Investigating the associations between TP53 variation and miscarriage: Follow up from the ALSPAC and Pelotas 1982 and 1993 birth cohorts in the GOYA Sample.

Fernando Pires Hartwig  
Prof Ellen Aagaard Nøhr, GOYA  
Prof Ana Maria Baptista Menezes, Pelotas 1993 birth cohort  
Prof’s Nicholas Timpson & George Davey Smith, ALSPAC

Aim
To investigate the associations between TP53 genetic variation (R72P SNP and haplotypes) and miscarriage in the GOYA (Genetics of Overweight Young Adults) cohort sample within the DNBC.

Introduction
TP53 is a gene that has been primarily studied in the context of cancer, in which its importance is hardly arguable due to the roles of p53 in regulating several pathways involved with the carcinogenic process. More recently, however, p53 has been found to have multiple roles, thus extending its importance to other diseases and health-related outcomes. Among these, there are pregnancy-related outcomes (especially miscarriage-related), which started to be studied in this regard after the publication of first functional link between p53 and pregnancy [1]. This finding was latter supported by a landmark study on the topic that showed that p53 has a role in the early embryonic development by regulating the expression of the LIF protein, thus having implications for embryonic implantation, pregnancy rates and litter size in mice [2].

These studies provided the incentive for genetic epidemiological investigations, which are relatively recent in the literature (the first was published in 2005 [3]). Most of the studies have focused on the Arginine72Proline (rs1042522) TP53 SNP, which is largely studied in the context of cancer [4]. Associations have been found for many pregnancy-related outcomes, including recurrent pregnancy loss and ICSI/IVF problems [3, 5, 6], although there is also evidence pointing to a null effect of this variant [7-9]. Although still inconclusive at the epidemiological level, the functional link between p53 and reproduction has strong biological plausibility and experimental support [10, 11], including the reported association between rs1042522 and LIF expression [5].

Furthermore, this SNP was recently associated with twinning rate [12], which is in accordance to earlier experimental findings regarding the roles of p53 in regulating LIF [2]. A plausible explanation for the inconclusiveness of the relationship between p53 and human reproduction is that rs1042522 is not the only component of the genetic variation in the TP53 locus that is important for pregnancy traits. In line with this, there is evidence that another TP53 SNP is associated with recurrent pregnancy loss [13]. Moreover, linkage disequilibrium analyses using Avon Longitudinal Study of Parents and Children (ALSPAC) data on 10 TP53 SNPs evidenced that rs1042522 does not capture overall TP53 genetic variability.

To investigate this in population-based samples, we have so far performed single-SNP (rs1042522, coded in five different models of genetic effect) in the ALSPAC and Pelotas (Brazil) 1982 and 1993 birth cohorts. Haplotype-trait (using 10 SNPs) associations (additive and dominant effects, since recessive effect was not feasible) were already performed in ALSPAC and are currently ongoing in Pelotas 1982 birth cohort. We hope to perform similar analyses in GOYA and then meta-analyze the results.

Study protocol
Data requested: We would like to include the 1948 control women and 1960 obese women in the GOYA cohort sample within the DNBC. For these women, we ask to use individual
genetic data (10 SNPs in the TP53 gene) and information about BMI and case/control status of the mother. Maternal outcome is number of miscarriages. In addition, information on a number of potential confounding factors is also required: maternal age at birth, maternal education, parity and maternal smoking during last pregnancy (i.e., referent to the children that is a DNBC member).

2) Genotyping:
10 SNPs in the TP53 gene: rs8073498, rs12951053, rs1625895, rs2909430, rs9895829, rs1042522, rs8079544, rs12602273, rs2078486 and rs8064946.

3) Statistical analysis:

3.1) The variables collected as potential confounders will be submitted to a statistically-oriented process of selection. The process is performed twice, having, as the dependent variable, the outcome (number of miscarriages, in a Poisson regression model) and each exposure (rs1042522 – coded as the codominant effect – and TP53 diplotypes; in a multinomial regression model), and all potential confounders (including case-control status for obesity) as dependent variables. If at least one of the dependent variables has P > 0.2, the least associated one is removed and the model re-fitted. This process will be repeated until all remaining independent variables have P > 0.2. Variables that remained in the final models of both the outcome and exposure for a given analysis were considered confounders and adjusted for subsequently.

3.2) Using number of miscarriages and having had at least one miscarriage as outcomes, the associations involving rs1042522, under five models of genetic effects (i.e., codominant, overdominant, additive, dominant and recessive), will be tested using the appropriate regression framework [Poisson and Negative binomial regression (estimating relative risks); and Poisson regression (estimating prevalence rates)].

3.3) 10 SNPs across the TP53 locus will be used for the haplotype analyses. Haplotype-trait associations (additive and dominant effects, since recessive is not feasible) will be performed based on a two-step iteration process: the posterior probabilities of pairs of haplotypes per subject are used as weights to update the regression coefficients, and the regression coefficients are used to update the haplotype posterior probabilities. It is, then, a fully likelihood-based approach that involves simultaneous haplotype frequency estimation and estimation of haplotype-outcome association. For both number of miscarriages and having had at least one miscarriage, Poisson regression will be used (estimating relative risks and prevalence rates, respectively).

3.4) Repeat 3.1-3.3, stratifying into reference (controls) and obese women (cases).

4) Logistics:

Phenotype data will be provided by Ellen Aagaard Nohr and sent to our GOYA collaborators at University of Bristol. Here, an anonymized individualized data set containing the 10 TP53 SNPs will be generated and sent to Fernando Hartwig at the Epidemiology Research Centre, Federal University of Pelotas, Brazil.
References