Heredity (2006) 97, 256–257

www.nature.com/hdy

NEWS AND COMMENTARY

Population genetics

Separating nurture from nature in estimating heritability

S Xu

Heredity (2006) **97,** 256–257. doi:10.1038/sj.hdy.6800859; published online 28 June 2006

Are human features more a product of Nature (genetics and heredity) or Nurture (culture and environment)? Recent work by Visscher *et al* (2006) is shedding light on the answers to an age-old question in a novel way. The question is difficult because many traits are quantitative: meaning that they are controlled by multiple genes, and may also be influenced by environment. It is difficult to dissect natural from nurturing influences because the members of human families, who share genetic affinities, also tend to share a common nurturing culture and physical environment. Furthermore, although animal and plant experiments can be designed to remove such common environmental effects (CEEs), designed experiments cannot be readily implemented for humans for ethical reasons.

However, Visscher *et al* (2006) demonstrate that nature and nurture can be separated without resorting to any special experimental design, through utilizing molecular markers. This study is the first attempt of its kind.

The innovation is to utilize marker data to measure the proportions of identical-by-descent (IBD) genes actually shared by siblings. Remember that although siblings share, on average, half their genes, some sibling pairs will share more and some will share fewer. Although some siblings will have shared similar environments, and some will have experienced dissimilar environments, the strength of these environmental effects will not be related to the IBD values. The association between the IBD value and the similarity of the siblings' phenotype can therefore be used to calculate a covariance and provide an unbiased heritability estimate.

Using a dense marker map and 3375 sib pairs of humans, they calculated the actual IBD values for all sib pairs. The mean and standard deviation of the actual IBD were 0.498 and 0.036, respectively, close to the theoretical values of 0.50 and 0.0384 (Kong *et al*, 2004). Incorporating the actual IBD into their variance component model, the authors estimated heritability of 0.80 for height (a typical quantitative trait). This estimate was free of bias and consistent with results from independent studies using other sources of information. Heritability runs from zero to one, so a value of 0.8 is considered very high. Human height, therefore, is highly heritable. Assume that, on average, the heights of men and women are 1.75 and 1.60 m, respectively. A married couple 1.80 and 1.65 m tall will give a child with an expected height deviating from the population mean by

$$0.8\times\frac{(1.80-1.75)+(1.65-1.60)}{2}=0.8\times0.05=0.04\text{m}.$$

This equation says that the average deviation of the parental heights from the population mean is 0.05 m, but only 80% (heritability) of the deviation can be transmitted to their child. The predicted height of the child would be $1.75+0.04=1.79$ m for a son and $1.60+0.04=1.64$ m for a daughter. In addition to this, heritability has many other uses (Falconer and Mackay, 1996).

The actual IBD is a variable rather than a constant, which can be demonstrated in the extreme case where the trait is controlled by a single locus. Let A_1A_2 and A_3A_4 be the genotypes for the parents at the locus of interest. If the genotypes of two sibs are A_1A_3 and A_2A_3, the IBD shared by the sibs is 1/2. However, there is a chance that the sibs may have genotypes of A_1A_3 and A_1A_3. In this case, the IBD is unity and the two sibs act like identical twins for this single gene trait. If the genotypes are A_1A_3 and A_2A_4, instead, the IBD is zero and the sibs behave like strangers. The variance of the IBD at a single locus is 1/8, the maximum possible variance that an IBD can have. The variance of the average IBD for k independent loci is $1/(8k)$. As k increases, the variance will decrease and approach zero as k goes to infinity.

However, the size of our genome is limited (35.8 Morgan (M)) and it only contains an equivalent of 85 independent loci (Visscher *et al*, 2006). This leads to a variance of the genome-wide IBD $1/(8\times85)=0.001475$. This variance is the key in the method for separating genetic and environmental resemblances. Because this variance is so tiny, the power is small – one needs 2500 sib pairs to estimate heritability of 0.8 with 90% power (Visscher *et al*, 2006). Therefore, the method is not appropriate for organisms with a large genome and performs better for organisms with a smaller genome.

An implicit assumption for the genome-wide IBD method is that all loci contribute equally to the polygenic variance. Consider the following quantitative trait locus (QTL) model,

$$\begin{aligned}\text{cov(sibs)}&=\sigma_c^2+\sum_{i=1}^{k}\pi_i\sigma_i^2\\&=\sigma_c^2+\frac{1}{k}\sum_{i=1}^{k}\pi_i(k\sigma_i^2)\end{aligned}$$

where σ_c^2 is the variance of CEE, σ_i^2 is the locus specific variance and π_i is the IBD of locus i. The genome-wide IBD model is

$$\text{cov(sibs)}=\sigma_c^2+\bar{\pi}(k\sigma^2)=\sigma_c^2+\bar{\pi}\sigma_A^2$$

where $\bar{\pi}=\frac{1}{k}\sum_{i=1}^{k}\pi_i$ is the average IBD of the genome and $\sigma_A^2=\sum_{i=1}^{k}\sigma_i^2$ is the polygenic variance. These two models are identical only if $\sigma_i^2=\sigma^2=\sigma_A^2/k$, for $i=1,\ldots,k$. If the assumption is violated, QTL mapping that identifies locus-specific variances is more efficient. However, QTL mapping is more complicated than polygenic analysis because each and every putative locus must be evaluated.

A more powerful and also more realistic method is the chromosome-wide IBD analysis (Visscher, 1996), where the polygenic variance simply takes the sum of all chromosome-specific variances. The variance of IBD for the largest human chromosome is 0.0199 (Visscher *et al*, 2006), which is 13.5 times the variance of the genome-wide IBD. One can imagine that such a chromosome-wide IBD analysis will increase the power significantly.

Handling a model with 22 variance components (22 chromosomes) is trivial. A more enhanced method is the multipoint method of Goldgar (1990). Here, the entire genome is divided into many regions and the IBD of each region is calculated using markers within that region. Under the region-wide IBD approach, Goldgar (1990) estimated the variance of each region and scanned all regions across the genome. This method can be extended to a model, including variances of all regions. The human

genome is about 35.8 M, which can be divided into 358 regions, assuming 10 cM per region. With the current statistical technology and computing power, a model with this many variances can still be handled with ease (Xu, 2003).

Joint estimation of variances for all regions may then not be called QTL mapping because no statistical tests are involved and all regions, regardless of the significance, contribute to the calculation of the polygenic variance. Visscher *et al* (2006) show us that, molecular markers, previously thought of being only useful for QTL mapping, can also help us study the inheritance of classically defined polygenic traits.

S Xu is at the Department of Botany and Plant Sciences, University of California, Riverside, CA 92521, USA.

e-mail: xu@genetics.ucr.edu

Falconer DS, Mackay TFC (1996). *Introduction to Quantitative Genetics*. Longman: London.
Goldgar DE (1990). *Am J Hum Genet* **47**: 957–967.
Kong X, Murphy K, Raj T, He C, White PS, Matise TC (2004). *Am J Hum Genet* **75**: 1143–1148.
Visscher PM (1996). *J Hered* **87**: 136–138.
Visscher PM *et al* (2006). *PLoS Genet* **2**: 0316–0325.
Xu S (2003). *Genetics* **163**: 789–801.

Editor's suggested reading

Bogdan M, Doerge RW (2005). Biased estimators of quantitative trait locus heritability and location in interval mapping. *Heredity* **95**: 476–484.
Luo L, Zhang Y-M, Xu S (2005). Focus article: a quantitative genetics model for viability selection. *Heredity* **94**: 347–355.
Xu C, Zhang Y-M, Xu S (2005). An EM algorithm for mapping quantitative resistance loci. *Heredity* **94**: 119–128.