Human Molecular Genetics, 2012, Vol. 21, Review Issue 1 **R18–R23**
doi:10.1093/hmg/dds363
Advance Access published on August 29, 2012

Functional evaluation of genetic variation in complex human traits

Derek T. Peters and Kiran Musunuru*

Department of Stem Cell and Regenerative Biology, Harvard University, Sherman Fairchild Biochemistry Bldg 160, 7 Divinity Ave., Cambridge, MA 02138, USA

Received August 16, 2012; Revised August 16, 2012; Accepted August 23, 2012

Genome-wide association studies and, more recently, next-generation sequencing studies have accelerated the investigation of complex human traits by providing a wealth of association data linking genetic variants to diseases and other phenotypic traits. These data promise to transform our understanding of the molecular pathways underlying complex human traits, but only if functional evaluation of the novel genetic variants is undertaken. Here, we review recent examples in which such functional evaluation has been attempted, with varying degrees of success, and we highlight new technological advances that should greatly enhance our ability to identify and dissect causal genotype–phenotype relationships.

INTRODUCTION

Genetic variation gives rise to the heritability of complex human traits, including predisposition to common diseases. Many disorders ranging from autoimmune conditions to neuropsychiatric and cardiovascular diseases exhibit substantial heritability, as do traits such as fasting glucose level and plasma lipid concentrations, which are proven causal risk factors for disease. Recent discoveries in the field of human genetics have provided unprecedented opportunities to elucidate the genes and molecular pathways that underlie complex traits. Genome-wide association studies (GWASs) have identified numerous novel genetic loci associated with diseases and risk factors, while exome sequencing and candidate gene sequencing studies have uncovered a number of putative causal mutations in familiar as well as novel genes.

Functionally evaluating human genetic variation that underlies complex traits requires the application of an array of experimental approaches. As this area of investigation continues to mature, pioneering functional studies have utilized traditional *in vitro* and *in vivo* techniques and model systems as well as recently developed sequencing, epigenetic and pluripotent stem cell technologies. Despite the wealth of genetic association data that have been generated in the past few years by increasing large GWASs as well as next-generation sequencing studies, to date there are relatively few examples of functional evaluation that has sought to clarify the connection between genotype and phenotype, and these examples demonstrate the advantages as well as pitfalls of the specific approaches that can be employed.

CHARACTERIZING COMMON DNA VARIANTS

GWASs identify loci in the human genome that contain common single-nucleotide polymorphisms (SNPs) that are associated with a complex trait of interest. Functional evaluation of GWAS loci begins with the search either for a causal gene within the implicated locus or for a causal DNA variant, which in turn may lead to the identification of a causal gene. We present two illustrative examples here.

Loci associated with blood lipids

GWASs of blood lipid concentrations, complex traits that are risk factors for cardiovascular disease, have identified 95 associated genomic loci, including loci containing known lipid genes, loci with established targets of lipid-lowering medications, and novel loci with no known connection to lipids (1). The last group of loci likely harbor genes that play previously unappreciated roles in the regulation of lipoprotein metabolism and have been the focus of a variety of functional studies.

One such locus, located on chromosome 1p13, contains genetic variants that are highly associated with plasma low-density lipoprotein cholesterol (LDL-C) concentration as well as myocardial infarction (MI) risk, but at the time of discovery were not in proximity to any obvious candidate causal

*To whom correspondence should be addressed. Tel: +1 6174965361; Fax: +1 6174968351; Email: kiranmusunuru@gmail.com

genes. Genetic mapping narrowed the region of LDL-C-associated genetic variation within the locus to a haplotype in an intergenic interval between the *CELSR2* and *PSRC1* genes; the haplotype was found to have a liver-specific association with the expression levels of *CELSR2*, *PSRC1* and a nearby gene, *SORT1* (2). This observation suggested that the haplotype might affect an enhancer. Extensive *in vitro* characterization in cultured hepatoma cells using luciferase reporter constructs showed that an SNP, located within a predicted C/EBP transcription-factor-binding site (TFBS), is responsible for the enhancer activity of the haplotype and is the likely causal variant. Additional *in vitro* experiments supported a genetic mechanism by which the minor allele of the causal SNP creates a C/EBP TFBS (or, conversely, the major allele disrupts the TFBS), thereby increasing *CELSR2*/*PSRC1*/*SORT1* expression in hepatocytes.

Once a putative causal gene is identified for a GWAS locus, *in vivo* overexpression and knockdown studies of the orthologous gene in mice can be a useful approach to establishing that variation in the gene's function likely underlies the GWAS trait. Motivated by the association between the suspected causal SNP and *SORT1* expression in the human liver, experiments were undertaken in which *Sort1* was overexpressed in the mouse liver via adeno-associated virus or knocked down by siRNA. These perturbations resulted in significant and opposite effects on blood LDL-C levels in the mice and together with *in vitro* studies in primary mouse hepatocytes, they strongly suggested that *SORT1* is the causal gene within the locus (2). In contrast, studies with *PSRC1* and *CELSR2* in mice argued against their involvement in lipoprotein metabolism (2).

Overexpression and knockdown of the mouse orthologs of suspected causal genes at other lipid-associated loci were used to show that *GALNT2*, *PPP1R3B* and *TTC39B* are likely to be regulators of high-density lipoprotein cholesterol levels in humans (1). In the case of *GALNT2*, this gene was the only one within the locus in question, making it the obvious candidate for functional testing. In contrast, *PPP1R3B* and *TTC39B* were prioritized for functional characterization above several other genes in their respective loci because their expression levels in the human liver were found to be associated with the genotypes of the GWAS SNPs, whereas the expression levels of the other genes were not.

Variation within a gene desert: the 9p21 locus

Characterization of the LDL-C-associated 1p13 locus illustrates how *in vitro* reporter experiments and *in vivo* mouse experiments can be used to identify a causal non-coding genetic variant and provide the foundation for the elucidation of a pathway connecting genotype to phenotype. In other instances where disease-associated variants fall within large non-coding regions of the genome, the underlying biology has proven more difficult to dissect.

Chromosome 9p21 contains common DNA variants that have been identified by various GWASs to be associated with MI risk, type 2 diabetes mellitus, melanoma or breast cancer, making 9p21 one of the most tantalizing loci identified by GWASs. MI-associated SNPs within the 9p21 locus are not associated with traditional risk factors for MI such as lipids, blood pressure and diabetes (3–5), but they are associated with abdominal aortic aneurysm and intracranial aneurysm (6), suggesting the existence of a novel pathway in vasculature that contributes to the pathophysiology of each of these diseases. These SNPs are located in a large (58 kb) gene-poor genomic locus, which includes part of a long non-coding RNA termed *CDKN2BAS* (also known as *ANRIL*), and lie more than 100 kb away from the closest protein-coding genes, *CDKN2A* and *CDKN2B*—cell cycle-related genes that encode cyclin-dependent kinase inhibitors—making the locus extremely challenging to characterize.

To investigate the role of this region, one group chose an approach based on a unique application of one of the most commonly used and powerful techniques in experimental biology: modification of the mouse genome. The investigators deleted a 70 kb non-coding interval on mouse chromosome 4 that is orthologous to the MI-associated region within the 9p21 locus, which resulted in decreased *Cdkn2a* and *Cdkn2b* expression in the mouse aorta and heart (7). Concordant with decreased expression of the cell cycle regulators *Cdkn2a* and *Cdkn2b*, knockout of the 70 kb interval resulted in increased proliferation and disrupted senescence of primary aortic smooth muscle cells in culture. However, weakening the relevance of these findings to MI risk in humans is the fact that only ~50% sequence homology exists between the deleted interval and the orthologous region in humans. Moreover, knockout of this interval had no effect on atherosclerosis progression; this could reflect either that the causal DNA variant detected by GWASs in humans may not exist in or may not be relevant in mouse or that human atherosclerotic disease is poorly phenocopied in mice—in either case, highlighting the potential limitations of studying human complex traits in non-human models.

Another group took a different approach to evaluate the 9p21 locus and its association with MI, performing an integrated analysis incorporating bioinformatics data sets and examining enhancer elements and long-range interactions involving the locus (8). First the authors analyzed DNase hypersensitivity and TFBSs as well as chromatin modification profiles—including marks associated with promoters of active genes, insulators and enhancers—in multiple human cell types to discover potential regulatory elements in the 9p21 locus. Several of these elements displayed *in vitro* enhancer activity in luciferase reporter assays. After cataloging additional DNA variants by sequencing the region in 50 individuals, they narrowed the list of candidate regulatory variants by focusing on those that fell within enhancer elements, were consistently associated with increased MI risk and disrupted consensus TFBSs. Two neighboring SNPs met these criteria, located within a predicted STAT1-binding site/enhancer preserved by the non-risk haplotype (and disrupted by the risk haplotype). Experiments in human vascular endothelial cells (HUVECs) and lymphoblastoid cell lines suggested a pathway through which interferon-gamma (IFN-γ) signals via STAT1 through the TFBS modified by the two SNPs, thereby resulting in the differential expression of *CDKN2BAS* and *CDKN2B* depending on which of the SNP alleles are present.

Extending their analysis of the STAT1 enhancer, the authors performed chromatin conformation capture in HUVECs combined with DNA selection and ligation, showing that the enhancer

participates in long-range interactions (up to a megabase) with the *CDKN2A*, *CDKN2B*, *MTAP* and *IFNA21* genes, especially upon IFN-γ treatment of the cells, suggesting a dependence on STAT1 activity. This study demonstrates the power of an integrated approach for the identification of transcriptional regulatory elements in a complex non-coding region of the genome, and it provides insight into how non-coding genetic variants can alter gene regulatory pathways, potentially at great distances. However, although the study suggests that there is interplay between inflammation, variants in the 9p21 locus and cell cycle regulation, it does not provide evidence that the implicated SNPs examined are the causal variants underlying the association of the 9p21 locus with MI or that IFN-γ signaling is the causal mechanistic link connecting the locus to the pathogenesis of MI. Indeed, no experimental approaches are undertaken to connect the STAT1 TFBS to atherosclerotic disease or some other phenotype that is directly relevant to MI.

CHARACTERIZING NOVEL CODING VARIANTS

Gene resequencing studies ask whether a gene of interest contributes to the pathogenesis of a complex human trait. Sequencing the coding region of a gene in a population may reveal a preponderance of variants present in a subset of individuals who exhibit a specific trait, suggesting an association between the gene's function and the observed trait. Whole-exome or whole-genome sequencing studies, which are unbiased, large-scale alternatives to the one-gene-at-a-time resequencing approach, aim to uncover novel causal mutations underlying a specific trait by performing next-generation DNA sequencing across the entire genome in a few affected individuals. Both approaches can provide strong evidence for an association between a specific gene and a trait; however, definitively establishing that a suspected gene contributes to the trait of interest requires functional validation. Moreover, in this context, it would be valuable to assess the functional impact of each of the actual coding variants that are discovered—some might prove to be highly debilitating to gene function, whereas others may be of no functional consequence.

In one study, the coding region of the *ANGPTL3* gene was resequenced in a population from the Dallas Heart Study to address whether *ANGPTL3* modulates blood triglyceride (TG) levels in humans (9), in light of prior evidence that the mouse ortholog plays this role. While several nonsense, frameshift and splice-site mutations were found in individuals within the lowest quartile of blood TG levels, suggesting that loss-of-function alleles are associated with low TG levels, a number of missense mutations of unclear functional consequence were also identified in individuals with low TG levels. *In vitro* experiments were performed in HEK293A cells to investigate the effects of each of the discovered mutations on ANGPTL3 expression, secretion and activity. Notably, the results of these experiments were compared with predictions made by two computational annotation programs (PolyPhen and SIFT), revealing ∼35% false-positive and 40% false-negative rates for the prediction of deleterious mutations—highlighting that, at the present time, computational methods cannot substitute for functional experiments in characterizing the effects of novel coding variants.

Functional analysis of novel variants has been successfully conducted in model organisms, where the effects of coding variants on more complex phenotypes can be assessed with moderate throughput. One group followed up on resequencing studies with functional analysis in zebrafish as well as mammalian cells to gain insight into the genetic architecture of ciliopathy disorders (10). The investigators characterized mutant versions of the human *TTC21B* gene containing novel variants identified in ciliopathy patients or in healthy controls by assaying the genes' ability to rescue a zebrafish ciliary phenotype (produced by knockdown of endogenous *Ttc21b* expression); they found that about one-quarter of the variants led to complete loss of function, about one-half resulted in partial loss of function and the remainder were benign (i.e. non-functional). Of note, the investigators found that there was no significant difference in the burden of novel *TTC21B* variants in ciliopathy cases versus controls. When they restricted the analysis to novel variants shown by their experiments to affect protein function, they found a striking enrichment of these functional variants in the cases.

These two studies highlight the utility of experimental data in understanding the relationships between genotypes and phenotypes in humans. They also reinforce the point that an appropriate experimental system for functional analysis should be chosen based on the nature of the trait being studied as well as the function of the gene of interest, when known.

EMERGING APPROACHES FOR FUNCTIONAL EVALUATION

Massively parallel enhancer characterization

The regulation of gene expression is integral to most biological processes. As exemplified by the 1p13 locus, genetic variation within regulatory elements can have significant phenotypic consequences. Advanced DNA synthesis and sequencing technologies have made possible the development of creative solutions to address the challenge of dissecting genetic regulatory sequences, which can be used to interrogate the non-coding regions identified by complex-trait GWASs.

The 'massively parallel reporter assay' (MPRA) approach is used to systematically interrogate each nucleotide position within a transcriptional regulatory element to identify functional TFBSs (11). Transcriptional activity using MPRA is assessed in cultured cells *in vitro* so it can be performed in an appropriate available cell type, accounting for cell-type-specific differential regulation. Briefly, in MPRA, enhancer variants (containing one or more desired nucleotide substitutions) for an element of interest are generated and coupled to distinguishing tags ('barcodes') using microarray-based DNA synthesis. Next, the variants are cloned into a plasmid backbone, an invariant promoter open reading frame segment is inserted, and the reporter plasmid pool is transfected into cells. Lastly, high-throughput RNA sequencing and counting of the distinguishing tags determines the relative quantities of RNA transcripts expressed from the plasmid pool and, by extension, the regulatory activities of the enhancer variants, defining the location of important regulatory positions.

'Massively parallel functional dissection' (MPFD) is a technology related to MPRA; however, it facilitates the

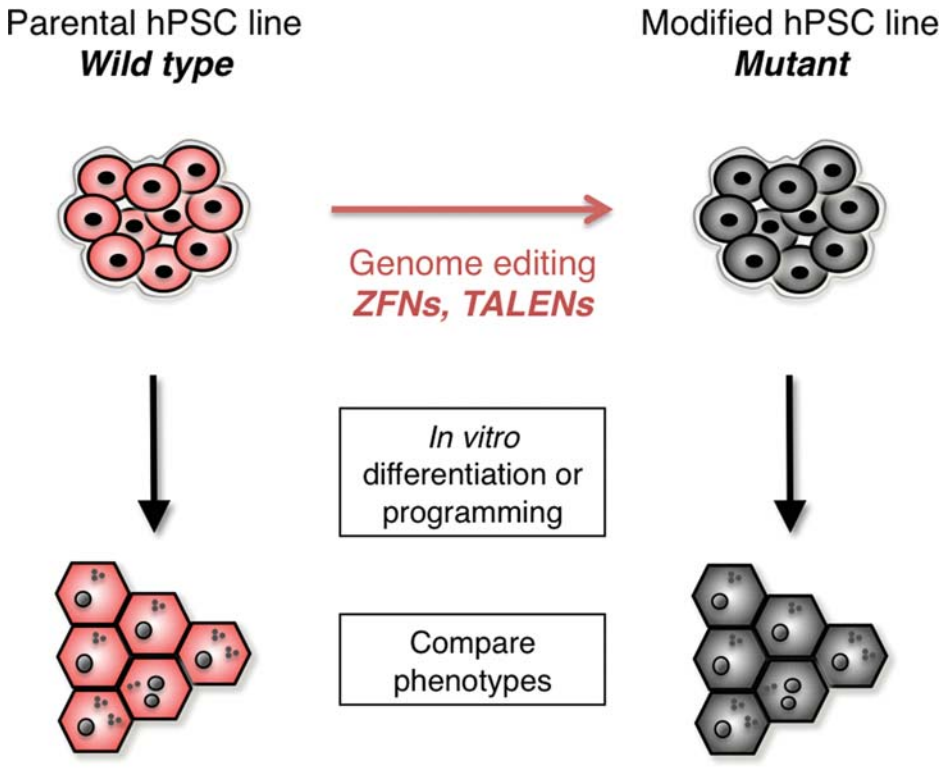

Figure 1. Use of hPSCs to fulfill a genetics version of Koch's postulates.

characterization of enhancer elements *in vivo* and involves the generation of a library of enhancer haplotypes with a programmable level of degeneracy without requiring individual synthesis of each enhancer variant (12). These features make MPFD preferable for the efficient examination of longer regulatory elements in an *in vivo* context. Notably, initial experiments using MPFD have shown that not all functional motifs are associated with predicted TFBSs and that evolutionary constraint (i.e. phylogenetic conservation) poorly predicts the magnitude of functional impact for specific nucleotide positions—two points that should be considered when analyzing regulatory regions in the human genome by computational methods alone.

The MPRA and MPFD technologies could readily be applied to identify causal DNA variants within regions of non-coding DNA identified by GWASs, assuming that the causal DNA variants act by modulating the expression of nearby genes. Identification of a potential causal regulatory variant could then help to nominate candidate causal genes that could be functionally tested, as was the case with the 1p13 locus.

Cellular models of human genetic variation fulfilling the genetics version of Koch's postulates

Despite the insights that were gained using the approaches described above, each system ultimately falls short of establishing a direct causal connection linking genetic variation to its associated complex trait. These studies are limited by several factors that have traditionally made conducting appropriate functional studies challenging. First, complex human diseases and phenotypes may be poorly replicated in animal model systems, and regulatory variants may not be evolutionarily conserved. Second, human clinical samples are not amenable for experimental characterization due to the limited availability of most tissue types. Finally, immortalized human cell lines may in some cases be inappropriate for the evaluation of human genetic discoveries because the impact of natural genetic variation may be obscured by potential genomic instability, karyotypic abnormalities and variability resulting from transformation and extensive passaging in cell culture.

Ongoing research provides clues as to the most appropriate way in which to study the genetic variation underlying complex traits. As observed with the 1p13 locus, genetic contributors to the development of complex traits may act in a tissue-specific manner. Indeed, a majority of common DNA variants that regulate gene expression appear to act in a cell-type-dependent fashion (13), and disease-associated DNA variants are often located in predicted cell-type-specific enhancer elements (14).

The challenges and insights described above suggest that the human pluripotent stem cell (hPSC) technology could be a powerful tool for the evaluation of human genetic variation. One significant advantage is that hPSCs can be genetically modified using genome editing. Engineered nucleases such as zinc finger nucleases or transcription activator-like effector nucleases—chimeric proteins with customized DNA-binding domains that bind to a specified site in the genome and create double-strand breaks—can facilitate precise modification of the genome in hPSCs (15,16). Furthermore, directed differentiation of hPSCs makes it possible to investigate

cellular phenotypes in cell types which are difficult to obtain from humans (e.g. neurons) yet are the most relevant for the traits being studied. While hPSCs have been used to establish cellular models of monogenic disorders (17), little has been published regarding the use of hPSCs to study the genetic nature of complex traits. Nonetheless, one can envision the use of hPSCs to conduct genetic characterization of potential causal DNA variants, to interrogate putative causal genes and to validate genotype–phenotype relationships in appropriate human cell types.

A number of studies have examined the impact of disease-associated genetic variation using human-induced pluripotent stem cells (iPSCs) derived from affected patients and control individuals (18–23). However, this approach has two important limitations: (i) reprogramming may introduce variability that contributes to phenotypic differences when individual iPSC lines are differentiated, particularly when relatively subtle phenotypes are being examined; (ii) the genetic backgrounds of disease-specific and control iPSC lines are not matched, making it difficult to attribute an observed phenotype to the influence of a specific genetic variant. These limitations can be overcome by using human iPSCs or human embryonic stem cells together with genome editing to fulfill the genetics equivalent of Koch's postulates: a causal variant should cause the associated phenotype when introduced into a wild-type cell. One can start with a well-characterized hPSC line; perform genome editing to either introduce or 'repair' a genetic variant, thereby creating an otherwise isogenic cell line that differs from the parental hPSC only with respect to the variant; differentiate the parental (control) and modified hPSC lines in parallel into the cell type of interest; and phenotype the differentiated cell lines (Fig. 1). With this experimental design, any phenotypic differences seen between the cell lines can be directly attributed to the genetic variant.

CONCLUSIONS

It is clear that to truly establish causal connections between genetic variation and complex human traits, specific DNA variants must be directly interrogated in an appropriate biological context. While this type of precise functional genetic characterization is the definitive step in establishing genotype–phenotype relationships, evaluating novel genetic associations often requires several stages of investigation, where a variety of tools and approaches are used to uncover causal DNA variants, genes and molecular mechanisms that contribute to the development of a complex trait. Creative applications of established techniques as well as a host of emerging technological developments promise to make such investigation sufficiently straightforward and efficient that it may well be feasible to cope with the wealth of data that is now emerging from GWASs and next-generation sequencing studies.

ACKNOWLEDGEMENTS

The authors thank Dr Sekar Kathiresan for useful comments on the manuscript.

Conflict of Interest statement. None declared.

FUNDING

The authors are supported by the National Institutes of Health (R00HL098364 to K.M.) and the Sternlicht Director's Fund Award for Graduate Students from the Harvard Stem Cell Institute to D.T.P.

REFERENCES

1. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J. *et al.* (2010) Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, **466**, 707–713.
2. Musunuru, K., Strong, A., Frank-Kamenetsky, M., Lee, N.E., Ahfeldt, T., Sachs, K.V., Li, X., Li, H., Kuperwasser, N., Ruda, V.M. *et al.* (2010) From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature*, **466**, 714–719.
3. Samani, N.J., Erdmann, J., Hall, A.S., Hengstenberg, C., Mangino, M., Mayer, B., Dixon, R.J., Meitinger, T., Braund, P., Wichmann, H.E. *et al.* (2007) Genomewide association analysis of coronary artery disease. *N. Engl. J. Med.*, **357**, 443–453.
4. McPherson, R., Pertsemlidis, A., Kavaslar, N., Stewart, A., Roberts, R., Cox, D.R., Hinds, D.A., Pennacchio, L.A., Tybjaerg-Hansen, A., Folsom, A.R. *et al.* (2007) A common allele on chromosome 9 associated with coronary heart disease. *Science*, **316**, 1488–1491.
5. Helgadottir, A., Thorleifsson, G., Manolescu, A., Gretarsdottir, S., Blondal, T., Jonasdottir, A., Jonasdottir, A., Sigurdsson, A., Baker, A., Palsson, A. *et al.* (2007) A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*, **316**, 1491–1493.
6. Helgadottir, A., Thorleifsson, G., Magnusson, K.P., Grétarsdottir, S., Steinthorsdottir, V., Manolescu, A., Jones, G.T., Rinkel, G.J., Blankensteijn, J.D., Ronkainen, A. *et al.* (2008) The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat. Genet.*, **40**, 217–224.
7. Visel, A., Zhu, Y., May, D., Afzal, V., Gong, E., Attanasio, C., Blow, M.J., Cohen, J.C., Rubin, E.M. and Pennacchio, L.A. (2010) Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature*, **464**, 409–412.
8. Harismendy, O., Notani, D., Song, X., Rahim, N.G., Tanasa, B., Heintzman, N., Ren, B., Fu, X.D., Topol, E.J., Rosenfeld, M.G. and Frazer, K.A. (2011) 9p21 DNA variants associated with coronary artery disease impair interferon-γ signalling response. *Nature*, **470**, 264–268.
9. Romeo, S., Yin, W., Kozlitina, J., Pennacchio, L.A., Boerwinkle, E., Hobbs, H.H. and Cohen, J.C. (2009) Rare loss-of-function mutations in ANGPTL family members contribute to plasma triglyceride levels in humans. *J. Clin. Invest.*, **119**, 70–79.
10. Davis, E.E., Zhang, Q., Liu, Q., Diplas, B.H., Davey, L.M., Hartley, J., Stoetzel, C., Szymanska, K., Ramaswami, G., Logan, C.V. *et al.* (2011) TTC21B contributes both causal and modifying alleles across the ciliopathy spectrum. *Nat. Genet.*, **43**, 189–196.
11. Melnikov, A., Murugan, A., Zhang, X., Tesileanu, T., Wang, L., Rogov, P., Feizi, S., Gnirke, A., Callan, C.G. Jr, Kinney, J.B. *et al.* (2012) Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat. Biotechnol.*, **30**, 271–277.
12. Patwardhan, R.P., Hiatt, J.B., Witten, D.M., Kim, M.J., Smith, R.P., May, D., Lee, C., Andrie, J.M., Lee, S.I., Cooper, G.M. *et al.* (2012) Massively parallel functional dissection of mammalian enhancers in vivo. *Nat. Biotechnol.*, **30**, 265–270.
13. Dimas, A.S., Deutsch, S., Stranger, B.E., Montgomery, S.B., Borel, C., Attar-Cohen, H., Ingle, C., Beazley, C., Gutierrez Arcelus, M., Sekowska, M. *et al.* (2009) Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science*, **325**, 1246–1250.
14. Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M. *et al.* (2011) Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*, **473**, 43–49.
15. Soldner, F., Laganière, J., Cheng, A.W., Hockemeyer, D., Gao, Q., Alagappan, R., Khurana, V., Golbe, L.I., Myers, R.H., Lindquist, S. *et al.* (2011) Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell*, **146**, 318–331.

16. Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C. *et al.* (2011) Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.*, **29**, 731–734.
17. Rashid, S.T., Corbineau, S., Hannan, N., Marciniak, S.J., Miranda, E., Alexander, G., Huang-Doran, I., Griffin, J., Ahrlund-Richter, L., Skepper, J. *et al.* (2010) Modeling inherited metabolic disorders of the liver using human induced pluripotent stem cells. *J. Clin. Invest.*, **120**, 3127–3136.
18. Israel, M.A., Yuan, S.H., Bardy, C., Reyna, S.M., Mu, Y., Herrera, C., Hefferan, M.P., Van Gorp, S., Nazor, K.L., Boscolo, F.S. *et al.* (2012) Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells. *Nature*, **482**, 216–220.
19. Dimos, J.T., Rodolfa, K.T., Niakan, K.K., Weisenthal, L.M., Mitsumoto, H., Chung, W., Croft, G.F., Saphier, G., Leibel, R., Goland, R. *et al.* (2008) Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science*, **321**, 1218–1221.
20. Ebert, A.D., Yu, J., Rose, F.F. Jr, Mattis, V.B., Lorson, C.L., Thomson, J.A. and Svendsen, C.N. (2009) Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature*, **457**, 277–280.
21. Lee, G., Papapetrou, E.P., Kim, H., Chambers, S.M., Tomishima, M.J., Fasano, C.A., Ganat, Y.M., Menon, J., Shimizu, F., Viale, A. *et al.* (2009) Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature*, **461**, 402–406.
22. Itzhaki, I., Maizels, L., Huber, I., Zwi-Dantsis, L., Caspi, O., Winterstern, A., Feldman, O., Gepstein, A., Arbel, G., Hammerman, H. *et al.* (2011) Modelling the long QT syndrome with induced pluripotent stem cells. *Nature*, **471**, 225–229.
23. Brennand, K.J., Simone, A., Jou, J., Gelboin-Burkhart, C., Tran, N., Sangar, S., Li, Y., Mu, Y., Chen, G., Yu, D. *et al.* (2011) Modelling schizophrenia using human induced pluripotent stem cells. *Nature*, **473**, 221–225.