



# A statistical framework for testing the causal effects of fetal drive

Nianjun Liu<sup>1</sup>, Edward Archer<sup>2</sup>, Vinodh Srinivasasainagendra<sup>1</sup> and David B. Allison<sup>1,2\*</sup>

<sup>1</sup> Department of Biostatistics, School of Public Health, University of Alabama at Birmingham, Birmingham, AL, USA

<sup>2</sup> Office of Energetics, School of Public Health, University of Alabama at Birmingham, Birmingham, AL, USA

## Edited by:

Lisa J. Martin, Cincinnati Children's Hospital Medical Center, USA

## Reviewed by:

Yiran Guo, Children's Hospital of Philadelphia, USA

Ge Zhang, Cincinnati Children's Hospital Medical Center, USA

Lisa J. Martin, Cincinnati Children's Hospital Medical Center, USA

## \*Correspondence:

David B. Allison, Office of Energetics and Department of Biostatistics, School of Public Health, University of Alabama at Birmingham, Ryals Building, Room 140J, 1665 University Boulevard, Birmingham, AL 35294, USA

e-mail: dallison@uab.edu

Maternal genetic and phenotypic characteristics (e.g., metabolic and behavioral) affect both the intrauterine milieu and lifelong health trajectories of their fetuses. Yet at the same time, fetal genotype may affect processes that alter pre and postnatal maternal physiology, and the subsequent health of both fetus and mother. We refer to these latter effects as 'fetal drive.' If fetal genotype is driving physiologic, metabolic, and behavioral phenotypic changes in the mother, there is a possibility of differential effects with different fetal genomes inducing different long-term effects on both maternal and fetal health, mediated through intrauterine environment. This proposed mechanistic path remains largely unexamined and untested. In this study, we offer a statistical method to rigorously test this hypothesis and make causal inferences in humans by relying on the (conditional) randomization inherent in the process of meiosis. For illustration, we apply this method to a dataset from the Framingham Heart Study.

**Keywords: genetics, fetal effects, human, statistics, causal inference, intrauterine environment**

## INTRODUCTION

A common belief is that genetic, phenotypic, and behavioral characteristics of mothers affect the metabolic milieu during pregnancy, and that this in turn affects both the short and long-term metabolic health of the offspring, including the risk of obesity. This is likely true and some evidence is supportive (Catalano et al., 1998; Plagemann, 2005). Yet there is a complementary possibility that has gone largely unconsidered and untested. Namely, that the genotype of the offspring may in fact be driving (in part) the physiology and behavior of the mother, and that different fetal genotypes have differential long-term consequences on both maternal and offspring metabolic health. Using a statistical method originally developed for testing for linkage in the presence of association (Allison, 1997), we can rigorously test this hypothesis and make causal inferences, even in humans, by relying on the (conditional) randomization inherent in the process of meiosis.

The maternal–fetal relationship can and has been viewed as both one of intimate cooperation and that of intergenerational competition (Trivers, 1974; Haig, 1993; Godfray, 1995). Foundational to these views is allocation of nutrient resources to mother and fetus as well as the alteration of maternal behavior, physiology, and anatomy for the benefit or detriment of mother or fetus. From a genetic perspective, the 'conflict of interest' may occur because the mother's and the fetus' genomes are correlated, but not identical (due to the random process of meiosis and to the presence of paternal genes in the fetus; Page, 1939; Douglas et al., 1959; Haig and Westoby, 1989; Moore and Haig, 1991; Haig, 1993, 2004; Moore, 2012). As the fetus (via the placenta) gains access to the maternal blood supply, the fetus can potentially influence the behavior, physiology, and anatomy of the mother, to 'canalize'

(Van Vleck, 1978; Gottlieb, 1991; Hallgrímsson et al., 2002) its ontogeny in a way that is well-suited to its genome. For example, a number of placental hormones manipulate maternal energy metabolism and blood supply for fetal benefit. Human placental lactogen (hPL; a diabetogenic, growth hormone-related placental hormone) plays a role in increasing maternal insulin resistance, thereby driving more energy to the fetus. Nevertheless, because hPL production is independent of maternal regulation, mothers respond to this metabolic challenge by increasing insulin production to maintain glycemic control. Additionally, increments in blood pressure increase the flux of nutrients across the intrauterine environment, and gestational hypertension induced in the first trimester leads to reduced perinatal mortality (Symonds, 1980; Hollegaard et al., 2013) and higher birth weights (Goddard et al., 2007). Nevertheless, preeclampsia (i.e., hypertension with proteinuria) is a major cause of both maternal and fetal mortality and morbidity (Haig, 1993). These results (Trivers, 1974; Haig, 1993; Godfray, 1995) suggest that fetal genes may in some cases be influencing maternal behavior, physiology, and anatomy in ways that influence the long-term health of mother, fetus, and/or both. In fact, there is recent evidence that fetal sex (a function of fetal genotype) has significant effects on maternal pre and postnatal phenotype [e.g., increased breast circumference (Galbarczyk, 2011), increased accumulation of adipose tissue, and breast milk energy content (Hinde, 2007; Powe et al., 2010; Hinde et al., 2014)]. These results and others (Takimoto et al., 1996; Kanayama et al., 2002; Tamimi et al., 2003; Dekker and Robillard, 2005; Wangler et al., 2005; Petry et al., 2007) support emerging evidence that fetal genotype affects maternal physiology both during and after pregnancy. Thus, some fetal genotypes may induce long-term

effects on maternal and fetus' postnatal health, in part via manipulation of maternal phenotype during pregnancy. Petry et al. (2007) introduced this idea, but did not offer a statistical test that could test the hypothesis and in addition, separate correlation from causation.

Identifying specific aspects of fetal genotypes (beyond the differential effects of sex chromosomes) influence maternal phenotype may be challenging for all the reasons that studying complex genetic effects are challenging in general, but also because a statistical model for testing such hypothesized effects has not yet been offered. Except in certain situations of controlled experimental crosses, observational studies showing an association between fetal genotype at a particular genomic locus and maternal phenotype merely indicate association (correlation) and not necessarily of causation. Therefore, we offer a statistical model, adapted from one developed by one of the authors in the context of family based association (i.e., "TDT") testing, to rigorously test this hypothesis in humans (and other diploid populations) and make causal inferences by relying on the randomization inherent in the process of meiosis. Although the model is adapted from genetic study, the principle has its theoretic root in the causal inference literature. We describe the method here and illustrate it with data from the Framingham Heart Study (FHS).

## MATERIALS AND METHODS

### STATISTICAL MODEL

The essential proposition underlying the validity of the method we offer is that, under Mendelian theory, at any genetic locus, conditional on the parents' mating types at that locus, the offspring's genotype at that locus is a random variable for which the probability of each possible genotype is equal for all offspring. Therefore when we condition on parents' mating types at a locus of interest, then we have effectively a randomized experiment in which offspring are randomly assigned to genotypes at that locus. Stated, equivalently, mothers are randomized to carry offspring of different genotypes at that locus. Hence, if we test for relations of offspring genotype with mother's phenotype during pregnancy (e.g., preeclampsia, gestational diabetes, weight gain, or for that matter any time after conception), we can reasonably draw causal inferences about the effects of fetus genotype on mother's pregnancy because of the aforementioned randomization, *if we condition on parental mating types at the locus under study* (Allison and Neale, 2001; Tiwari et al., 2008). Note that this causal inference refers to the effect of the fetus genotype, not necessarily the fetus' genotype at the locus being used in the analysis. More specifically, the causal effect identified may be due to the fetal genotype at the locus under study or to the fetal genotype at another locus physically linked to and in disequilibrium with the locus under study. By this reasoning, the sex of offspring is also randomized in mammals, thus the finding that women carrying male fetuses had higher rates of gestational diabetes mellitus, indicates that having a fetus with a Y chromosome causes changes in mother's physiology.

The above principle has been formalized in the literature of causal inference (Rubin, 1974, 1977; Hernán and Robins, 2006; Greenland and Robins, 2009). Using Rubin's (1977) language and notation, if the experimental units are assigned into two treatment

groups solely on the basis of a covariate,  $X$ , and random factors then causal inference can be made and in addition, causal effects of treatments can be estimated without bias. The assignment to treatment group means that if two units have the same value of  $X$ , then they either must receive the same treatment or must be randomly assigned to treatments (not necessarily with the same probability). The critical point is that the probability that a study unit is assigned to one treatment rather than others is a function only of the values of  $X$  in the sample and purely random factors. Here,  $X$  can either be univariate or multivariate. The principle has been further explored for observational data explicitly in Hernán and Robins (2006). They argued that in ideal randomized experiments, causal inference can be made because the randomization ensures that the exposed and unexposed are exchangeable; whereas in observational studies, because the exposed and the unexposed are not generally exchangeable therefore causal inference cannot be generally made. They reviewed a condition that permits causal inference from observational studies, that is, the conditional randomization which guarantees conditional exchangeability. These works lay a theoretical foundation for our method. Specifically, here the study unit is parents-child trio; the parental mating types which are functions of the two parental genotypes, ( $g_F, g_M$ ), are the  $X$  in Rubin's (1977) paper and the covariate used to assign study units to treatment groups; the genotype of children,  $g_k$ , is the treatment although we have three groups here (i.e., three types of genotype); mother's phenotype is the dependent variable  $Y$  using Rubin's notation. Therefore conditional on parents' mating types at a locus, then the offspring are randomly assigned to genotypes at that locus. Equivalently, the study unit, the parents-child trio, is randomly assigned to the genotypes of offspring solely based on the parental mating types. Based on the aforementioned previous work, causal inference can be made on the effect of offspring's genotype (i.e., treatment effect in Rubin's words) on mother's phenotype (i.e., the dependent variable  $Y$ ). Note that the causal inference is to 'treatment assignment' which in this case is the offspring's genotype at the locus under study and/or the other loci physically linked to and correlated (in disequilibrium) with it (Allison, 1997). This is analogous to a randomized trial of a treatment (e.g., a diet involving eating more meat) in which strictly speaking the causal inference is about treatment assignment rather than meat *per se* because one cannot separate the effect of meat *per se* from other factors which may covary with assignment to more meat consumption such as the need to chew more. This is the reason why the class of genetic tests we are relying on are referred to as transmission *disequilibrium* tests.

Rubin (1977) discussed two general methods to estimate the causal effect. Because our primary interest is hypothesis testing, we choose to use a linear model as follows:

$$Y = \beta_0 + \mathbf{X} \boldsymbol{\beta}_1 + \beta_k g_k + \beta_m m + \varepsilon$$

where  $Y$  is mother's phenotype during pregnancy,  $\mathbf{X}$  is a matrix for covariates including environmental factors,  $g_k$  is genotype of fetus,  $m$  is the parental mating types at the same locus,  $\varepsilon$  is random error. In order to test the hypothesis, we only need to test  $H_0 : \beta_k = 0$ . Note that the genotypes should be treated as

categorical in the model (Rubin, 1977; Allison, 1997) so the test of null hypothesis is a two degrees of freedom test. Here,  $Y$  is assumed to be continuous. The same principle can be used when  $Y$  is other type of data (e.g., categorical, count, or survival), by use of a link function (i.e., in a generalized linear model framework; Nelder and Baker, 2004). Note also that the causal inference about the effects of fetal genotype on the mother's pregnancy is based on the randomization achieved by conditioning on parental mating types at the locus under study, resulting in the inclusion of mating types  $m$  in the model (Rubin, 1977). Given the above principle, from the point of view of genetic study, the statistical model is in spirit the same as the method proposed by one of the authors (Allison, 1997), for testing such hypotheses. This work can be seen as an application of a previously validated and analytically derived method (Model Q5) to a novel context (Allison, 1997; Ewens et al., 2008; Of course, it is also an application of the work of Rubin, 1977). Given that the theoretical work in the literature and the extensive simulation studies have already demonstrated both the validity of concept and the performance of the model (Allison, 1997) we will not perform the simulations here. The analysis of data below is for illustration of this novel application, rather than a test of the previously established method (Allison, 1997), nor an attempt to draw biological conclusions from the small sample dataset.

## DATA

For illustration, we applied the concept and the model to a dataset from the FHS. The study began in 1948 with 5,209 adult subjects (i.e., the first generation) from Framingham, Massachusetts, and is now on its third generation of participants. The Offspring Study (i.e., the second generation) was initiated in 1971. A sample of 5,124 men and women, consisting of the offspring of the Original Cohort and their spouses was recruited. The recruitment of the third generation participants who had at least one parent in the Offspring Study and would be at least 20 years old by the close of the first exam cycle was started in 2001. A recruitment target of 4,095 Gen III participants was achieved by July of 2005. In order to draw causal inference using the model as described above, we need information from both parents and offspring (although extensions to allow for missing data from one parent with substitution of sibling data are possible (Dudbridge, 2008)), and most importantly, measurements during mothers' pregnancy. Because there is little information about pregnancy for the original cohort, we had to use only the offspring generation (i.e., the second generation) and their offspring (the third generation). We focus on metabolic related measurements during pregnancy, specifically, gestational hypertension, diabetes, and weight gain which are all available in FHS. We chose to focus our examination on maternal metabolic

function because the large variability of maternal phenotypic in response to the metabolic sequelae of pregnancy may be influenced by fetal genotypes. Gestational hypertension and diabetes are binary. For weight gain, there is only information about gaining more than 30 pounds during pregnancy or not in FHS. We need information (including genotype information) from both parents and children. After applying all the criteria, there are 109 families left, with a total of 282 children, clearly a small sample, but suitable for illustrative purposes. In the data, there are four observations of gestational hypertension, no observation of gestational diabetes, and 49 observations of weight gain in excess of 30 pounds. In the following analysis, we only use the first child in each family to avoid the complexities introduced by potential correlations among siblings.

Although information on gestational hypertension, diabetes, and weight gain is available in FHS, in the final dataset we obtained, the number of observations of hypertension and diabetes are small. Therefore we only focus on weight gain, which is binary with value 1 if the mother reported to have >30 pounds weight gain during pregnancy and 0 otherwise, that is:

$$y = I(\text{had over 30 lbs of weight gain})$$

where  $I()$  is the indicator function.

FTO is a gene located in chromosome region 16q12.2. Studies have revealed association of single nucleotide polymorphisms (SNPs) in this gene with obesity. The genotype we tested are four SNPs (i.e., rs9930506, rs9939609, rs1121980, and rs8050136) in FTO gene which have been shown to be associated with obesity (Dina et al., 2007; Frayling et al., 2007; Price et al., 2008; Li et al., 2010; Liu et al., 2013). We performed a test of Hardy-Weinberg equilibrium and no significant departure was found.

## RESULTS

To analyze the data, we used logistic regression. We performed analyses including and excluding mother's age at children's birth in the model. For the analyses we conducted, none of the mother's age at children's birth was significant (at  $\alpha = 0.05$ ). Therefore they were not included in the final models. **Table 1** depicts the results from the logistic regression for SNP rs9930506. The results from additional analyses are summarized in **Table 2**.  $P$ -values for the genotypes are from two degrees of freedom test of the two dummy variables for each genotype. Neither of the results is significant at  $\alpha$  level of 0.05. This may indicate that the variants of these four SNPs carried by children do not have effect on mothers' weight gain during pregnancy. Of course, given the small sample size, there is a possibility that there is not enough power to detect

**Table 1 | Results from logistic regression for SNP rs9930506 from data with only the first child in each family.**

Independent variable	Intercept	Fetal genotype		Mating type				
		0	1	0	1	2	3	
Coefficient	-1.504	-0.049	0.239	0.177	-0.527	0.200	0.860	0.133
$P$ -value	0.054	0.928		0.974				

$P$ -values for the genotypes are from two degrees of freedom test.

such effect. Given that this is an illustration of the application of a previously validated method, and the small sample size, we did not intend to arrive at any biological conclusions from this analysis.

## DISCUSSION

The hypothesis that the fetal genotype affects maternal physiology and behavior represents a complementary view about not only some gestational diseases (e.g., pregnancy-induced hypertension and gestational diabetes) of mothers, but also long-term effects on metabolic health of both mother and offspring. In addition, the effects of this process may be cyclic over generations (Petry et al., 2007). This hypothesis does not contradict the common belief that association between fetal growth and diseases in pregnancy results from effects of the mother's genotype and/or environment on her physiology which subsequently affect the fetus. Rather, both of these causal paths can, in principle, be active.

Both animal studies and human studies have shown the effect of fetal genotype on mother's physiology during pregnancy (Galbarczyk, 2011; Hinde et al., 2014). This provides indirect evidence for our hypothesis. However, the direct test of this hypothesis is very difficult, especially in human studies. We illustrate that the process of meiosis provides a natural randomization which can be used in statistical analysis for causal inference and offer a simple statistical model for such analysis. This statistical model can be viewed either as stemming from causal inference literature or stemming from our previous work in genetic study. Of course, the theoretical work in causal inference provides foundation and our previous work in genetic study provides intuition and specific model (Model Q5 in Allison, 1997). Although Model Q5 proposed by Allison (1997) is conditional on three mating types excluding families with both homozygous parents, this can actually be released, leading to the more general model as proposed in this work. Allison's (1997) Q5 model excludes families with both parents homozygous because such families add no information to the estimation of the elements of  $\beta$ . However, under the homoscedasticity assumption, they do contribute to the precision of the estimation of residual variance and therefore can be included, and their inclusion should result in slightly greater power under the alternative hypothesis. From the point of view of causal inference (Rubin, 1977), in principle causal inference can be made when conditional on parental mating types, given that we do not consider parent-of-origin effects (that is we assume that alleles in offspring have the same effects regardless of which parent those alleles were inherited from), because the probability of offspring genotypes can be fully determined by

parental mating types. Therefore in our model we conditional on parental mating types. By using this analytic method, if a significant finding is obtained, one can justifiably claim that either the fetal genotype at the locus in question or a genotype physically linked to it has a causal effect on the maternal phenotype. It should be noted that many factors may contribute to the maternal phenotype. Our model only tests if fetal genotype is one of the factors, and if it is, it is a causal factor. To illustrate the principle, we analyzed a dataset from FHS using the model we proposed.

It should be noted that random mating is not assumed for our model. The model is based on the randomization of offspring via the random assignment to genotypes at the locus under study. The randomization holds *if we condition on parental mating types at that locus* under Mendelian theory (Allison and Neale, 2001; Tiwari et al., 2008). The inclusion of the mating types in the model is necessary for this "conditioning" to achieve the randomization, as pointed out by Rubin (1977). Because the fetal genotype is derived from parental genotypes, they are highly correlated. This may raise the concerns related to collinearity in numerical computation. However, collinearity is caused by high linear correlations between independent variables in the model. Although fetal genotype is correlated with parental genotypes, the linear correlations via Pearson correlation coefficients suggest that they are not prohibitively high. In our model, we treat the offspring's genotypes and parental mating types as categorical so dummy variables are created for each of them. Therefore the correlation between parental mating types and fetal genotypes is modest. In the data we analyzed above, the correlation coefficients are less than 0.474 between fetal genotype and paternal mating types for the four SNPs, respectively. These low to moderate correlations may not induce undue issues of collinearity. Nevertheless, we recommend investigators assessing collinearity in their datasets and interpreting results accordingly.

Our model is based on a linear model (more general, generalized linear model) and assumes samples with unrelated study units. The model can be extended to handle data with related study units, such as biological siblings, by the use of a linear mixed model (or more general, generalized linear mixed model). The correlation between the related individuals can be dealt with via a random effect by using a matrix of kinship coefficients (a kinship coefficient is a measure of degree of genetic correlation between two individuals; Yu et al., 2006; Kang et al., 2010; Liu et al., 2011). In this work, we only use a linear model to test the fetal drive effects. Causation can be inferred because of the conditional randomization in the process of meiosis. More complicated models, such as structural equations models, may capture more about the complexity of the data and may work more efficiently. However, we focus on causal inference; therefore care must be taken in order to test causation with more complicated models. This is beyond the scope of this work and is a good topic for future research.

There are no significant findings in our data analysis. This may be because that the four SNPs we tested did not have deleterious effects on maternal metabolic function, or our analysis lacked the requisite power to detect an effect. Statistical power is influenced

**Table 2 | P-values of children's genotypes from the analyses with only the first child in each family.**

Statistical model	SNPs			
	rs9930506	rs9939609	rs1121980	rs8050136
Logistic regression	0.928	0.946	0.944	0.965

*P-values for the genotypes are from two degrees of freedom test.*

by many factors including effect size and sample size. Metabolic function is a complex phenotype, and we did not expect a large effect size for any of its risk factors. We only tested four SNPs in FTO gene due to its relationship with obesity, in a small sample, and think that a genome-wide scan in a much larger sample is warranted for more reliable inferences.

Although our statistical method makes the direct test of our hypothesis possible in human studies, it should be noted that conducting such studies necessitate a great deal of time and effort because both parents and offspring need to be included in the study and relevant physiological measurements must be made in the mothers during pregnancy. The effort is worthwhile because if the hypothesis is validated, our understanding of some gestational metabolic conditions may be shifted to a new level. New strategies may be developed to prevent and reduce morbidity and possibly mortality in both the mother and offspring and their future descendants.

## ACKNOWLEDGMENTS

Supported by NIH grants P30DK056336, T32DK062710, R25DK099080, R01DK052431, R01GM081488, R01HL092173, S10RR026723, and P60AR064172.

## REFERENCES

- Allison, D. B. (1997). Transmission-disequilibrium tests for quantitative traits. *Am. J. Hum. Genet.* 60, 676–690.
- Allison, D. B., and Neale, M. C. (2001). Joint tests of linkage and association for quantitative traits. *Theor. Popul. Biol.* 60, 239–251. doi: 10.1006/tpbi.2001.1544
- Catalano, P. M., Thomas, A. J., Huston, L. P., and Fung, C. M. (1998). Effect of maternal metabolism on fetal growth and body composition. *Diabetes Care* 21(Suppl. 2), B85–B90.
- Dekker, G. A., and Robillard, P. Y. (2005). Preeclampsia: a couple's disease with maternal and fetal manifestations. *Curr. Pharm. Des.* 11, 699–710. doi: 10.2174/1381612053381828
- Dina, C., Meyre, D., Gallina, S., Durand, E., Korner, A., Jacobson, P., et al. (2007). Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat. Genet.* 39, 724–726. doi: 10.1038/ng2048
- Douglas, G. W., Thomas, L., Carr, M., Cullen, N. M., and Morris, R. (1959). Trophoblast in the circulating blood during pregnancy. *Am. J. Obstet. Gynecol.* 78, 960–973.
- Dudbridge, F. (2008). Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum. Hered.* 66, 87–98. doi: 10.1159/000119108
- Ewens, W. J., Li, M., and Spielman, R. S. (2008). A review of family-based tests for linkage disequilibrium between a quantitative trait and a genetic marker. *PLoS Genet.* 4:e1000180. doi: 10.1371/journal.pgen.1000180
- Frayling, T. M., Timpson, N. J., Weedon, M. N., Zeggini, E., Freathy, R. M., Lindgren, C. M., et al. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 316, 889–894. doi: 10.1126/science.1141634
- Galbarczyk, A. (2011). Unexpected changes in maternal breast size during pregnancy in relation to infant sex: an evolutionary interpretation. *Am. J. Hum. Biol.* 23, 560–562. doi: 10.1002/ajhb.21177
- Godfray, H. C. (1995). Evolutionary theory of parent-offspring conflict. *Nature* 376, 133–138. doi: 10.1038/376133a0
- Goddard, K. A., Tromp, G., Romero, R., Olson, J. M., Lu, Q., Xu, Z., et al. (2007). Candidate-gene association study of mothers with pre-eclampsia, and their infants, analyzing 775 SNPs in 190 genes. *Hum. Hered.* 63, 1–16. doi: 10.1159/000097926
- Gottlieb, G. (1991). Experiential canalization of behavioral development: theory. *Dev. Psychol.* 27, 4–13. doi: 10.1037/0012-1649.27.1.4
- Greenland, S., and Robins, J. M. (2009). Identifiability, exchangeability and confounding revisited. *Epidemiol. Perspect. Innov.* 6:4. doi: 10.1186/1742-5573-6-4
- Haig, D. (1993). Genetic conflicts in human pregnancy. *Q. Rev. Biol.* 68, 495–532. doi: 10.1086/418300
- Haig, D. (2004). Genomic imprinting and kinship: how good is the evidence? *Annu. Rev. Genet.* 38, 553–585. doi: 10.1146/annurev.genet.37.110801.142741
- Haig, D., and Westoby, M. (1989). Parent-specific gene expression and the triploid endosperm. *Am. Nat.* 134, 147–155.
- Hallgrímsson, B., Willmore, K., and Hall, B. K. (2002). Canalization, developmental stability, and morphological integration in primate limbs. *Am. J. Phys. Anthropol.* (Suppl. 35), 131–158. doi: 10.1002/ajpa.10182
- Hernán, M. A., and Robins, J. M. (2006). Estimating causal effects from epidemiological data. *J. Epidemiol. Community Health* 60, 578–586. doi: 10.1136/jech.2004.029496
- Hinde, K. (2007). First-time macaque mothers bias milk composition in favor of sons. *Curr. Biol.* 17, R958–R959. doi: 10.1016/j.cub.2007.09.029
- Hinde, K., Carpenter, A. J., Clay, J. C., and Bradford, B. J. (2014). Holsteins favor heifers, not bulls: biased milk production programmed during pregnancy as a function of fetal sex. *PLoS ONE* 9:e86169. doi: 10.1371/journal.pone.0086169
- Hollegaard, B., Byars, S. G., Lykke, J., and Boomsma, J. J. (2013). Parent-offspring conflict and the persistence of pregnancy-induced hypertension in modern humans. *PLoS ONE* 8:e56821. doi: 10.1371/journal.pone.0056821
- Kanayama, N., Takahashi, K., Matsuura, T., Sugimura, M., Kobayashi, T., Moniwa, N., et al. (2002). Deficiency in p57Kip2 expression induces preeclampsia-like symptoms in mice. *Mol. Hum. Reprod.* 8, 1129–1135. doi: 10.1093/molehr/8.12.1129
- Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S. Y., Freimer, N. B., et al. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 42, 348–354. doi: 10.1038/ng.548
- Li, M., Zhang, T., Liu, Y., and Xu, P. R. (2010). [The research of association between gene rs9930506 polymorphism and Hazakh children with overweight or obesity in Xinjiang]. *Zhonghua Yu Fang Yi Xue Za Zhi* 44, 1106–1110.
- Liu, C., Mou, S., and Cai, Y. (2013). FTO gene variant and risk of overweight and obesity among children and adolescents: a systematic review and meta-analysis. *PLoS ONE* 8:e82133. doi: 10.1371/journal.pone.0082133
- Liu, N., Zhao, H., Patki, A., Limdi, N. A., and Allison, D. B. (2011). Controlling population structure in human genetic association studies with samples of unrelated individuals. *Stat. Interface* 4, 317–326. doi: 10.4310/SII.2011.v4.n3.a6
- Moore, T. (2012). Review: parent-offspring conflict and the control of placental function. *Placenta* 33(Suppl.), S33–S36. doi: 10.1016/j.placenta.2011.11.016
- Moore, T., and Haig, D. (1991). Genomic imprinting in mammalian development: a parental tug-of-war. *Trends Genet.* 7, 45–49. doi: 10.1016/0168-9525(91)90230-n
- Nelder, J. A., and Baker, R. J. (2004). “Generalized linear models,” in *Encyclopedia of Statistical Sciences* (John Wiley & Sons Inc.). doi: 10.1002/0471667196.ess0866.pub2
- Page, E. W. (1939). The relation between hydatid moles, relative ischemia of the gravid uterus, and the placental origin of eclampsia. *Am. J. Obstet. Gynecol.* 37, 291–293.
- Petry, C. J., Ong, K. K., and Dunger, D. B. (2007). Does the fetal genotype affect maternal physiology during pregnancy? *Trends Mol. Med.* 13, 414–421. doi: 10.1016/j.molmed.2007.07.007
- Plagemann, A. (2005). Perinatal programming and functional teratogenesis: impact on body weight regulation and obesity. *Physiol. Behav.* 86, 661–668. doi: 10.1016/j.physbeh.2005.08.065
- Powe, C. E., Knott, C. D., and Conklin-Brittain, N. (2010). Infant sex predicts breast milk energy content. *Am. J. Hum. Biol.* 22, 50–54. doi: 10.1002/ajhb.20941
- Price, R. A., Li, W. D., and Zhao, H. (2008). FTO gene SNPs associated with extreme obesity in cases, controls and extremely discordant sister pairs. *BMC Med. Genet.* 9:4. doi: 10.1186/1471-2350-9-4
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* 66, 688–701. doi: 10.1037/h0037350
- Rubin, D. B. (1977). Assignment to treatment group on the basis of a covariate. *J. Educ. Stat.* 2, 1–26. doi: 10.2307/1164933
- Symonds, E. M. (1980). Aetiology of pre-eclampsia: a review. *J. R. Soc. Med.* 73, 871–875.
- Takimoto, E., Ishida, J., Sugiyama, F., Horiguchi, H., Murakami, K., and Fukamizu, A. (1996). Hypertension induced in pregnant mice by placental renin and maternal angiotensinogen. *Science* 274, 995–998. doi: 10.1126/science.274.5289.995
- Tamimi, R. M., Lagiou, P., Mucci, L. A., Hsieh, C. C., Adami, H. O., and Trichopoulos, D. (2003). Average energy intake among pregnant women carrying

- a boy compared with a girl. *BMJ* 326, 1245–1246. doi: 10.1136/bmj.326.740.1.1245
- Tiwari, H. K., Barnholtz-Sloan, J., Wineinger, N., Padilla, M. A., Vaughan, L. K., and Allison, D. B. (2008). Review and evaluation of methods correcting for population stratification with a focus on underlying statistical principles. *Hum. Hered.* 66, 67–86. doi: 10.1159/000119107
- Trivers, R. L. (1974). Parent-offspring conflict. *Am. Zool.* 14, 249–264. doi: 10.1093/icb/14.1.249
- Van Vleck, L. D. (1978). A genetic model involving fetal effects on traits of the dam. *Biometrics* 34, 123–127. doi: 10.2307/2529597
- Wangler, M. F., Chang, A. S., Moley, K. H., Feinberg, A. P., and Debaun, M. R. (2005). Factors associated with preterm delivery in mothers of children with Beckwith-Wiedemann syndrome: a case cohort study from the BWS registry. *Am. J. Med. Genet. A* 134A, 187–191. doi: 10.1002/ajmg.a.30595
- Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208. doi: 10.1038/ng1702
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 13 May 2014; accepted: 17 December 2014; published online: 13 January 2015.
- Citation: Liu N, Archer E, Srinivasasainagendra V and Allison DB (2015) A statistical framework for testing the causal effects of fetal drive. *Front. Genet.* 5:464. doi: 10.3389/fgene.2014.00464
- This article was submitted to *Applied Genetic Epidemiology*, a section of the journal *Frontiers in Genetics*.
- Copyright © 2015 Liu, Archer, Srinivasasainagendra and Allison. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.