Epigenetics and Maternal Smoking

Background

The prevalence of maternal cigarette-smoking during pregnancy in the U.S. remains, by some estimates, as high as 14%. This figure has been alarmingly persistent, despite strident warnings by health authorities about the effect of smoking on the developing fetus.

Indeed, it is thought more generally that many complex disorders have origins in utero, resulting from adaptations to the gestational environment and/or from exposure to toxicants, such as those that result from maternal smoking during pregnancy. For example, exposures may affect the epigenetic regulation of imprinted genes that are critical to normal growth and development, thereby increasing the risk of adverse health outcomes. Imprinted genes are those whose regulation is either maternally or paternally controlled. Normally, one of the two alleles of an imprinted gene is silenced via methylation[1] at differentially methylated regions (DMRs) located on that allele. If the gene is maternally expressed, the paternal allele is silenced; if it is paternally expressed, the maternal allele is silenced. Hence when such a DMR is assayed, a normal individual's methylation fraction (measured using the DNA of many cells from the subject) should be near 50%.

[1] Ask a biologist, or see, e.g. `http://en.wikipedia.org/wiki/DNA_methylation`. If you want to read about a particularly well-studied example of imprinting, try reading up on Prader–Willi and Angelman syndromes.

A well-characterized imprinted domain on human chromosome band 11p15.5 contains the genes for paternally expressed Insulin-like Growth Factor II (IGF2) and maternally expressed H19. Deregulation of IGF2 expression has been linked to growth disorders, obesity, and cancer. Imprinted expression and transcription of IGF2 are regulated in large part through the patterns of differential methylation of at least two regulatory DMRs, one of which is located near the H19 promoter (H19 DMR) and the other upstream of the three IGF2 promoters that are subject to imprinting (IGF2 DMR). Both DMRs have been shown to exhibit altered methylation in cigarette smoking-related malignancies.

The goal of this analysis is to examine the influence of a developing fetus' in-utero exposure to maternal cigarette smoking byproducts on percent methylation at the IGF2 DMR using umbilical-cord blood samples. The data are from a multi-ethnic birth cohort study designed to measure the effects of early exposures on both epigenetic profiles and on phenotype. The IGF2R

DMR methylation is measured as the percent of the subject's (offspring's) alleles that are methylated in a sample of their DNA and, hence, takes values between 0 and 100. Methylation at this locus was assayed twice for 294 of the 314 subjects and only once for the remaining 20. The 608 DNA samples were arranged on 22 different plates of 96 wells each, and each plate was processed separately. Each plate has eight rows, designated A through H; and twelve columns, designated 1 through 12. Only a subset of wells were used on each plate.

Data

The data are in the (tab-delimited) file "epigen.txt." You can read in the data with the command

```
epigen = read.table("epigen.txt", header=TRUE, sep="\t")
```

There are 314 rows, where each row corresponds to one subject (offspring and mother), with columns corresponding to:

age: Maternal age coded lt30 (younger than 30 at delivery), 30to39 (30 to 39 years old at delivery), and ge40 (more than 40 years old at delivery).

BMI: Maternal body mass index measured as weight before pregnancy (in kg) divided by height (in meters) squared; coded 0 (less than 30) or 1 (greater than or equal to 30).

smoke: Maternal cigarette smoking during pregnancy. Coded 0 if the mother did not smoke; and 1 if the mother smoked during early pregnancy and stopped, or if the mother smoked throughout pregnancy.

gestage: Gestational age of the infant. Coded 1 if less than 37 weeks and 0 if greater than or equal to 37 weeks.

sex: Infant's sex (1=male, 0=female).

edu: Mother's education level (coded ltHS for less than high school, ltCollege for high school/GED, and geCollege for at least some college)

race: Mother's race/ethnicity (AA=African American, EA=Caucasian or Other).

methyl1: the first replicate measurement of the subject's (fetus') methylation level.

methyl2: the second replicate measurement of the subject's methylation level. 20 subjects do not have a second measurement.

plate1: (and plate2) plates on which the subjects first and second replicate measurements were made, respectively. plate2 is missing for 20 subjects, who have only a single measurement.

row1: (and row2) row in which the subjects first and second replicate measurements were placed on the plates whose IDs appear in plate1 and in plate2, respectively. row2 is missing for 20 subjects.

column1: (and column2) column in which the subjects first and second replicate measure- ments were placed on the plates whose IDs appear in plate1 and in plate2, respectively. column2 is missing for 20 subjects.

well1: (and well2) well in which the subjects first and second replicate measurements were placed on the plates whose IDs appear in plate1 and in plate2, respectively. well2 is missing for 20 subjects.

Objectives

A fetus' methylation level may depend on sex and race and may be associated with maternal smoking. It may also be associated with unmeasured exposures related to maternal age, maternal BMI, maternal education level, and/or gestational age. Explore and analyze the data, characterizing the relationships between percent methylation in the offspring and maternal smoking status, adjusting for any other relevant covariates. It is also of interest to determine if the relationship between smoking and methylation is the same across race and sex.