A Case for Core Genes

Chunyu Liu[1*]

[1] Department of Psychiatry, University of Illinois at Chicago.

*Correspondence: Chunyu Liu, Email: liucy@uic.edu.

Gene networks have been used to investigate the complex relationships between genetic factors in disease risk. Functional modules within these networks contain some "core" or "hub" genes, which are the most highly connected "nodes". They are related to the other genes in the module one way or another. The idea of "guilt by association" suggests that the "peripheral genes" associated with these core genes play an indirect role in disease risk (e.g. via regulation) when the network module is associated with disease risk. Boyle *et al.*[1] further exploits this phenomenon to propose the omnigenic model, which posits that almost any active gene with regulatory variants in at least one tissue associated with a disease contributes to disease risk. Here, we advocate an emphasis on core genes as a more practical approach than an omnigenic view.
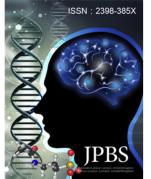
Boyle *et al.* expounds on the idea that the combined effects of peripheral genes on disease risk are greater than the disease risk associated with core genes because the former far outnumber the latter in quantity. The paper seeks to highlight the importance of genes that may go overlooked because they are less connected in gene modules. The omnigenic argument is further supported by the evidence that risk factors for a single disease are distributed across the genome rather than clustered in a few pathways.

Indeed, we do not expect core genes or any single module alone to explain the total cause of any complex disease. Still, core genes will likely prove ultimately more efficient and meaningful than the broader-lensed view of the omnigenic model. Regardless of whether or not the theory of omnigenics is valid, core genes are likely the main drivers of structure and stability of a given network module.

The "edges," or links between genes in a network, can represent interactions at the level of DNA, RNA, or protein, as well as protein-protein interactions, regulatory relationships, etc. Different gene associations carry different weights, some of which may barely be significant, and some genes are directly associated, while others may be linked by several degrees of separation. Because there exist so many peripheral genes with variable qualities, the task of calculating and interpreting how they could all fit together in a functional sense is likely to be a highly daunting and difficult statistical task.

Interpreting the significance of the connections between core and other genes is also certainly fraught with complication. However, their connections to many other genes make for greater certainty of a core gene's significance in a function. While the additive effects of core genes' expressions alone may not account for the majority of risk variance, more insight can likely be gained by altering expression of

a core gene than by altering the expression of any one peripheral gene. For example, it can be argued that knocking out or knocking down a core gene will have more drastic effects on a network's overall expression than will doing the same thing to a gene with fewer connections. In a similar vein, it would be impractical to try to design a drug targeting all risk areas identified for a disease. Targeting fewer nodes that are the most likely to lead to altered expression in other areas might be more effective.

While a focus on core genes is not expected to provide a complete understanding of genetic etiology, we contend that core genes are the best place to begin investigating functions of gene networks and their relevance to psychiatric genetics, in large part because core genes are likely the strongest candidates for effective drug targets. Going forward, we believe that the structure, function, and behavior of biological networks can best be understood through the identification and analysis of subnetworks built around core genes.

Questions for Dr. Jonathan Pritchard:

Please clarify whether the following conclusions are what were intended by the omnigenic model. These have been the interpretations many readers have drawn about the model:

1) According to the omnigenic model, every active gene expressed in a disease-associated tissue contributes to a phenotype. In the case of most psychiatric disorders, this would mean that all active regions expressed in brain contributes to disease risk.

2) Peripheral genes are more important than hub genes in gene networks because they make a greater collective contribution.

3) What positive improvements and valuable findings would the application of the omnigenic model bring to the study of the genetics of complex disorders.

REFERENCES

1. Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell. 2017; 169(7): 1177-1186.