

Missing heritability and stochastic genome alterations

Henry H. Q. Heng

A recent Viewpoint article (Eichler, E. E. *et al.* Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Rev. Genet.* **11**, 446–450 (2010))¹ proposes several explanations for the missing heritability in complex traits, including epigenetics, rare alleles, copy number variants, complex genome architecture and environmental uncertainty. I argue that, according to the ‘genome theory’², the missing heritability lies in stochastic genome alterations that are associated with unstable genomes and environmental insults and not by mutations in specific sets of genes.

The gene theory states that defective genes are the main cause of disease and should be readily identifiable. By contrast, the genome theory states that the interaction of the genome (genes plus genomic topology) with the environment defines the genetic system, and that most individual genes are neither independent information units nor common factors in disease.

Several considerations justify looking at the entire genome when searching for missing heritability. First, the genome and the gene have different or even conflicting functions (the genome defines and constrains a system, whereas genes modify and diversify it)³. Second, macro-evolutionary selection acts on the package (genome) rather than

the parts (genes)². And third, the *de novo* locus-specific rate of genomic rearrangement is at least 100- to 10,000-fold greater than the rate of point mutations⁴.

Finally, several common diseases, including autism and Alzheimer’s disease, are associated with altered karyotypes^{5,6}. This observation suggests similarities in the bases of common disease and cancer⁷, which arise from stochastic genome alterations rather than common gene mutations^{8,9}. Unlike Mendelian diseases, in which highly penetrant genetic defects are detectable within a patient population, the molecular evolution of cancer can only be explained by including all molecular mechanisms in the entire patient population^{10,11}. Similar to cancer, the progression of many common diseases is influenced by genome–environment–time interactions. Clearly, most common diseases are not caused by commonly shared genetic and environmental factors^{7,10}.

If we interpret common diseases as being the result of stochastic genome alterations, then missing heritability can be found at three transitional levels: first, between the population and the individual (many rare genome variations stochastically occur during each generation but are disregarded at the population level as ‘insignificant noise’); second, between germline and somatic cells

(most diseases occur at the somatic cell level and so there is increased variation between germline and somatic cells); and third, between normal and diseased cells (the genome integrity typical of normal cells is often lost in pathological conditions).

In TABLE 1 genetic diseases are classified into four categories based on their degree of genome instability and genetic penetrance. As most genome alterations are individually rare, the rationale of identifying average profiles among populations and applying them back to individuals is problematic. Establishing new strategies to monitor system instability and to identify the key level of genetic organization responsible for diseases is now urgently needed. The information from individual cells and cell populations should be combined and patterns of system dynamics compared (including divergent or convergent phases). These strategies will deliver the benefit of increasing our understanding of numerous diseases whose phenotypes are very different but whose underlying cause, genomic instability, may be the same.

Henry H. Q. Heng is at the Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, Detroit, Michigan 48201, USA.
e-mail: hheng@med.wayne.edu

doi:10.1038/nrg2809-c3

Published online 13 October 2010

Table 1 | Genetic disease classification

Disease type	Genetic factors in patient populations	Disease type prevalence	Relative genome stability	Examples
A	Commonly shared	Infrequent	Stable	Typical Mendelian diseases (for example, sickle cell anaemia)
B	Commonly shared	Infrequent	Unstable	Familial cancer syndromes
C	Rare	Infrequent	Stable	Charcot–Marie–Tooth neuropathy
D	Rare	Prevailing	Unstable	Sporadic cancers and neurological and/or behavioural disorders

Disease types A and B have been associated with higher penetrance in patient populations; these have been the primary focus of research into the causes of inherited and complex disease, and have been the targets of genome-wide association studies. For type C diseases, personal genome sequencing is useful¹². For type D diseases, whole-genome sequencing might not be clinically useful, as the stochasticity of genome alteration reduces clinical predictability based on genetic profiles. If genome-wide association studies are to be applied to types C and D, the use of smaller, homogenous subpopulations will be key. Notably, an observation that is often ignored in common diseases is that abnormal genomes often present in mosaic and not in clonal form^{5,6}.

- Eichler, E. E. *et al.* Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Rev. Genet.* **11**, 446–450 (2010).
- Heng, H. H. The genome-centric concept: resynthesis of evolutionary theory. *Bioessays* **31**, 512–525 (2009).
- Heng, H. H. Elimination of altered karyotypes by sexual reproduction preserves species identity. *Genome* **50**, 517–524 (2007).
- Lupski, J. R. Genomic rearrangements and sporadic disease. *Nature Genet.* **39**, S43–S47 (2007).
- Ye, C. J., Liu, G., Bremer, S. W. & Heng, H. H. The dynamics of cancer chromosomes and genomes. *Cytogenet. Genome Res.* **118**, 237–246 (2007).
- Iourov, I. V., Vorsanova, S. G. & Yurov, Y. B. Chromosomal mosaicism goes global. *Molecular Cytogenet.* **1**, 26 (2008).
- McClellan, J. & King, M. C. Genetic heterogeneity in human disease. *Cell* **141**, 210–217 (2010).
- Heng, H. H. *et al.* Stochastic cancer progression driven by non-clonal chromosome aberrations. *J. Cell. Physiol.* **208**, 461–472 (2006).
- Heng, H. H. Cancer genome sequencing: the challenges ahead. *Bioessays* **29**, 783–794 (2007).
- Heng, H. H. *et al.* The evolutionary mechanism of cancer. *J. Cell. Biochem.* **109**, 1072–1084 (2010).
- Ye, C. J. *et al.* Genome based cell population heterogeneity promotes tumorigenicity: the evolutionary mechanism of cancer. *J. Cell. Physiol.* **219**, 288–300 (2009).
- Lupski, J. R. *et al.* Whole-genome sequencing in a patient with Charcot–Marie–Tooth neuropathy. *N. Engl. J. Med.* **362**, 1181–1191 (2010).

Competing interests statement

The author declares no competing financial interests.

FURTHER INFORMATION

Henry H. Q. Heng’s homepage:
<http://genetics.wayne.edu/faculty/heng/index.php>