Approaches to Tag a Gene or Variation Need Updates

Yijia Zhou

January 9, 2017

Abstract

In recent years, researchers searched for essential genes or pathogenic variations by simply scanning for mutations that are more common in people with disease than in those without or knocking out to see weather cells survive or not. But since some human genetic variations are relatively rare, this method can lead to false positive results. And without concerning adaptive evolution, some "essential genes" in fact are something cells can throw away.

1 Introduction

Working out which mutations are actually linked to illness will be a long and arduous task. Early efforts to discover the genetic underpinnings of disease worked on families with a recurring particular disease, generation after generation. But in recent years, researchers have switched tactics: for instance, searching for evidence of pathogenicity by scanning for mutations that are more common in people with disease than in those without. It is becoming clear that many human genetic variations are relatively rare, and when researchers do not examine large enough groups of people with and without disease when scanning for pathogenic mutations, they are likely to mistakenly conclude that particular variants of interest turn up only in people with disease. The truth may be that they just haven't looked hard enough for these variants elsewhere.[1]

However, the Exome Aggregation Consortium (ExAC) and collaborators report the exome sequences of 60,706 individuals, collected from diverse studies: a venture 5,000 times larger than our initial study.[4] With these huge datasets, we now can operate more reliable inspection about the variations we concern.

A history of systematic gene deletion studies, conducted primarily in prokaryotes, has identified biological features that distinguish essential from non-essential genes. But new researches suggests that essentiality is not an intrinsic property of a gene but is influenced by genetic and environmental factors[5].

2 Validate existing pathogenic variation researches

2.1 The deadly mutations that weren't

In August this year, MacArthur's group published its analysis of ExAC data in Nature[4], revealing that many mutations thought to be harmful are probably not. In one analysis, the group identified 192 variants that had previously been thought to be pathogenic, but turned out to be relatively common. The scientists reviewed papers about these variants, looking for plausible evidence that they actually caused disease, but could find solid evidence for only nine of them. Most are actually benign, according to standards set by the American College of Medical Genetics and Genomics, and many have now been reclassified as such.1

2.2 Along with medical research and practice

Similar work promises to have direct impacts on medical practice. In a companion paper[8], geneticist Hugh Watkins of the University of Oxford, UK, looked at genes associated with certain types of cardiomyopathy that cause gradual weakening of the heart muscle, which can lead to sudden death undetected, and it has become fairly common to check relatives of people with the conditions for genetic mutations associated with them. Those found to have a genetic risk are sometimes counselled to get an implanted defibrillator, which delivers electrical shocks to the

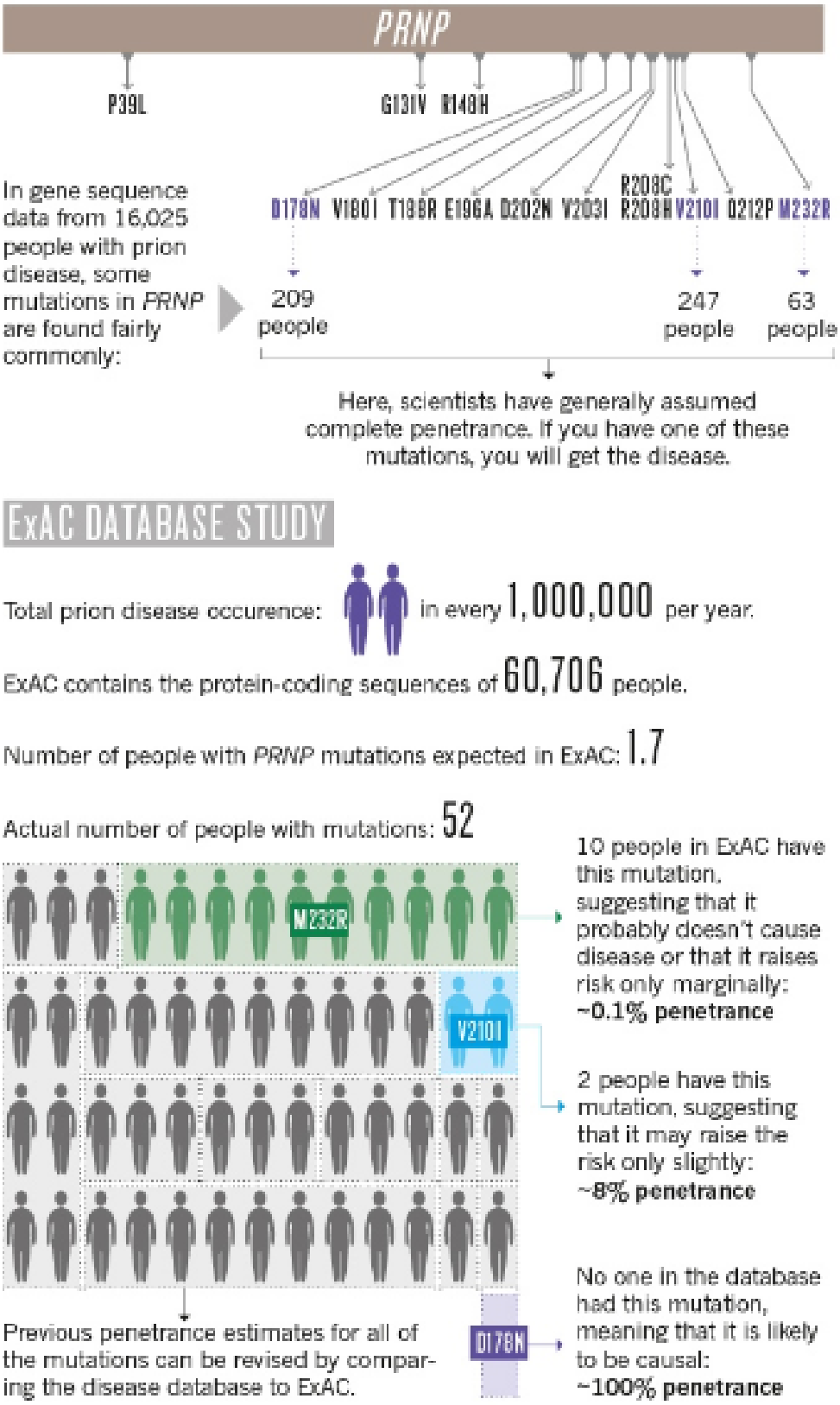

Figure 1: The deadly mutations that weren't.
Prion diseases are rare neurodegenerative disorders caused by misfolded prion proteins. About 63 mutations in the gene PRNP have been linked to them. But until now it has been difficult to estimate how likely it is that a given variant will result in disease, a measure known as penetrance. Data compoled by the Exome Aggregation Consortium (ExAC) can help.[2]

heart if it seems to be beating abnormally. Watkins checked the ExAC database for information on genes that have been associated with these heart conditions, and found that many mutations are much too common among healthy people to be pathogenic. About 60 genes had been implicated as harbouring pathogenic mutations that cause one form of the disease; Watkins' analysis revealed that 40 of these probably bear no link.

3 Approaches for ongoing valid researches

3.1 Apply more rigorous tests to future research

ExAC is quietly becoming a standard tool in medical genetics. Clinical labs around the world now check it before telling a patient that a particular glitch in their genome might be making them ill. If the mutation is common in ExAC, it's unlikely to be harmful.

3.2 Discover genes involved in rare diseases

The way how exome sequencing could be used to identify Mendelian-disease genes or to diagnose Mendelian disease has been described before[3]. Because there are tens of thousands of genetic variants in an exome, these strategies depended on effectively filtering out common variants, which are not likely to cause Mendelian disorders. At that time, databases of common variants were uneven and of suspect quality. Although Exome Sequencing Project (ESP) greatly improved the situation by uniformly and systematically cataloguing both common and rare variants across the exome, ExAC is an order of magnitude larger, and so enables better filtering. This is especially relevant for exome sequencing of non-European, non-African-American individuals, because ExAC provides greater sampling of individuals from outside the United States than ESP does[7].

3.3 Human knockouts

ExAC is revealing a lot about genes through the frequency of mutations. MacArthur and his team found 3,200 genes that are almost never severely mutated in any of the ExAC genomes — a signal that these genes are important. And yet 72% of them have never before been linked to disease[4]. Researchers are eager to study whether some of these genes play unappreciated parts in illness.

Conversely, the group has found nearly 180,000 instances of mutations so severe that they should render their protein products completely inactive. Scientists have long studied genes by knocking them out in animals such as mice, so that they don't work. By looking at the symptoms that develop, they can study what the genes do. But that has never been possible in humans. Now, researchers are eager to study these natural human knockouts to understand what they can reveal about how diseases develop or may be cured. MacArthur and other researchers are gearing up to prioritize which human knockout genes to study and how best to contact the people carrying them for further study.

4 Evolvable and Non-evolvable Essential Genes

Our current definition of essential genes is based on whether or not gene knockout cells survive. Researchers now reveal a new class of essential genes — called 'evolvable essential genes' — that can overcome loss of function with time by evolving alternative, usually unrelated, cellular processes driven by aneuploidy[6].

Thus, a quantitative redefinition of gene essentiality that incorporates both viability and evolvability is required. Distinguishing non-evolvable from evolvable essential genes should be considered when ranking potential drug targets to minimize drug resistance.

References

[1] Rethink the links between genes and disease. *Nature*, 538(7624):140–140, oct 2016.

[2] Erika Check Hayden. A radical revision of human genetics. *Nature*, 538(7624):154–157, oct 2016.

[3] M. Choi, U. I. Scholl, W. Ji, T. Liu, I. R. Tikhonova, P. Zumbo, A. Nayir, A. Bakkaloglu, S. Ozen, S. Sanjad, C. Nelson-Williams, A. Farhi, S. Mane, and R. P. Lifton. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences*, 106(45):19096–19101, nov 2009.

[4] Monkol Lek, Konrad J. Karczewski, Eric V. Minikel, Kaitlin E. Samocha, Eric Banks, Timothy Fennell, Anne H. O'Donnell-Luria, James S. Ware, Andrew J. Hill, Beryl B. Cummings, Taru Tukiainen, Daniel P. Birnbaum, Jack A. Kosmicki, Laramie E. Duncan, Karol Estrada, Fengmei Zhao, James Zou, Emma Pierce-Hoffman, Joanne Berghout, David N. Cooper, Nicole Deflaux, Mark DePristo, Ron Do, Jason Flannick, Menachem Fromer, Laura Gauthier, Jackie Goldstein, Namrata Gupta, Daniel Howrigan, Adam Kiezun, Mitja I. Kurki, Ami Levy Moonshine, Pradeep Natarajan, Lorena Orozco, Gina M. Peloso, Ryan Poplin, Manuel A. Rivas, Valentin Ruano-Rubio, Samuel A. Rose, Douglas M. Ruderfer, Khalid Shakir, Peter D. Stenson, Christine Stevens, Brett P. Thomas, Grace Tiao, Maria T. Tusie-Luna, Ben Weisburd, Hong-Hee Won, Dongmei Yu, David M. Altshuler, Diego Ardissino, Michael Boehnke, John Danesh, Stacey Donnelly, Roberto Elosua, Jose C. Florez, Stacey B. Gabriel, Gad Getz, Stephen J. Glatt, Christina M. Hultman, Sekar Kathiresan, Markku Laakso, Steven McCarroll, Mark I. McCarthy, Dermot McGovern, Ruth McPherson, Benjamin M. Neale, Aarno Palotie, Shaun M. Purcell, Danish Saleheen, Jeremiah M. Scharf, Pamela Sklar, Patrick F. Sullivan, Jaakko Tuomilehto, Ming T. Tsuang, Hugh C. Watkins, James G. Wilson, Mark J. Daly, Daniel G. MacArthur, and Exome Aggregation Consortium. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, 536(7616):285–291, aug 2016.

[5] Liesbet Lieben. Redefining gene essentiality. *Nature Publishing Group*, 17(December):2015, dec 2015.

[6] Gaowen Liu, Mei Yun Jacy Yong, Marina Yurieva, Kandhadayar Gopalan Srinivasan, Jaron Liu, John Soon Yew Lim, Michael Poidinger, Graham Daniel Wright, Francesca Zolezzi, Hyungwon Choi, Norman Pavelka, and Giulia Rancati. Gene Essentiality Is a Quantitative Property Linked to Cellular Evolvability. *Cell*, 163(6):1388–1399, dec 2015.

[7] Jay Shendure. Human genomics: A deep dive into genetic variation. *Nature*, 536(7616):277–278, aug 2016.

[8] Roddy Walsh, Kate L. Thomson, James S. Ware, Birgit H. Funke, Jessica Woodley, Karen J. McGuire, Francesco Mazzarotto, Edward Blair, Anneke Seller, Jenny C. Taylor, Eric V. Minikel, Exome Aggregation Exome Aggregation Consortium, Daniel G. MacArthur, Martin Farrall, Stuart A. Cook, and Hugh Watkins. Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and 60,706 reference samples. *Genetics in Medicine*, aug 2016.