

SHORT REPORT

Measuring environmental factors can enhance the search for disease causing genes?

T Dwyer, A-L Ponsonby, J Stankovich, L Blizzard, S Easteal

J Epidemiol Community Health 2004;58:613–615. doi: 10.1136/jech.2003.012971

The value of the concurrent measurement of environmental factors in studies aimed at the discovery of disease causing genes has been questioned on the grounds that such an approach fails to increase study power. This report discusses the issue and shows with examples from the recent literature that the examination of a gene disease association within an environmental subgroup can provide enhanced opportunities for detecting gene effects. The concurrent collection of environmental as well as genetic factors in studies of disease aetiology may enhance study informativeness and validity in several ways, including an increase in the power of the study to detect gene disease associations.

The draft of the human genome sequence greatly enhances the prospect of identifying genes that contribute to complex disease.¹ However, for most common diseases, individual genetic contributions are unlikely to be large. For example, a review of cancer epidemiology noted that only a few genetic polymorphisms seem to change risk substantially, with systematic meta-analyses revealing that many genes involved in the metabolism of mutagens have no or small effects, with summary odds ratios for polymorphisms less than two.² In addition, gene effects are embedded in complex causal systems that include environmental, behavioural, and developmental factors. Identifying individual gene effects has been difficult, and most reports have not held up under scrutiny.³ One response to this lack of success has been the adoption of population based studies that incorporate environmental as well as genetic measures—the argument for the approach resting on the availability of concurrent environmental measures with this design. The value of such studies has been questioned,⁴ an important element of the argument being that power to identify environmental or genetic effects is not increased in this way. An important assertion in support of this case is:—“Even if there are subgroups of genetically susceptible individuals, and the size of effect associated with an environmental exposure varies with genotype, the direction of this effect is unlikely to vary with genotype. This lack of variation limits the gain in statistical power”.⁴ We have examined the recent literature and show that marked variation in effect between strata has been observed to occur more frequently than might have been anticipated, and that this scenario, at least in relation to the direction of genetic associations within environmental strata, has already been encountered in human studies with an observable gain in power resulting from stratification.

RESULTS

Table 1 shows results from a recent study investigating relations between low birth weight, maternal smoking, and the *CYP1A1* gene.⁵ Previous to this study, maternal smoking

had been well established as a risk factor for low birth weight.⁶ The *CYP1A1* gene was considered a candidate for influencing birth weight among smokers because it controls a metabolic enzyme for chemicals in cigarette smoke.⁷ A priori, we would expect any gene polymorphism to exert its strongest influence in mothers who smoke.

This proved to be true as shown in table 1.

While there is no significant association between birth weight and genotype in an analysis unstratified by smoking status ($p = 0.972$), there is a significant genotype-birth weight association among babies of smoking mothers ($p = 0.016$). Stratification for the environmental measure has increased power because the magnitude of the association is much stronger among mothers who smoke, thus revealing an association.

Another study shows a similar pattern.⁸ The disease is Parkinson's disease, with a putative environmental risk factor—exposure to pesticides.⁹ The authors genotyped six polymorphisms in four genes, which code for glutathione transferases (GSTs). GST enzymes are part of the widespread glutathione system for xenobiotic detoxification but it is not yet known if inter-individual differences in handling toxicants are attributable to polymorphisms of the genes coding the enzymes themselves or of the genes coding for the receptors or transcription factors that regulate enzyme expression.¹⁰

Among the pesticide exposed group, there was a significant association with one of the six polymorphisms ($p = 0.004$) but no significant association overall ($p = 0.370$) (table 2). The magnitude of the association in the exposed group is more than fourfold that of the association overall (table 2). A third example is provided by a recent study investigating relations between antisocial behaviour, childhood maltreatment, and a common polymorphism in a neurotransmitter metabolising enzyme monoamine oxidase A (MAOA) that influences expression levels of the enzyme (table 3).¹¹ The environmental factor, childhood maltreatment, is a well known risk factor for antisocial behaviour.¹² Increased aggression has been observed in MAOA knockout mice,¹³ and human antisocial behaviour is associated with a rare mutation that prevents production of any enzyme.¹⁴ However, prior evidence concerning the relation between the common regulatory polymorphism and antisocial behaviours was inconclusive. Using linear regression analysis, the study found no significant overall effect of the polymorphism on antisocial behaviour ($p = 0.89$), the effect of the polymorphism was significantly modified by childhood maltreatment ($p = 0.01$).¹¹ Again, the gene-disease association would not have been revealed without examining the environmental exposure. Among people who were mistreated as children, the low MAOA activity variant was associated with a

Abbreviations: GST, glutathione transferase; MAOA, monoamine oxidase A

Table 1 An association between maternal CYP1A1 genotype and low birth weight is only observed in the subgroup of babies born to mothers who smoke during pregnancy⁵

Maternal smoking status during pregnancy	Maternal CYP1A1 genotype	Number of babies <2500 g	Number of babies ≥2500 g	Odds ratio (p value)*
All mothers	AA	91	334	1.01 (p=0.972)
	Aa or aa	68	248	
Smoking mothers only	AA	18	57	2.58 (p=0.016)
	Aa or aa	22	27	

*p Values calculated using Woolf's formula for the standard error of the log odds ratio.²²

Table 2 An association between a polymorphism in the GSTP1 gene and Parkinson's disease is strongest in the subgroup of subjects exposed to pesticides⁸

Exposure status	GSTP1 genotype at codon 105	Parkinson's disease cases	Controls	Odds ratio (p value)*
All subjects	Ile/Ile	33	39	1.31 (p=0.0370)
	Ile/Val, Val/Val	62	56	
Subjects exposed to pesticides only	Ile/Ile	7	14	5.33 (p=0.004)
	Ile/Val, Val/Val	32	12	

*p Values calculated using Woolf's formula for the standard error of the log odds ratio.²²

significant increase in antisocial behaviour (change in standardised score of 0.64, $p = 0.02$). Among people who were not mistreated, the low MAOA activity variant tended to be associated with a decrease in antisocial behaviour (change in score of -0.17 , $p = 0.14$). For this gene the direction of the effect differed by environmental exposure (table 3). Similar findings of differences in direction of an environmental effect in different variant carriers, or different genotype effects within different environmental strata, have been reported in other recent publications.^{15, 16} In one of these examples, a polymorphism for alcohol dehydrogenase was associated with different effects on HDL cholesterol at different levels of alcohol consumption.¹⁵ In the second, risk of obesity associated with carbohydrate consumption differed between carriers of alternative polymorphisms for the β_2 adrenoceptor gene.¹⁶

DISCUSSION

In the examples cited, the inclusion of measures of both environmental factors and genes has enhanced the search for gene-disease or environment-disease associations by increasing statistical power. That is, even though the sample size was smaller within the environmental subgroups, the larger magnitude of the gene effect under specific environmental conditions enabled the gene-disease association to be detected. This finding of a different gene-disease association in different strata of the environmental factor is consistent with research in other species. The phenotypes that constitute

many classic "mutations" in *Drosophila* and other species are in fact the result of genotype-environment interactions. Their expression depends on factors such as temperature and diet. The importance of such interactions in naturally occurring variation has long been understood by evolutionary biologists¹⁷⁻¹⁹ and there is no clear reason why humans should be an exception.

The evidence from the human examples presented here suggests that the broader inclusion of environmental measures in genetic studies, aimed at gene discovery, may be justified even if statistical power only is considered. Other issues, although not the focus of this paper, also support a move towards the inclusion of environmental measures in population based genetic studies. A consideration of the environmental contexts in which genetic susceptibility is suspected to be most evident could lead to confirmation of a hypothesised gene-environment interaction and thus may increase the confidence that the genetic or environmental factor under study is a true cause of the disease. This will assist in the identification of genuine gene-disease associations, in a field where false positive genetic association studies have been a major problem.³ Environmental determinants of disease may also confound gene-disease associations and where negative confounding occurs the association will be obscured. Without concurrent measures of these environmental factors adequate control for confounding will not be possible.

Table 3 The effect of MAOA activity on antisocial behaviour depends on childhood history of maltreatment ($p = 0.01$ for interaction)¹¹

Childhood maltreatment status	Increase in standardised antisocial behaviour score among subjects with low MAOA activity levels compared with other subjects*	p Value for change in behaviour score†
All subjects (n = 442)	0.01	$p = 0.89$
None (64%)	-0.17	$p = 0.14$
Probable (28%)	0.17	$p = 0.11$
Severe (8%)	0.64	$p = 0.02$

*The antisocial behaviour score has been standardised over the entire group of 442 subjects to have a mean of 0 and standard deviation of 1; †p values calculated using *t* tests

The future search for genes within a context that includes environmental factors as covariates may involve different research strategies. If environmental data are difficult to obtain retrospectively, cohort studies, such as BioBank UK, will be required.²⁰ Recognition of the potential importance of including environmental factors in genetic studies has also led to the development of new analytical methods for gene linkage studies that include environmental information.²¹

In conclusion, epidemiologists and others involved in human studies should be aware that the concurrent collection of environmental as well as genetic data in studies of disease aetiology could enhance study informativeness and validity in several ways, including a possible increase in statistical power. A focus on very large population based genetic studies without consideration of the environmental context may not provide the level of study informativeness required to identify and understand how genes contribute to disease.

Authors' affiliations

T Dwyer, A-L Ponsonby, J Stankovich, L Blizzard, S Easteal, Menzies Centre for Population Health, University of Tasmania, Australia

A-L Ponsonby, National Centre for Epidemiology and Population Health, Australian National University

S Easteal, John Curtin School of Medical Research, Australian National University

Funding: none.

Conflicts of interest: none declared.

Correspondence to: Professor T Dwyer, Menzies Centre for Population Health Research, University of Tasmania, Private Bag 23, Hobart, Tasmania 7001, Australia; t.dwyer@utas.edu.au

Accepted for publication 21 October 2003

REFERENCES

- 1 **Khoury MJ**, Dorman JS. The human genome epidemiology network (The HuGE Network). *Am J Epidemiol* 1998;**148**:1–3.

- 2 **Peto J**. Cancer epidemiology in the last century and the next decade. *Nature* 2001;**411**:390–5.
- 3 **Anonymous**. Freely associating. *Nat Genet* 1999;**22**:1–2.
- 4 **Clayton D**, McKeigue PM. Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet* 2001;**358**:1356–60.
- 5 **Wang X**, Zuckerman B, Pearson C, *et al*. Maternal cigarette smoking, metabolic gene polymorphism, and infant birth weight. *JAMA* 2002;**287**:195–202.
- 6 **Kramer MS**. Intrauterine growth and gestational duration determinants. *Pediatrics* 1987;**80**:502–11.
- 7 **Xu X**, Kelsey KT, Wiencke JK, *et al*. Cytochrome P450 CYP1A1 MspI polymorphism and lung cancer susceptibility. *Cancer Epidemiol Biomarkers Prev* 1996;**5**:687–92.
- 8 **Menegon A**, Board PG, Blackburn AC, *et al*. Parkinson's disease, pesticides, and glutathione transferase polymorphisms. *Lancet* 1998;**352**:1344–6.
- 9 **Tanner CM**. The role of environmental toxins in the etiology of Parkinson's disease. *Trends Neurosci* 1989;**12**:49–54.
- 10 **Their R**, Bruning T, Roos PH, *et al*. Markers of genetic susceptibility in human environmental hygiene and toxicology: The role of selected CYP, NAT and GST genes. *Int J Hyg Environ Health* 2003;**206**:149–71.
- 11 **Caspi A**, McClay J, Moffitt TE, *et al*. Role of genotype in the cycle of violence in maltreated children. *Science* 2002;**297**:851–4.
- 12 **Widom CS**. The cycle of violence. *Science* 1989;**244**:160–6.
- 13 **Cases O**, Seif I, Grimsby J, *et al*. Aggressive behavior and altered amounts of brain serotonin and norepinephrine in mice lacking MAOA. *Science* 1995;**268**:1763–6.
- 14 **Brunner HG**, Nelen M, Breakefield XO, *et al*. Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science* 1993;**262**:578–80.
- 15 **Hines LM**, Stamper MJ, Jing Ma PH, *et al*. Genetic variation in alcohol dehydrogenase and the beneficial effect of moderate alcohol consumption on myocardial infarction. *N Engl J Med* 2001;**344**:549–55.
- 16 **Martínez JA**, Corbalán MS, Sánchez-Villegas A, *et al*. Obesity risk is associated with carbohydrate intake in women carrying the Gln27Glu β_2 -adrenoceptor polymorphism. *J Nutr* 2003;**133**:2549–54.
- 17 **Leips J**, Mackay TF. Quantitative trait loci for life span in *Drosophila melanogaster*: interactions with genetic background and larval density. *Genetics* 2000;**155**:1773–88.
- 18 **Lynch M**, Walsh B. *Genetics and analysis of quantitative traits*. Massachusetts: Sinauer, 1998.
- 19 **Schmalhausen II**. *Factors of evolution. The theory of stabilizing selection*. Chicago: University of Chicago Press, 1949.
- 20 **Banks E**, Meade T on behalf of the Protocol Development Committee for BioBank UK. Study of genes and environmental factors in complex diseases. *Lancet* 2002;**359**:1156.
- 21 **Gauderman WJ**, Morrison JL, Siegmund KD. Should we consider gene x environment interaction in the hunt for quantitative trait loci? *Genet Epidemiol* 2001;**21**(suppl 1):S831–6.
- 22 **Woolf B**. On estimating the relationship between blood group and disease. *Ann Hum Genet* 1955;**19**:251–3.