Systematic analysis of pleiotropy during *C. elegans* early embryogenesis

Lihua Zou[1,2], Brian Ross[3], Sira Sriswasdi[3], Patrycja E. Missiuro[3], Jun Liu[1] and Hui Ge[3]

[1]Department of Statistics, Harvard University, Cambridge MA 02138, USA;[2]Dana-Farber Cancer Institute, Harvard Medical School, Boston MA 02115, USA;[3]Whitehead Institute, 9 Cambridge Center, Cambridge MA 02142, USA

Pleiotropy refers to the phenomenon of a single gene controlling multiple distinct and seemingly unrelated phenotypic effects. Pleiotropy reflects the fact that some genes in the genome perform multiple biological functions. Traditionally, loss-of-function phenotypes are examined for individual genes. The recent availability of high-throughput loss-of-function datasets may lead to the opportunity for systematic identification of pleiotropic genes. However, it is still not clear how effects arising from loss of a single function can be differentiated from the effects of losing multiple functions. It is also a challenge to provide mechanistic interpretations for such phenotypic complexity.

C. elegans is especially amenable to genome-wide loss-of-function analysis because of its known genome sequence, well-characterized anatomy and the convenience of RNA interference (RNAi) technique. The *C. elegans* early embryo is an excellent system for studying mitotic cell divisions. Sonnichsen *et al* performed whole-genome RNAi experiments and identified 661 genes involved in early embryogenesis [1]. A series of phenotypic characters were used to represent the absence and presence of cellular defects. By comparing the number of cellular defects each gene shows in the phenotypic profiles with that of randomly permuted datasets, we find that significantly more genes exhibit complex phenotypes during *C. elegans* early embryogenesis.

We aim to uncover pleiotropic genes during *C. elegans* early embryogenesis. Using pre-defined functional categories as seeds, we identify the sets of phenotypic characters, or "signatures" of defects, which best represent each functional category. We generate gene modules which contain these signatures. Some genes are grouped into one module while others are grouped into multiple modules. We define the number of phenotypic modules a gene belongs to as its pleiotropy index. Many kinases, for example, are highly pleiotropic according to the index.

We investigate whether a global mechanistic interpretation can be reached for the phenomenon of pleiotropy. Since many cellular events in early development may be mediated by protein-protein interactions, we examine the properties of highly pleiotropic proteins in interactome networks [2]. We develop a "circuit between-ness" model to estimate the extent to which a given protein contributes to connecting the remaining proteins in the interactome. We find that highly pleiotropic proteins tend to show higher circuit between-ness and lower density of local neighborhood compared to other essential proteins involved in early embryogenesis. We propose that highly pleiotropic proteins act as "module connectors" in interactome networks, which is a fundamental reason for their loss-of-function phenotype complexity.

[1] Sonnichsen B et al (2005). Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. Nature 434:462-469.
[2] Li S et al (2004). A Map of the Interactome Network of the Metazoan *C. elegans*. Science 303:540-543.