This Month in Genetics

Kathryn B. Garber[1,*]

A Natural Experiment

During pharmaceutical development, researchers might spend substantial time and effort on a candidate compound that has the desired effect on a protein target only to find out that it does not have the desired outcome on the whole body. To potentially avoid this type of dead end, one can start with a natural experiment to decide whether a particular biomarker is a good target for drug development. In a recent example of this approach, two groups used natural outliers for a biomarker of interest to explore its outcome. These population-based studies focused on apolipoprotein C3 (ApoC3), which inhibits hydrolysis of triglyceride-containing lipoproteins and thereby increases plasma triglyceride levels. Rare genetic variation that lowers ApoC3 functionality would thus be expected to reduce triglyceride levels. Although an association between triglyceride levels and cardiovascular disease was already known, these groups wanted to see whether a lifetime of low triglyceride levels would translate into a lower risk of cardiovascular disease. Both studies found a similar 40% reduction in risk in groups of individuals heterozygous for one of four *APOC3* alleles with reduced function. This natural experiment makes ApoC3 a plausible pharmacological target for reducing risk of heart disease. In fact, antisense oligonucleotides to *APOC3* are being explored for this purpose.

The TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute. (2014). N. Engl. J. Med. 371, 22–31.

Jorgensen et al. (2014). N. Engl. J. Med. 371, 32–41.

Mixed Up

In each human cell, hundreds of copies of the mtDNA sequence are present, and they are not necessarily identical. This phenomenon of a mixture of sequences, known as heteroplasmy, contributes to the unique inheritance of mtDNA variation and makes it difficult to precisely predict the inheritance of mitochondrial disease. Beyond these affected families, there are limited data for estimating the prevalence of heteroplasmy, but recent studies suggest that it might be widespread in the general human population. This is borne out in the recent work by Ye et al., who used sequence data on more than 1,000 individuals from the 1000 Genomes Project. They report that 90% of this population has detectable heteroplasmy. Although most people have low-level heteroplasmy, sequences are in a ratio as high as 90/10 in 44% of the sample. Disease-associated sites in mtDNA are overrepresented in this heteroplasmy and seem to be subject to purifying selection. However, selection is less efficient than it would be on nuclear genetic variation, so it does not effectively remove disease-associated variation. As cells divide over a lifetime, the level of heteroplasmy could shift toward pathogenic variation, and the authors wonder whether this has the potential to contribute to age-related diseases.

Ye et al. (2014). Proc. Natl. Acad. Sci. USA. Published online July 7, 2014. http://dx.doi.org/10.1073/pnas.1403521111.

Do We Have the Power?

Is one mutation like any other? That is, is the disease risk conferred by any pathogenic variant the same as that conferred by any other in the same gene? Biologically, this is not universally true because there can be a range of activity from null to normal for a gene product. Practically, however, we almost always lump together all pathogenic variation in the same gene and assign it the same level of disease risk. Through a series of power calculations, Shirts et al. considered the feasibility of instead examining each variant individually. This allowed them to estimate the sample sizes needed for classifying rare variation in cancer genes and accurately estimating relative risk associated with these individual variants. With population-based strategies, tens of thousands to millions of subjects will be needed for classifying cancer variants. Required sample sizes are much lower for family-based strategies, but these would hinge on the ability to recruit the right individuals. Their results leave Shirts et al. to conclude that much of this rare missense variation will remain unclassified in the near term.

Shirts et al. (2014). Genet. Med. 16, 529–534.

Forward and Back

Lacking some of the mechanisms that eukaryotic cells use to generate diversity, some viruses use ribosomal frameshifting to produce alternate protein domains. The same type of mechanism in higher eukaryotes hasn't been documented but—on the basis of what we know from frameshift mutations—would instead be likely to result in nonsense-mediated decay when a premature stop codon is encountered in the −1 reading frame. This

[1]Department of Human Genetics, Emory University School of Medicine, Atlanta, GA 30322, USA
*Correspondence: kgarber@genetics.emory.edu
http://dx.doi.org/10.1016/j.ajhg.2014.07.010.

prediction is borne out in work by Belew et al., who document so-called programmed ribosomal frameshift (PRF) events in *CCR5*, a gene most famous for encoding an HIV coreceptor. *CCR5* was a likely candidate for PRF on the basis of the juxtaposition of a homopolymeric sequence on which the ribosome could slip and a sequence that would be predicted to form a complicated RNA secondary structure. In cell culture, placing this sequence upstream of a reporter gene put in a −1 frame led to reporter expression with an efficiency of approximately 10%. The PRF mechanism is stimulated by a miRNA that interacts with the *CCR5* PRF signal sequence, forming a triplex structure. As expected, shifting the reading frame resulted in decreased CCR5 and its mRNA via nonsense-mediated decay. Additional experiments suggest that PRF also occurs on other cytokine receptors, but how widespread this novel mechanism of gene regulation is in mammals is still unclear.

Belew et al. (2014). Nature. Published online July 9, 2014. http://dx.doi.org/10.1038/nature13429.

A Vulnerable Gene

Haploinsufficiency for any of a number of genes that encode ribosomal proteins causes Diamond-Blackfan anemia (DBA). How a defect in a crucial cellular process leads to a phenotype that is largely limited to erythrocytes has puzzled researchers. The first clue came when mutations in the gene encoding the transcription factor GATA1were found in a small number of individuals who had been clinically diagnosed with DBA. Ludwig et al. used these individuals to understand the relationship between protein translation and *GATA1*. They discovered that this gene is particularly sensitive to reduced ribosome abundance, probably because of its highly structured 5′ UTR. As such, typical DBA-causing mutations result in reduced GATA1 levels and a downstream reduction in the expression of GATA1 target genes. Expressing *GATA1* in bone marrow cells from these individuals increases production and differentiation of erythroid cells, so it appears that this is a major downstream contributor to the DBA phenotype.

Ludwig et al. (2014). Nat Med. 20, 748–753.

This Month in Our Sister Journals

Using LD to Infer Population History

The extent of linkage disequilibrium (LD) is influenced by natural selection, mutation and recombination rates, and population size. Our current understanding of the response of LD to changes in population size is rather crude, affecting our ability to make inferences on the basis of LD measures. Rogers looked in depth at the effects of changes in population size on LD measures and found different signatures in the LD curves when a population declines or grows and when it expands after a period of equilibrium versus after a temporary decline in size. Comparison of these curves to data from a European population suggests a history of population growth.

Rogers. (2014). Genetics. Published online June 6, 2014. http://dx.doi.org/10.1534/genetics.114.166454.