Natl Acad Sci U S A*. 2006; 103:13463–7. [PubMed: 16938879]
 43. Dolan SM, Hollegaard MV, Merialdi M, et al. Synopsis of preterm birth genetic association studies: the preterm birth genetics knowledge base (PTBGene). *Public Health Genomics*. 2010; 13:514–23. [PubMed: 20484876]
 44. Bezold KY, Karjalainen MK, Hallman M, Teramo K, Muglia LJ. The genomics of preterm birth: from animal models to human studies. *Genome Med*. 2013; 5:34. [PubMed: 23673148]
 45. Kim Y, Zerwas S, Trace SE, Sullivan PF. Schizophrenia genetics: where next? *Schizophr Bull*. 2011; 37:456–63. [PubMed: 21505112]
 46. Yang J, Benyamin B, McEvoy BP, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet*. 2010; 42:565–9. [PubMed: 20562875]
 47. Ku CS, Loy EY, Pawitan Y, Chia KS. The pursuit of genome-wide association studies: where are we now? *J Hum Genet*. 2010; 55:195–206. [PubMed: 20300123]
 48. Singleton AB, Hardy J, Traynor BJ, Houlden H. Towards a complete resolution of the genetic architecture of disease. *Trends in genetics : TIG*. 2010; 26:438–42. [PubMed: 20813421]
 49. Eberle MA, Ng PC, Kuhn K, et al. Power to detect risk alleles using genome-wide tag SNP panels. *PLoS Genet*. 2007; 3:1827–37. [PubMed: 17922574]
 50. Hindorf LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A*. 2009; 106:9362–7. [PubMed: 19474294]
 51. Uzun A, Dewan AT, Istrail S, Padbury JF. Pathway-based genetic analysis of preterm birth. *Genomics*. 2013; 101:163–70. [PubMed: 23298525]
 52. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009; 461:747–53. [PubMed: 19812666]
 53. Button KS, Ioannidis JP, Mokrysz C, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 2013; 14:365–76. [PubMed: 23571845]
 54. Lango Allen H, Estrada K, Lettre G, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*. 2010; 467:832–8. [PubMed: 20881960]
 55. Speliotes EK, Willer CJ, Berndt SI, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet*. 2010; 42:937–48. [PubMed: 20935630]
 56. Maher B. Personal genomes: The case of the missing heritability. *Nature*. 2008; 456:18–21. [PubMed: 18987709]
 57. Pennell CE, Jacobsson B, Williams SM, et al. 2007; 196:107–18.
 58. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet*. 2010; 11:415–25. [PubMed: 20479773]
 59. Kaufman JS, Cooper RS, McGee DL. Socioeconomic status and health in blacks and whites: the problem of residual confounding and the resiliency of race. *Epidemiology*. 1997; 8:621–8. [PubMed: 9345660]
 60. Lu MC, Chen B. Racial and ethnic disparities in preterm birth: the role of stressful life events. *AmJObstetGynecol*. 2004; 191:691–99.
 61. Goldenberg RL, Cliver SP, Mulvihill FX, et al. Medical, psychosocial, and behavioral risk factors do not explain the increased risk for low birth weight among black women. *Am J Obstet Gynecol*. 1996; 175:1317–24. [PubMed: 8942508]
 62. Gravlee CC. How race becomes biology: embodiment of social inequality. *Am J Phys Anthropol*. 2009; 139:47–57. [PubMed: 19226645]
 63. Burris HH, Collins JW Jr. Race and preterm birth--the case for epigenetic inquiry. *Ethn Dis*. 2010; 20:296–9. [PubMed: 20828105]
 64. McGowan PO, Sasaki A, D'Alessio AC, et al. Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nat Neurosci*. 2009; 12:342–8. [PubMed: 19234457]

65. McGowan PO, Meaney MJ, Szyf M. Diet and the epigenetic (re)programming of phenotypic differences in behavior. *Brain Res.* 2008; 1237:12–24. [PubMed: 18694740]
66. Waterland RA. Early environmental effects on epigenetic regulation in humans. *Epigenetics.* 2009; 4:523–5. [PubMed: 19923895]
67. Rutten BP, Mill J. Epigenetic mediation of environmental influences in major psychotic disorders. *Schizophr Bull.* 2009; 35:1045–56. [PubMed: 19783603]
68. Champagne FA. Epigenetic influence of social experiences across the lifespan. *Dev Psychobiol.* 2010; 52:299–311. [PubMed: 20175106]
69. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet.* 2003; 33(Suppl):245–54. [PubMed: 12610534]
70. Miller G, Chen E, Cole SW. Health psychology: developing biologically plausible models linking the social world and physical health. *Annu Rev Psychol.* 2009; 60:501–24. [PubMed: 19035829]
71. Hunter RG, McEwen BS. Stress and anxiety across the lifespan: structural plasticity and epigenetic regulation. *Epigenomics.* 2013; 5:177–94. [PubMed: 23566095]
72. Dolinoy DC, Weidman JR, Jirtle RL. Epigenetic gene regulation: linking early developmental environment to adult disease. *Reprod Toxicol.* 2007; 23:297–307. [PubMed: 17046196]
73. Wang H, Ogawa M, Wood JR, et al. Genetic and epigenetic mechanisms combine to control MMP1 expression and its association with preterm premature rupture of membranes. *Hum Mol Genet.* 2008; 17:1087–96. [PubMed: 18178580]
74. Adkins RM, Krushkal J, Tykavsky FA, Thomas F. Racial differences in gene-specific DNA methylation levels are present at birth. *Birt Defects Res A Clin Mol Teratol.* 2011; 91:728–36.
75. Pares SE, Conneely KN, Kilaru V, et al. Fetal DNA Methylation Associates with Early Spontaneous Preterm Birth and Gestational Age. *PLoS ONE.* 2013; 8:e67489. [PubMed: 23826308]
76. Romero R, Espinoza J, Gotsch F, et al. The use of high-dimensional biology (genomics, transcriptomics, proteomics, and metabolomics) to understand the preterm parturition syndrome. *BJOG : an international journal of obstetrics and gynaecology.* 2006; 113(Suppl 3):118–35. [PubMed: 17206980]
77. Parimi N, Tromp G, Kuivaniemi H, et al. Analytical approaches to detect maternal/fetal genotype incompatibilities that increase risk of pre-eclampsia. *BMC Med Genet.* 2008; 9:60. [PubMed: 18598365]
78. Manuck TA, Lai Y, Meis PJ, et al. Admixture mapping to identify spontaneous preterm birth susceptibility loci in African Americans. *Obstet Gynecol.* 2011; 117:1078–84. [PubMed: 21508746]
79. Pasaniuc B, Zaitlen N, Lettre G, et al. Enhanced statistical tests for GWAS in admixed populations: assessment using African Americans from CARE and a Breast Cancer Consortium. *PLoS Genet.* 2011; 7:e1001371. [PubMed: 21541012]
80. Kang G, Gao G, Shete S, et al. 2011; 2
81. Tsai HJ, Yu Y, Zhang S, et al. Association of genetic ancestry with preterm delivery and related traits among African American mothers. *Am J Obstet Gynecol.* 2009; 201(94):e1–10.
82. Glymour MM, Tchetgen EJ, Robins JM. Credible Mendelian randomization studies: approaches for evaluating the instrumental variable assumptions. *Am J Epidemiol.* 2012; 175:332–9. [PubMed: 22247045]
83. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol.* 2003; 32:1–22. [PubMed: 12689998]
84. Kramer MR, Hogue CR. What causes racial disparities in very preterm birth? A biosocial perspective. *Epidemiol Rev.* 2009; 31:84–98. [PubMed: 19477907]
85. York TP, Strauss JF 3rd, Neale MC, Eaves LJ. Estimating fetal and maternal genetic contributions to premature birth from multiparous pregnancy histories of twins using MCMC and maximum-likelihood approaches. *Twin Res Hum Genet.* 2009; 12:333–42. [PubMed: 19653833]

Figure 1.

Race-specific genetic and environmental variance component estimates with 95% confidence intervals from York *et. al.*, PLoS ONE 2010;5:e12391.²² AA = African American, EA = European American.

Table 1

Summary of Standardized Genetic and Environmental Estimates for Birth Outcomes from Studies Using Population Samples of European Ancestry.

Study	Outcome	Sample Location	Sample Years	No. Births	Fetal Genetic (95% CI)	Maternal Genetic (95% CI)	Familial Environment (95% CI)	Pregnancy-specific Environment (95% CI)
Lunde et al., 2007	Gestational Age at Birth ^a	Norway	1967-2004	328,809	0.11 (0.10,0.13)	0.14 (0.13,0.16)	0.13 (0.12,0.15)	0.61 (0.60,0.62)
York et al., 2010	Gestational Age at Birth ^b	United States	1989-2008	603,092	0.35 (0.18,0.52)	0.13 (0.06,0.20)	0.07 (0.01,0.14)	0.45 (0.36,0.53)
York et al., 2013	Gestational Age at Birth ^b	Sweden	1987-2008	244,056	0.13 (0.68,0.18)	0.23 (0.21,0.25)	0.10 (0.07,0.13)	0.56 (0.53,0.59)
Svensson et al., 2009	Pre-term Birth ^b	Sweden	1992-2004	536,637	0.05 (0.0,0.23)	0.25 (0.23,0.27)	0.18 (0.16,0.20)	0.52 (0.41,0.58)
Oberg et al., 2013	Post-term Birth	Sweden	1992-2004	475,429	0.26 (0.13,0.35)	0.21 (0.16,0.26)	0.02 (0.0,0.07)	0.51 (not reported)

^aExcluded births from cesarean section and extremely growth-restricted infants (<1,400 grams).

^bSpontaneous onset births considered. York *et. al.*¹⁹ and Svensson *et. al.*²⁵ report small differences in parameter estimates when indicated births are omitted.