Published in final edited form as:
Int J Epidemiol. 2015 August ; 44(4): 1130–1132. doi:10.1093/ije/dyv056.

Understanding the Candidate Gene × Environment Interaction Debate: Epistemological or evidential divide?

Marcus R. Munafò[1,2]

[1]MRC Integrative Epidemiology Unit (IEU) at the University of Bristol, Bristol, United Kingdom

[2]UK Centre for Tobacco and Alcohol Studies, School of Experimental Psychology, University of Bristol, Bristol, United Kingdom

It is clear that the candidate gene × environment (cG×E) interaction literature has proved controversial, perhaps nowhere more so than in relation to the putative moderating effect of 5-HTTLPR genotype on the association between exposure to stressful life events and the subsequent development of major depressive disorder. Tabery [1] argues that much of this controversy can be traced to the distinct scientific traditions of, on the one hand, identifying genetic and environmental sources of variance (after Fisher) and, on the other hand, understanding biological mechanisms (after Hogben). There is certainly some truth in this, reflected in a number of reviews and commentaries from either side of this epistemological divide. However, an alternative (or additional) perspective is that the approaches that Tabery characterizes as exemplifying these two traditions – genomewide association studies (GWAS) for variance partitioning and candidate gene studies for mechanism elucidation – differ fundamentally in evidential value, in other words their ability to produce robust, reproducible findings.

It is now generally accepted that candidate gene studies have produced very few reproducible findings, for two main reasons. First, the sample sizes used were simply too small to detect the very small effects that we now know to expect in relation to common genetic variation [2]. The resultant low statistical power meant that most claimed associations very either false, or at the very least reflected grossly over-estimated any true associations [3, 4]. Second, candidate gene studies were predicated on the known (or presumed) neurobiology of the phenotype of interest. In some cases, this approach was successful; for example, reduced or null activity variants in genes encoding drug metabolizing enzymes (e.g., *CYP2A6* for nicotine and *ALDH2* for alcohol) appear to have relatively strong effects on the corresponding drug consumption behaviours [5]. However, in many cases it seems our knowledge was incomplete – many loci identified by subsequent GWAS are located within genes that were *not* prominent candidate genes. This same basic conclusion appears to hold regardless of whether the phenotypes investigated are diagnostic categories or mechanistic intermediate phenotypes [6, 7].

Corresponding author: Marcus R. Munafò, School of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol BS8 1TU, United Kingdom. T: +44.117.9546841; F: +44.117.9288588 marcus.munafo@bristol.ac.uk.

Europe PMC Funders Author Manuscripts

The reasons for the demonstrable superiority of GWAS over candidate gene studies as a means of identifying common genetic variants associated with complex behavioural and disease phenotypes are themselves interesting. The combination of a profound multiple testing burden (given the very large number of variants simultaneously tested within GWAS), growing awareness that the effects associated with common variants would be very small, and the poor reproducibility of many of the claimed associations reported in the candidate gene literature necessitated an important cultural change, from research conducted in individual laboratories to research conducted collaboratively across large, international consortia. Applying a common, hypothesis-free analytical approach across multiple studies (thereby achieving large sample sizes and adequate statistical power), and meta-analysing and reporting all results (thereby protecting against publication bias), provides a relatively unbiased source of information, in stark contrast to the output of candidate gene studies [8]. As a result, GWAS has been remarkably successful in identifying common genetic variants associated with a range of complex disease and behavioural phenotypes.

Of course, the success of GWAS in identifying genetic main effects does not mean that genetic influences do not moderate the impact of environmental exposures – on the contrary, the interplay of genetic and environmental influences is likely to be common. A trivial example is cigarette smoking – a genetic influence on heaviness of smoking can only be expressed if one becomes a smoker in the first place (i.e., is exposed to smoking). What is at issue is whether the current evidence for *specific* cG×E effects is sufficiently robust – in other words has the cG×E literature fared any better than the wider candidate gene literature? The principal argument of those critical of the cG×E literature is simply that it does not, and instead recapitulates many of the limitations of the wider candidate gene literature. Sample sizes are for the most part too small to detect credible effects, while the background testing of multiple genotypes, environmental exposures and their interactions, with only selected effects (typically those that achieve $P < 0.05$) ultimately being reported, leads to strong publication bias [9].

Those critical of the cG×E literature argue that it has followed the same trajectory as the wider candidate gene literature: initial promise and excitement, followed by numerous failures to replicate key findings and a growing appreciation that putative findings generated by this literature are simply not robust. Special pleading – that large studies bring inadequate phenotypic characterization, for example, or that that the effects observed may be specific to certain populations – is common, again similarly to wider the candidate gene literature, but rarely empirically justified [10]. In other words, it is not yet able to provide useful insights, either in terms of population risk or in terms of mechanism. Fundamentally, we need to know whether cG×E effects hold robustly at the population level, since all cG×E results reported to date apply inferential statistics designed to investigate group-level effects [11]. It may be that most cG×E effects are in fact so unique, idiosyncratic and multi-factorial as to essentially be stochastic, and therefore simply not amenable to analysis within this framework. However, if this is the case then many cG×E effects will remain essentially unknowable, at least with the statistical tools currently available to us [11]. Reducing the debate to one between, on the one hand statisticians, and on the other biologists, neglects the fact that the latter still ultimately rely on Fisher's venerable P-value in order to draw inferences.

If the evidence for specific cG×E effects is weak, why does the belief in them persist? One explanation is that those critical of the relevant literature are simply mistaken. Another (as argued by Tabery) is that these differing interpretations reflect different methodological traditions. However, belief in a particular literature may depend on factors beyond the evidence itself. For example, authors of studies that report evidence in favour of a particular phenomenon are more likely to believe that an ambiguous meta-analysis supports their conclusions, compared with methodologists (who, presumably, are more impartial) [12]. Selective citation of studies that support a phenomenon, compared to studies that do not, can also give a distorted impression of the strength of evidence for that phenomenon [13]. There is an extensive literature on the distorting impact of financial vested interests on scientific findings, but less attention has been paid to the role of confirmation bias and other psychological factors in science. We are understandably reluctant to retreat from our publicly-stated positions, and interpret ambiguous evidence in a way that supports our preconceived beliefs, often unconsciously. It is therefore unsurprising that the debate around the cG×E literature has become so polarized, since this process is presumably operating on both sides of the epistemological divide described by Tabery.

What hope is there for cG×E research? Here, the answer is simple – our candidates should generally be identified via GWAS, rather than via our existing understanding of the relevant biology. Only under very specific circumstances would a G×E effect be present in the absence of a detectable main effect of genotype [14]. Loci identified via GWAS can then be pursued further, to both explore underlying mechanisms and elucidate the population health impact of genetic and environmental risk factors. This approach has been successfully applied to *FTO* in the context of physical activity and obesity risk [15], and recently extended to include multiple loci combined in a genetic risk score [16]. G×E effects that hold at the population level may be rare or very small [17], but the message is clear – we will not reliably identify them through the *a priori* selection of candidate genes for investigation. However, even here, we will need to be wary of issues such as the scale-dependence of interaction effects [14], simply because effects that can appear or disappears as a function of the statistical model employed are ultimately likely to be uninteresting. We therefore need to be wary of G×E research that is informed by GWAS but nevertheless recapitulates many of the other problems that beset the cG×E literature.

Acknowledgements

MRM is a member of the UK Centre for Tobacco Control Studies, a UKCRC Public Health Research: Centre of Excellence. Funding from British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, and the National Institute for Health Research, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged. This work was supported by the Medical Research Council (MC_UU_12013/1, MC_UU_12013/6).

References

1. Tabery J. Debating interaction: the history, and an explanation. Int J Epidemiol. in press.
2. Burton PR, et al. Size matters: just how big is BIG?: Quantifying realistic sample size requirements for human genome epidemiology. Int J Epidemiol. 2009; 38(1):263–73. [PubMed: 18676414]
3. Button KS, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci. 2013; 14(5):365–76. [PubMed: 23571845]

4. Ioannidis JP. Why most discovered true associations are inflated. Epidemiology. 2008; 19(5):640–8. [PubMed: 18633328]
5. Wang JC, Kapoor M, Goate AM. The genetics of substance dependence. Annu Rev Genomics Hum Genet. 2012; 13:241–61. [PubMed: 22703173]
6. Flint J, Munafo MR. The endophenotype concept in psychiatric genetics. Psychol Med. 2007; 37(2): 163–80. [PubMed: 16978446]
7. Munafo MR, Flint J. The genetic architecture of psychophysiological phenotypes. Psychophysiology. 2014; 51(12):1331–2. [PubMed: 25387716]
8. Gaunt TR, Davey Smith G. eNOS and coronary artery disease: Publication bias and the eclipse of hypothesis-driven meta-analysis in genetic association studies. Gene. 2014
9. Duncan LE, Keller MC. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. Am J Psychiatry. 2011; 168(10):1041–9. [PubMed: 21890791]
10. Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes. Lancet. 2003; 361(9360):865–72. [PubMed: 12642066]
11. Davey Smith G. Epidemiology, epigenetics and the 'Gloomy Prospect': embracing randomness in population health research and practice. Int J Epidemiol. 2011; 40(3):537–62. [PubMed: 21807641]
12. Panagiotou OA, Ioannidis JP. Primary study authors of significant studies are more likely to believe that a strong association exists in a heterogeneous meta-analysis compared with methodologists. J Clin Epidemiol. 2012; 65(7):740–7. [PubMed: 22537426]
13. Greenberg SA. How citation distortions create unfounded authority: analysis of a citation network. BMJ. 2009; 339:b2680. [PubMed: 19622839]
14. Taylor AE, et al. Mendelian randomization in health research: Using appropriate genetic variants and avoiding biased estimates. Economics & Human Biology. 2014; 13:99–106. [PubMed: 24388127]
15. Kilpelainen TO, et al. Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. PLoS Med. 2011; 8(11):e1001116. [PubMed: 22069379]
16. Ahmad S, et al. Gene x physical activity interactions in obesity: combined analysis of 111,421 individuals of European ancestry. PLoS Genet. 2013; 9(7):e1003607. [PubMed: 23935507]
17. Yang J, et al. FTO genotype is associated with phenotypic variability of body mass index. Nature. 2012; 490(7419):267–72. [PubMed: 22982992]