PRION DISEASE

"Big data" gets personal

Matthew S. Lebo,[1,2,3] Sheila Sutti,[4] Robert C. Green[2,3,4,5]*

A woman at risk for a rare fatal prion disease retrained in science with her husband and joined a bioinformatics team to analyze shared data, providing insights about variant penetrance and possible treatments (Minikel *et al.*, this issue).

Rare genetic diseases are tough to research. Efforts are often underfunded and hampered by small sample sizes that yield imprecise associations between genetic mutations and risk estimates. No treatments exist for many of these diseases—including some of the most deadly—so that even early detection seems futile. Yet, while each disease is rare, the additive percentage of individuals in the general population who harbor a rare disease gene variant is quite high. Global sharing of genetic data sources is increasing the number of cases available for analysis, while at the same time, patients and their families have become both advocates and active researchers in unprecedented ways. In this issue of *Science Translational Medicine,* the research findings of Minikel *et al.* (*1*) bear witness to each of these themes. In strictly scientific terms, the new findings better refine variant pathogenicity and penetrance of disease across disease-causing mutations in the *prion protein* (*PRNP*) gene. Of perhaps greater consequence, the story of this research stands as a testament to the generosity of unrequited data sharing from thousands of individuals and scientists around the globe and to the unflinching devotion of a young couple to altering their seemingly unalterable genetic destiny.

DETERMINING THE PENETRANCE OF SPECIFIC VARIANTS

For clinicians and clinical laboratories, establishing genetic-variant pathogenicity in uncommon diseases can be challenging because of the rarity of affected families and the even greater rarity of families with plausible pathogenic variants. This conundrum is compounded further in the Mendelian space by incomplete or unknown penetrance, that is, the probability that a carrier of a purported disease-causing genotype will indeed develop the disease (*2*). The inability to accurately determine the pathogenicity and penetrance of a particular variant confuses and frustrates efforts to counsel patients, even for genes that are well known and definitively associated with a disease. For instance, truncating variants in the A-band of the *titin* (*TTN*) gene are known to cause dilated cardiomyopathy (DCM) in some families but occur in the general population at a much higher frequency than the prevalence of DCM itself (*3*).

Taking advantage of large data collections of both affected and unaffected individuals, Minikel and colleagues (*1*) have used the frequencies of specific variants among families with prion disease to estimate variant-level penetrance. To this end, they interrogated more than 16,000 affected individuals and more than 60,000 presumably unaffected controls, then replicated some of their findings in an even larger cohort of >500,000 genotyped individuals.

The authors began with 63 variants in the *PRNP* gene that had previously been published as pathogenic for prion disease and compared the observed frequencies of these variants in presumed controls with what would be expected on the basis of the incidence of disease and genetic etiology. They confirmed that four variants are clearly pathogenic with approximately 100% penetrance, whereas another three, previously thought to be pathogenic with high penetrance, were likely benign. Three additional variants were identified as neither benign nor fully penetrant, but could now be assigned more quantitative estimates of their penetrance, ranging from 0.1% to 10%.

Focusing on genetic variants that result in a truncated prion protein, the authors refined our understanding of the likely disease mechanism for *PRNP* pathogenic variants by identifying patterns in the location of these variants within the gene. Keeping in mind that *PRNP* has a single exon (no RNA splicing is required) such that a truncated protein will persist without nonsense-mediated decay, the data revealed that truncating genetic variants found in presumably healthy individuals were limited to the 5′ portion of the gene, whereas variants in affected individuals (likely with higher penetrance) were clustered within the 3′ portion. This finding suggests that an abnormal protein must be long enough to produce a dominant gain-of-function effect that leads to disease; and that nonproduction of the prion protein (PrP) as seen with 5′ mutations is not, in and of itself, disease producing or lethal. This discovery has important treatment implications, because it suggests that a therapy tuned to reducing expression of PrP could be safely tolerated by an individual receiving the therapy. For example, using gene editing to disable the entire gene might prevent the disease without harming the patient.

AN UNUSUAL RESPONSE TO A RARE DISEASE

Eric Minikel did not stumble on prion disease research by looking for an interesting postgraduate project. He and his wife, coauthor Sonia Vallabh, are engaged in a very personal and existential battle against this disease. In 2010, Sonia's mother passed away in her 50s after visual disturbances progressed to weakness, dementia, and death. An autopsy revealed the prion disease fatal familial insomnia, and genetic testing determined that Sonia had inherited a well-known pathogenic *PRNP* mutation (p.D178N) for this dominant disorder from her mother, giving Sonia an unambiguous death sentence in her fourth or fifth decade. Whereas many at-risk individuals might have taken more conventional paths of denial, resignation, or hope leavened with spirituality or research advocacy, Sonia and Eric took a more active role. Within three years, they had left their respective positions as an engineer and a lawyer and both entered into scientific Ph.D. programs at Harvard Medical School. They have already contributed to better characterizing of the onset of prion diseases (*4*) and are deeply engaged in international collaborations and patient advocacy groups. In the study published in this issue (*1*), Sonia and Eric discovered cruel confirmation that the *PRNP* variant that Sonia carries is nearly 100% penetrant.

But for at least one other family, the burden of carrying a pathogenic variant has already been radically altered in a positive way by the new findings. Two of us (R.C.G. and S.S.) have been following an individual in our genetics clinic whose mother died from prion disease

[1]Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115, USA. [2]Harvard Medical School, Boston, MA 02115, USA. [3]Partners HealthCare Personalized Medicine, Boston, MA 02115, USA. [4]Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA. [5]Broad Institute of MIT and Harvard, Boston, MA 02115, USA.
*Corresponding author. E-mail: rcgreen@genetics.med.harvard.edu
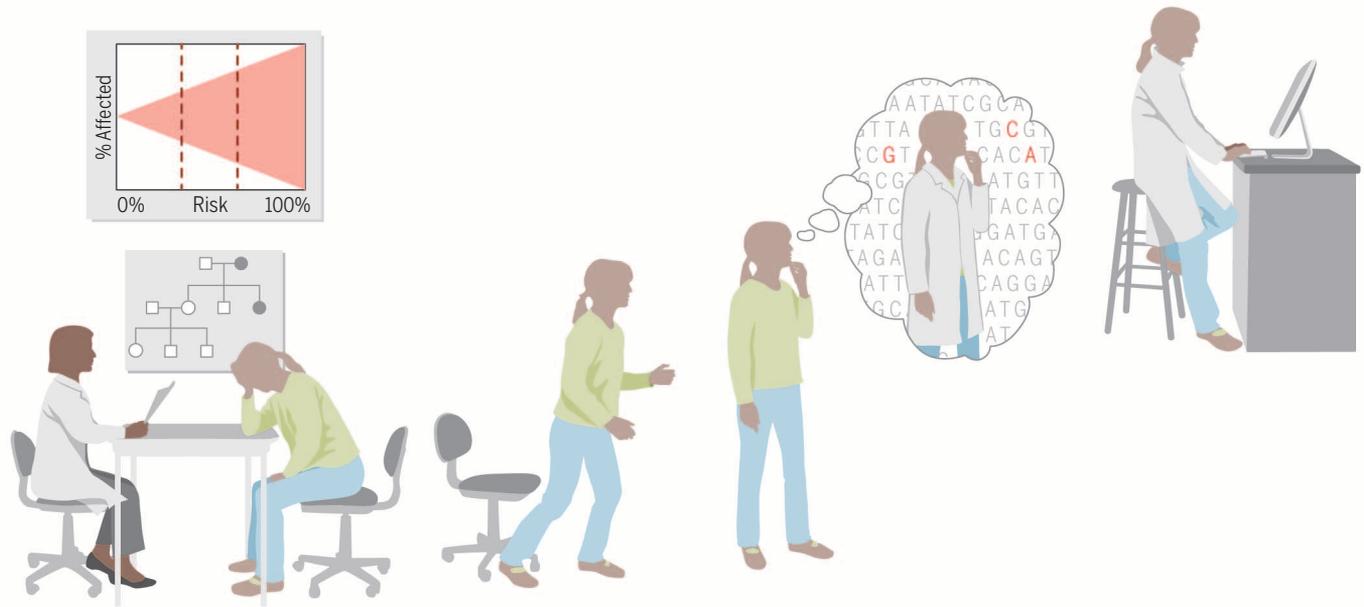
Patient empowerment and unconventional activism. A physician explains the concept of penetrance to a patient and informs her that she has a pathogenic mutation that will result in a rare, incurable and fatal genetic disorder. The patient ponders her fate and, in a rare form of patient activism, becomes a medical-genomics researcher dedicated to seeking new treatments for her disease.

attributed to the p.E196A variant that was previously presumed to be pathogenic but, through this new work, has now been shown likely to be benign. Upon acceptance of the Minikel *et al.* paper by *Science Translational Medicine* (*1*), we placed one of the most exhilarating phone calls a clinician can make—reassuring an untested adult that her mother's mutation was likely benign. Although the tragic demise of our patient's mother remains raw and, even in some sense, less explicable than before, the specter of inheriting a fatal mutation has been mitigated for her, for her children, and for their children to come.

MASSIVELY SCALABLE DATA MEET ARTISANAL ANNOTATION

The collection and sharing of case-level data through the goodwill of more than 16,000 families affected by prion diseases and the generosity of the clinicians and scientists who diagnosed and submitted the data were crucial to the success of the Minikel *et al.* study (*1*). So, too, was the startling power of the ExAC database (http://biorxiv.org/content/early/2015/10/30/030338) created through an international consortium by rigorously collating ancestry-specific variant frequencies in more than 60,000 individuals, from which extremely rare variant frequencies could, for the first time, be estimated. For this analysis, where consent permitted, contributors of individual

cohorts to ExAC were able to reach back into individual case histories and provide enough phenotypic information to determine the case status or family history of individuals with exceedingly rare *PRNP* variants, thus adding to the power of the control data. As demonstrated by the multinational authors and cohorts, the power for such in-depth studies can only be amassed by pooling cases across continents. The rise of federated sharing will enhance the utility of such collections; for example, the Global Alliance for Genomics and Health has made the linkage and sharing of case-level data one of its initial objectives through its newly formed Matchmaker Exchange (www.matchmakerexchange.org).

The Minikel *et al.* paper also highlights the precarious state of variant classification for most rare variants and diseases and the limitations of existing literature (*5*). Clinical laboratories must parse previously published reports and try to detect misclassifications in existing databases as they struggle to classify incompletely penetrant variants, especially when estimates of penetrance begin to resemble odds ratios (below 3 or 2) typically associated with more common complex-risk alleles. Even when databases are accurate about variant-disease associations, existing estimates of penetrance have been determined almost exclusively among affected families and are likely much lower for suspected variants in families without

a robust family history. The resulting classification paradigms make the best predictions that they can, but often the hours of literature analysis translate into only the vague and frustrating category of "variant of uncertain significance," or VUS. Our own work has shown that, even for hypertrophic cardiomyopathy, one of the most commonly tested for dominant genetic conditions, 42% of variants identified in patients are VUSs, which represents almost 20% of cases (*6*). This highlights the need for large case/control populations and for unbiased phenotyping and family histories of general-population cohorts such as that envisioned by President Obama's Precision Medicine Initiative (www.whitehouse.gov/precision-medicine). The increased use of automation, combined with large-scale data, will enable laboratorians and geneticists to better determine the pathogenicity of variants. The findings of Minikel *et al.* (*1*) emphasize the complex nature of low-penetrance alleles, which, because of their relatively high frequency in the general population, may not be classified as pathogenic without strong additional evidence.

It is important to note that while the work of Minikel and colleagues demonstrates the power of statistical approaches and large datasets, the authors were able to make strong assertions about the pathogenicity and penetrance of only 10 of the 63 variants that they studied. Although these variants are some of

the most common in their cohort, in ~15% of genotype-positive individuals the prognosis remains unclear. Clinical laboratories currently identify VUSs twice as often as pathogenic variants, and the vast majority of variants are observed only once or a handful of times in clinical laboratories. Therefore, much more data [such as those housed in ClinGen and other resources (www.clinicalgenome.org)] are needed to make approaches such as the one described herein more broadly applicable.

ACTIONABILITY AND PATIENT EMPOWERMENT

More than just teasing out *PRNP* variant pathogenicity and demonstrating an important framework for rare diseases, this work speaks to the confusing concept of actionability in genomic testing. In the strictest sense, Sonia's test results would never have been considered medically actionable, as prion disease is rapidly progressive, universally fatal, and completely untreatable at this time. By many published standards and expert perspectives, Sonia should have been heavily counseled, if not dissuaded, about the potential psychological damage of learning whether or not she carried this mutation, with the underlying assumption that no medical good could result from such knowledge. But the notion of actionability has proven to be far less parochial than current conventions would suggest. Accumulating evidence from disclosure of genetic risk for a single condition (*7*), for panels of genetic risk markers (*8*, *9*), and for potentially devastating unanticipated findings (*10*) suggests that when individuals wish to know their genetic risks, they largely manage the information in highly constructive ways. The power of sharing such knowledge with patients might, in some cases, be exaggerated by genetic testing companies in the emerging bioindustrial complex, but patient empowerment also appears to be a tangible, quantifiable construct that can stir fundamental personal changes and breakthrough science.

Eric Minikel and Sonia Vallabh have redefined the notion of actionability for prion disease in a way that could lead to treatment and prevention. It seems both naïve and patronizing to suggest that anyone who is properly informed should not be permitted access, if they wish, to any genetic information about themselves, regardless of whether treatments are currently available. Eric, Sonia, and their colleagues have given us a shining example of work that redefines actionability and challenges genetic destiny—offering hope not only for themselves but to everyone else who bears this burden.

REFERENCES AND NOTES

1. E. Vallabh Minikel, S. Minikel Vallabh, M. Lek, K. Estrada, K. E. Samocha, J. F. Sathirapongsasuti, C. Y. McLean, J. Y. Tung, L. P. C. Yu, P. Gambetti, J. Blevins, S. Zhang, Y. Cohen, W. Chen, M. Yamada, T. Hamaguchi, N. Sanjo, H. Mizusawa, Y. Nakamura, T. Kitamoto, S. J. Collins, A. Boyd, R. G. Will, R. Knight, C. Ponto, I. Zerr, T. F. J. Kraus, S. Eigenbrod, M. Calero, J. de Pedro-Cuesta, S. Haïk, J.-L. Laplanche, E. Bouaziz-Amar, J.-P. Brandel, S. Capellari, P. Parchi, A. Poleggi, A. Ladogana, A. H. O'Donnell-Luria, K. J. Karczewski, J. L. Marshall, M. Boehnke, M. Laakso, K. L. Mohlke, A. Kähler, K. Chambert, S. McCarroll, P. F. Sullivan, C. M. Hultman, S. M. Purcell, P. Sklar, S. J. van der Lee, A. Rozemuller, C. Jansen, A. Hofman, R. Kraaij, J. G. J. van Rooij, M. A. Ikram, A. G. Uitterlinden, C. M. van Duijn, Exome Aggregation Consortium (ExAC), M. J. Daly, D. G. MacArthur, Quantifying prion disease penetrance using large population control cohorts. *Sci. Transl. Med.* **8**, 322ra9 (2016).
2. D. N. Cooper, M. Krawczak, C. Polychronakos, C. Tyler-Smith, H. Kehrer-Sawatzki, Where genotype is not predictive of phenotype: Towards an understanding of the molecular basis of reduced penetrance in human inherited disease. *Hum. Genet.* **132**, 1077–1130 (2013).
3. A. M. Roberts, J. S. Ware, D. S. Herman, S. Schafer, J. Baksi, A. G. Bick, R. J. Buchan, R. Walsh, S. John, S. Wilkinson, F. Mazzarotto, L. E. Felkin, S. Gong, J. A. MacArthur, F. Cunningham, J. Flannick, S. B. Gabriel, D. M. Altshuler, P. S. Macdonald, M. Heinig, A. M. Keogh, C. S. Hayward, N. R. Banner, D. J. Pennell, D. P. O'Regan, T. R. San, A. de Marvao, T. J. Dawes, A. Gulati, E. J. Birks, M. H. Yacoub, M. Radke, M. Gotthardt, J. G. Wilson, C. J. O'Donnell, S. K. Prasad, P. J. Barton, D. Fatkin, N. Hubner, J. G. Seidman, C. E. Seidman, S. A. Cook, Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease. *Sci. Transl. Med.* **7**, 270ra6 (2015).
4. E. V. Minikel, I. Zerr, S. J. Collins, C. Ponto, A. Boyd, G. Klug, A. Karch, J. Kenny, J. Collinge, L. T. Takada, S. Forner, J. C. Fong, S. Mead, M. D. Geschwind, Ascertainment bias causes false signal of anticipation in genetic prion disease. *Am. J. Hum. Genet.* **95**, 371–382 (2014).
5. H. M. McLaughlin, O. Ceyhan-Birsoy, K. D. Christensen, I. S. Kohane, J. Krier, W. J. Lane, D. Lautenbach, M. S. Lebo, K. Machini, C. A. MacRae, D. R. Azzariti, M. F. Murray, C. E. Seidman, J. L. Vassy, R. C. Green, H. L. Rehm; for The MedSeq Project, A systematic approach to the reporting of medically relevant findings from whole genome sequencing. *BMC Med. Genet.* **15**, (2014).
6. A. A. Alfares, M. A. Kelly, G. McDermott, B. H. Funke, M. S. Lebo, S. B. Baxter, J. Shen, H. M. McLaughlin, E. H. Clark, L. J. Babb, S. W. Cox, S. R. DePalma, C. Y. Ho, J. G. Seidman, C. E. Seidman, H. L. Rehm, Results of clinical genetic testing of 2,912 probands with hypertrophic cardiomyopathy: Expanded panels offer limited additional sensitivity. *Genet. Med.* **17**, 880–888 (2015).
7. R. C. Green, J. S. Roberts, L. A. Cupples, N. R. Relkin, P. J. Whitehouse, T. Brown, S. L. Eckert, M. Butson, A. D. Sadovnick, K. A. Quaid, C. Chen, R. Cook-Deegan, L. A. Farrer; R. S. Group; REVEAL Study Group, Disclosure of APOE genotype for risk of Alzheimer's disease. *N. Engl. J. Med.* **361**, 245–254 (2009).
8. D. A. Carere, P. Kraft, K. A. Kaphingst, J. S. Roberts, R. C. Green; for the PGen Study Group, Consumers report lower confidence in their genetics knowledge following direct-to-consumer personal genomic testing. *Genet. Med.* **26**, (2015).
9. D. A. Carere, M. P. Couper, S. D. Crawford, S. S. Kalia, J. R. Duggan, T. A. Moreno, J. L. Mountain, J. S. Roberts, R. C. Green; PGen Study Group, Design, methods, and participant characteristics of the Impact of Personal Genomics (PGen) Study, a prospective cohort study of direct-to-consumer personal genomic testing customers. *Genome Med.* **6**, 96 (2014).
10. J. L. Vassy, D. M. Lautenbach, H. M. McLaughlin, S.-W. Kong, K. D. Christensen, J. Krier, I. S. Kohane, L. Z. Feuerman, J. Blumenthal-Barby, J. S. Roberts, L. S. Lehmann, C. Y. Ho, P. A. Ubel, C. A. MacRae, C. E. Seidman, M. F. Murray, A. L. McGuire, H. L. Rehm, R. C. Green; MedSeq Project, The MedSeq Project: A randomized trial of integrating whole genome sequencing into clinical medicine. *Trials* **15**, 85–97 (2014).

Citation: M. S. Lebo, S. Sutti, R. C. Green, "Big data" gets personal. *Sci. Transl. Med.* **8**, 322fs3 (2016).

Science Translational Medicine

"Big data" gets personal

Matthew S. Lebo, Sheila Sutti and Robert C. Green

ARTICLE TOOLS	http://stm.sciencemag.org/content/8/322/322fs3
RELATED CONTENT	http://stm.sciencemag.org/content/scitransmed/8/322/322ra9.full http://stm.sciencemag.org/content/scitransmed/7/291/291fs25.full http://stm.sciencemag.org/content/scitransmed/7/290/290ps13.full http://stm.sciencemag.org/content/scitransmed/7/311/311ra174.full http://stm.sciencemag.org/content/scitransmed/7/314/314fs47.full http://stm.sciencemag.org/content/scitransmed/7/299/299ra123.full http://stm.sciencemag.org/content/scitransmed/7/305/305ps20.full http://stm.sciencemag.org/content/scitransmed/4/115/115rv1.full http://stm.sciencemag.org/content/scitransmed/5/202/202ed15.full http://stm.sciencemag.org/content/scitransmed/8/322/322ps2.full http://stm.sciencemag.org/content/scitransmed/8/322/322pc1.full http://science.sciencemag.org/content/sci/356/6333/12.full
REFERENCES	This article cites 9 articles, 2 of which you can access for free http://stm.sciencemag.org/content/8/322/322fs3#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions