

It's all druggable

Current medicines are small chemicals that target the activity of proteins in the body, resulting in therapeutic and off-target changes in metabolism and gene expression. Now, however, gene expression levels can be raised or lowered at almost any point in the genome by sequence-targeted therapeutics, and the effects of these agents can be tested in preclinical models generated by gene editing of experimental animals and human cells. Genome-wide association studies (GWAS) of disease predisposition, metabolism and gene expression have a key role in explaining how current protein-targeting drugs work and in using the regulatory variation in human genomes to guide the therapeutic future of targeted gene regulation.

On page 193 (see also p. 175), Nir Hacohen and colleagues use gene editing of human cells to identify two new host pathways (target sulfation and cell aggregation) that can be interrupted to block HIV-1 infection. There are many ways that these discoveries might be advanced to clinical trial, via small molecule drugs, monoclonal antibodies or somatic genome editing. However, a general approach that opens the whole genome to therapeutic intervention makes use of sequence-targeted oligonucleotides. Recently, the US Food and Drug Administration (FDA) approved the use of modified oligonucleotides to treat pediatric spinal muscular atrophy by upregulation of *SMN2* gene expression (*Science* <http://dx.doi.org/10.1126/science.aal0476>, 2016). The same principle is being tested in clinical trials to try to downregulate *HTT* expression in Huntington disease (for example, ClinicalTrials.gov [NCT02519036](https://clinicaltrials.gov/ct2/show/study/NCT02519036)). This technology is in principle very widely applicable: for example, 72% of transcriptional units are sense–antisense, such that experimental manipulation of antisense transcripts can regulate mRNA expression (*Science* [309, 1564–1566, 2005](https://doi.org/10.1126/science.1256446)). Protein levels can be readily modulated by antisense targeting of the upstream ORFs present in a proportion of protein-coding genes (*Nat. Biotechnol.* **34**, 875–880, 2016).

The 'druggable genome' is not in fact genomic but a protein-based concept (*Nat. Rev. Drug Discov.* **1**, 727–730, 2002, and *Nat. Methods* **10**, 1209–1210, 2013). Apart from vaccination, immunotherapies and extracellular targeting of proteins by means of monoclonal antibodies, most current drugs are small molecules whose competitive binding to proteins makes them specific therapeutic agonists or antagonists. As such, their interactions are a particular subset of the interventions that are possible, for surrounding each existing drug target are other members of its functional biochemical pathway. Raising or lowering the activity of some of these will produce the same effect as the drug. The gene encoding the target protein is surrounded in gene expression space by a module of co-regulated transcripts, functional RNAs and proteins. The global metabolic effects of a drug, too, can be elicited by genetically altering the expression levels of one or more genes in key tissues either by naturally occurring variants and sporadic mutations or now therapeutically.

It has been suggested that genetic evidence from rare disease-causing mutations and regulatory variants alike improves the success of conventional pharmaceutical development (*Nat. Genet.* **47**, 856–860, 2015); however, carriers of common regulatory variants with quantitative effects on gene expression are far easier to recruit for statistically powered genomic investigations. Rare Mendelian mutations, although frequently very informative, are too few and far between to be compared directly in their effects on the body with the variety of drugs undergoing trials. Common regulatory variants identified by GWAS may be much more readily compared to drug effects. For example, the common rs12916[T] allele in *HMGCR* ([NC_000005.10:g.75360714T>C](https://www.ncbi.nlm.nih.gov/variation/packages/track/locus?hgvs=NC_000005.10:g.75360714T>C)) is associated with lower expression of *HMGCR* in liver and lower plasma LDL cholesterol levels relative to the similarly common rs12916[C] variant. The enzyme encoded by *HMGCR* is the target of the cholesterol-lowering statin family of medicines. Würtz *et al.* (*J. Am. Coll. Cardiol.* **67**, 1200–1210, 2016) found a close correlation between the magnitudes of metabolite alterations over time produced by statin treatment and those resulting from genetic reduction in *HMGCR* expression due to rs12916[T] alleles (scaled to their respective on-target effect of lowering LDL cholesterol). This work suggests that GWAS data linking disease loci to tissue-specific transcription and metabolite levels via expression quantitative traits and metabolomics have a broader role in specifying drug targets, off-target effects and mechanisms of action. A preprint in *bioRxiv* (<http://dx.doi.org/10.1101/066027>, 2016) would suggest that this approach could be more generally applicable.

We should therefore be inspired but not limited by current concepts of pharmaceutical design, as the potential targets for therapeutic intervention now include almost any transcribed sequence in the genome in any tissue. The genomic and metabolic perturbations of currently effective drugs form an important nucleus of knowledge upon which we can assemble the framework of a much larger strategy. We should in consequence make sure that quantitative tissue-specific regulation of transcripts and metabolites can be connected to diseases and individual genome variation via extension of the highly reproducible and productive GWAS approaches. ■