Phoebe Dace and Gregory M. Findlay*

Reducing uncertainty in genetic testing with Saturation Genome Editing

Abstract: Accurate interpretation of human genetic data is critical for optimizing outcomes in the era of genomic medicine. Powerful methods for testing genetic variants for functional effects are allowing researchers to characterize thousands of variants across disease genes. Here, we review experimental tools enabling highly scalable assays of variants, focusing specifically on Saturation Genome Editing (SGE). We discuss examples of how this technique is being implemented for variant testing at scale and describe how SGE data for *BRCA1* have been clinically validated and used to aid variant interpretation. The initial success at predicting variant pathogenicity with SGE has spurred efforts to expand this and related techniques to many more genes.

Keywords: variants of uncertain significance, functional assays, genotype–phenotype, CRISPR, Saturation Genome Editing

Introduction

Technologies to cost-effectively ascertain human genetic variation have greatly advanced our understanding of genotype–phenotype associations and the molecular basis of human disease. However, for each gene in which variants have been associated with disease it often remains unknown which of *all possible* variants in and near the gene will, indeed, cause disease.

This is evidenced clinically by large numbers of "variants of uncertain significance" (VUS) observed in commonly sequenced genes (Fig. 1) [1]. Considering nearly all single nucleotide variants (SNVs) compatible with life are predicted to exist in at least one living individual [2], the true number of variants of unknown disease association is far greater than the total number of VUS observed. Furthermore, VUS rates are particularly high in non-European populations, who are poorly represented in many of the largest genetic databases [3].

*Corresponding author: Gregory M. Findlay, The Genome Function Laboratory, The Francis Crick Institute, 1 Midland Rd, London, United Kingdom, e-mail: greg.findlay@crick.ac.uk
Phoebe Dace, The Genome Function Laboratory, The Francis Crick Institute, 1 Midland Rd, London, United Kingdom

For many diseases, specific management strategies may improve outcomes once a definitive genetic diagnosis is reached. In the realm of cancer predisposition syndromes, for instance, increased screening and prophylactic procedures can substantially increase overall survival for carriers of pathogenic variants [4]. Additionally, new classes of targeted therapeutics have proven effective for treating cancers of specific genetic etiologies [5–7]. Therefore, to maximize the value of genetic testing, it is imperative to be able to accurately predict the molecular and phenotypic consequences of any given variant.

This is, however, a monumental challenge. With over 6 million variants observed per whole genome sequenced [8], accurately identifying which variants contribute to each of thousands of monogenic disease phenotypes is hard enough for single patients, let alone whole populations. The magnitude of the challenge posed by rare variant interpretation necessitates leveraging orthogonal strategies. Sampling DNA from more individuals will continue to gradually increase the statistical power of genetic associations. However, for the foreseeable future classical genetics approaches will continue to prove inadequate for enumerating *all* variants that cause disease phenotypes. After all, most variants with phenotypic effects driving monogenic disease are exceedingly rare [9].

When rare variants are encountered in disease genes, computational models to predict their effects are routinely used to aid interpretation (e. g., CADD [10], REVEL [11]). These models, increasingly derived using machine learning [12–14], typically leverage combinations of sequence conservation and biochemical data to predict deleteriousness. This makes them unmatched in regards to scalability, as virtually every variant in the genome can be assigned a pre-computed score. However, their predictions are typically not accurate enough to be used without additional evidence [15].

MAVEs test thousands of genetic variants per experiment

To better understand sequence–function relationships, many research teams have contributed to a suite of experimental methods now collectively known as multiplex as-
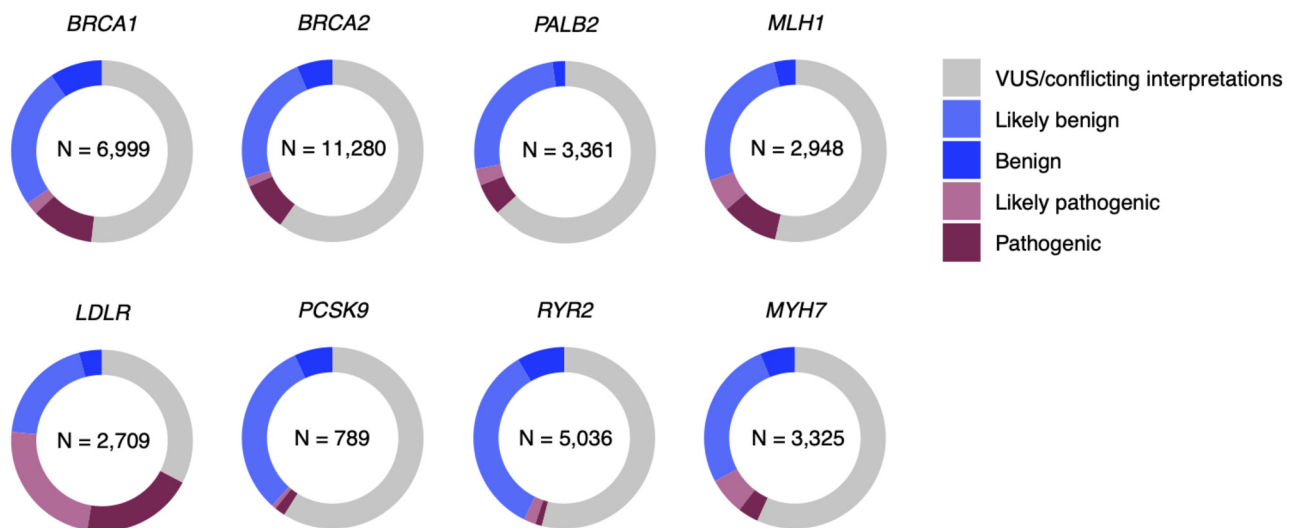
Figure 1: ClinVar variants in commonly sequenced genes. The number of variants reported as of September 2022 is shown for each of eight genes for which secondary findings are returned [37]. Variants of uncertain significance (VUS) or with conflicting interpretations are shown in gray. For most genes, VUS outnumber variants deemed pathogenic or benign.

says of variant effect (MAVEs) [16]. MAVEs allow thousands of different variants to be assessed in a pooled format to reveal each variant's functional consequence. The highly scalable nature of MAVEs relieves the bottleneck of testing variants individually. Yet, developing MAVEs that faithfully reflect variant effects on disease processes is challenging, as is translating experimental results into actionable clinical evidence.

Next-generation sequencing (NGS) platforms have enabled exponential growth in human genetic data by drastically reducing the cost of sequencing DNA. Sequencing technology is now used routinely to quantify the relative abundances of different DNA sequences present in complex samples. Powerful methods based on quantification via NGS include single-cell RNA sequencing to measure gene expression in individual cells and CRISPR knockout screens to ask which genes drive cellular phenotypes. A common feature of MAVEs is that they similarly leverage the scalability afforded by NGS for counting DNA molecules, as precise quantification of variants present in experimentally manipulated samples can enable accurate inference of functional effects.

In a typical MAVE experiment, hundreds to hundreds of thousands of variants are generated in a pooled fashion using a strategy such as custom oligonucleotide synthesis or error-prone PCR [17, 18]. This "variant library" is cloned into an expression vector designed such that each variant may influence a specific molecular or cellular phenotype. NGS is used to "read out" the experiment, i.e., to quantify variants' effects on phenotype by measuring dif-

ferences in read counts between variants tested. Deep mutational scanning (DMS) is one type of MAVE used to ask how missense variants affect, for instance, protein stability or enzymatic activity [19]. Likewise, massively parallel reporter assays (MPRAs) reveal how variants in regulatory sequences such as promoters or enhancers alter transcript levels [20, 21]. The key to a successful MAVE is being able to quantify each specific variant's effect with NGS, such that all variants can be tested in a pooled format, thereby drastically reducing cost and processing time.

For a MAVE to aid clinical variant interpretation, the assay must faithfully distinguish variants associated with disease. This can be challenging because pathogenic variants within a single gene often exert effects through diverse molecular mechanisms. Nonetheless, many recent MAVEs have predicted variant pathogenicity with high accuracy [22–25], suggesting MAVE data will be valuable for classifying new variants.

Developing Saturation Genome Editing to test variants at the endogenous locus

One key limitation of many functional assays is that variants are tested outside their endogenous genomic context. For instance, missense variants are typically expressed in cDNA vectors lacking introns and endogenous regulatory elements. This compaction of sequence space can

increase assay throughput but may confound interpretation because pathogenic variants can affect any number of molecular processes involved in gene function (e. g., transcription, RNA splicing, protein folding, protein function). Different variants affecting the same gene may require different assays depending on their mechanism of action, for instance, one assay for splicing and another for enzymatic activity.

To overcome these limitations, we developed Saturation Genome Editing (SGE) as a way to systematically test variants in the context of the human genome [26]. In a typical SGE experiment, all possible SNVs in a short region (up to 150 bp) are tested in a single pool, along with any other variants of interest, such as in-frame insertions and deletions (Fig. 2). Variants are cost-effectively synthe-

sized in oligonucleotide pools, meaning any short variant (<50 bp) can be engineered. The variant library is amplified, cloned into "donor" plasmids to facilitate homology-directed DNA repair (HDR), and introduced to cells using CRISPR gene editing [27]. Each successfully edited cell receives a single variant from the library. By genetically modifying millions of cells at a time, each variant in the library is created many times independently, minimizing potential effects of off-target CRISPR editing.

Once variants are introduced to cells, a functional selection is applied to discriminate variant effects. We initially leveraged the fact that across human cell lines a subset of expressed genes are required for proliferation [28]. We showed that loss-of-function variants in these "essential" genes can be distinguished from functionally neu-
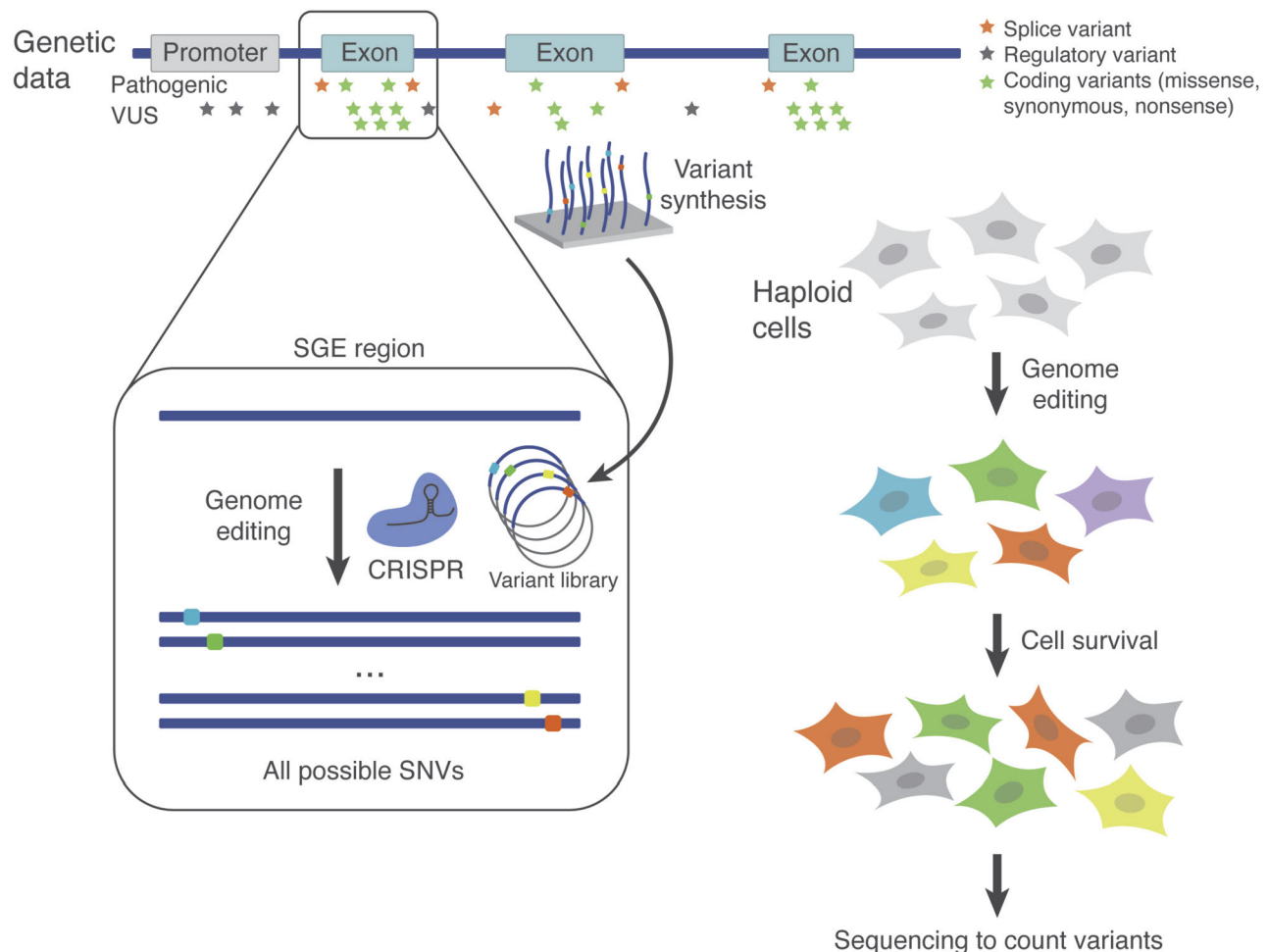

Figure 2: The Saturation Genome Editing method for testing human variants. Variant libraries are designed to include all single nucleotide variants in a genomic region of interest, such as an exon of a gene harboring many VUS. Variants are synthesized in an oligo pool, cloned into a library, and integrated to the genome of human cells with CRISPR technology. In the HAP1 essentiality assay (right), variants are created in a gene required for cell growth. Loss-of-function variants lead to growth defects and are depleted from the population over time. Next-generation sequencing is used to quantify variant abundances at multiple timepoints, such that variant effects can be deduced from sequencing data.

tral variants by performing SGE of the intron lariat debranche *DBR1* in a haploid human line, HAP1 [29]. Allowing edited cells to proliferate in culture and measuring variant depletion in the cell population over time revealed nonsense variants and missense variants at the enzyme's active site to be highly deleterious [26]. We also demonstrated that the same experimental workflow could be used to ask specific molecular questions, such as how different variants impact splicing.

Analyzing thousands of *BRCA1* variants with SGE

We next applied SGE to study variants in *BRCA1*, a tumor suppressor gene emblematic of the challenge posed by clinical variant interpretation. BRCA1 is involved in maintaining genomic stability and DNA repair by homologous recombination and germline mutations leading to loss-of-function cause hereditary breast and ovarian cancer. Like many frequently tested genes, *BRCA1* harbors thousands of VUS, typically rare missense and exon-proximal intronic variants with unknown effects. Accurately identifying pathogenic *BRCA1* variants is crucial for ensuring patients and their families are presented with options to reduce their cancer risk [4].

We prioritized assaying all possible SNVs in BRCA1's RING and BRCT domains, which harbor most of the gene's established pathogenic missense variants [22]. To do this, we applied an optimized SGE protocol to characterize a total of 3,893 SNVs across 13 exonic regions, culminating in a variant effect map of the two domains (Fig. 3A). Variants such as nonsense variants expected to cause loss-of-function consistently scored lowly, whereas only ~1% of synonymous variants proved deleterious to cells. Missense variants tended to either score lowly or have no effect on function, but a small fraction scored intermediately.

The clinical impact of SGE data

It was initially unclear how useful the *BRCA1* SGE data would be for variant interpretation, chiefly because HAP1 cells had not been used previously as a model for *BRCA1* function and lack clear relevance to breast and ovarian cancer. Therefore, analyzing how accurately the SGE assay adjudicated well-established pathogenic and benign variants was critical for determining utility. Comparing the SGE data for *BRCA1* variants to established interpretations in ClinVar [1] revealed pathogenic variants were identified with >95% specificity and sensitivity [22] (Fig. 3B). Crucially, this analysis revealed the assay to be highly accurate independent of variant type (e.g., nonsense, missense, synonymous, intronic), owing to the fact variants were edited into the human genome with CRISPR.

Several recent studies have independently validated the utility of SGE data for adjudicating clinically observed variants in conjunction with other lines of evidence. A comprehensive comparison of different BRCA1 assays revealed SGE to be among the most accurate despite having much greater scalability [30]. An analysis of variants observed in a breast cancer cohort revealed women with a VUS defined as loss-of-function by SGE had indistinguishable clinical characteristics to women harboring known pathogenic variants [31]. Another group asked whether SGE data predicted *BRCA1*-related cancer risk in an unbiased cohort of over 92,000 individuals. Indeed, women with variants measured to be loss-of-function had much higher risk of breast and ovarian cancer [32]. Lastly, Fayer et al. found that approximately half of the VUS in *BRCA1* tested via SGE could be clinically reclassified as "likely pathogenic" or "likely benign" [33], illustrating the potential to substantially improve the value of genetic testing for many patients.

Recent implementations of SGE and future advances

The clinical utility of *BRCA1* SGE data has spurred efforts to apply the method to other disease genes, including *CARD11* and *DDX3X* [34, 35]. In both studies, SGE again proved accurate at predicting disease risk conferred by variants, suggesting simple cellular models may be sufficient for determining variant pathogenicity across a range of genetic disorders. Like *BRCA1*, many other tumor suppressors are essential in the HAP1 line [36], indicating the same assay used for *BRCA1* may prove valuable for resolving variants in many more genes (e.g., *BRCA2*, *PALB2*, *RAD51D*, *RAD51C*) [37]. Indeed, efforts to perform SGE on many more cancer predisposition genes are ongoing [38]. It will remain critical that each new data set for a gene be carefully calibrated against clinical evidence. Towards this end, continued open sharing of genetic data and assay data will be essential for ensuring maximal impact. Initiatives to facilitate sharing include MaveDB [39], CanVIG [40], and the BRCA Exchange [41], and functional data
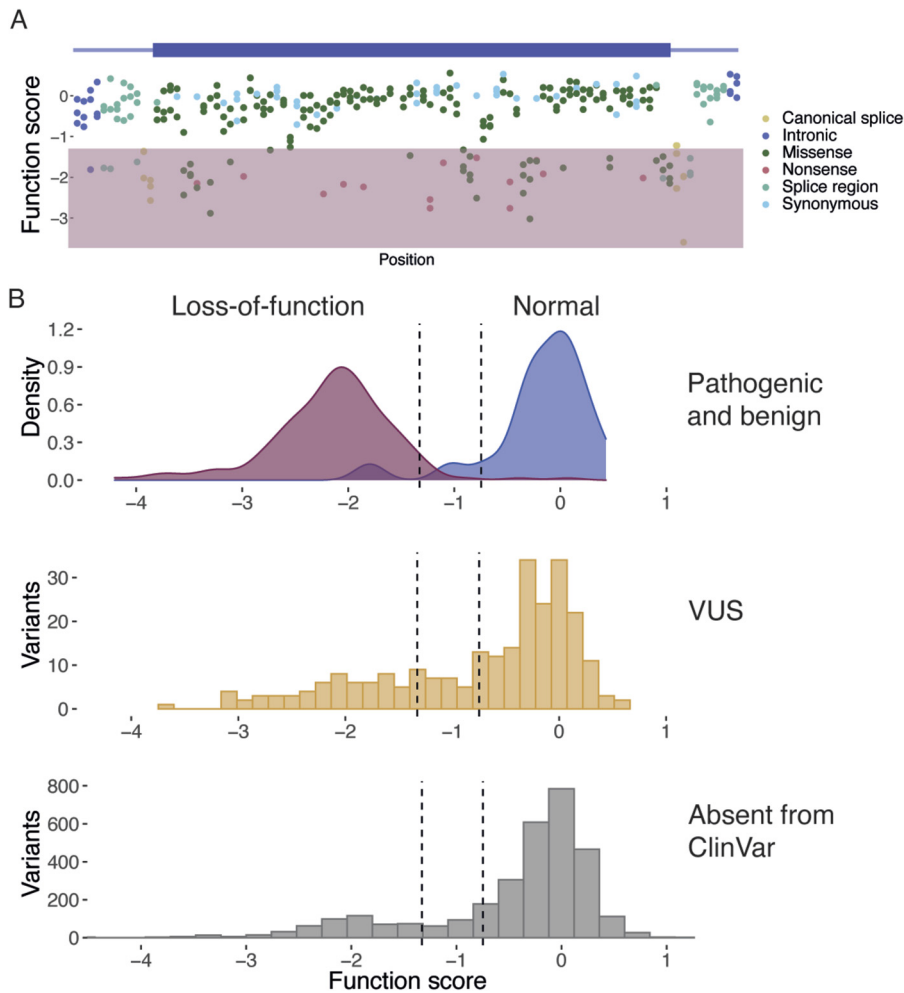
Figure 3: SGE of *BRCA1* improves clinical variant interpretation. SGE was used to produce a variant effect map of BRCA1 RING and BRCT domains, comprising function scores for each of 3,893 SNVs across 13 exons. (A) Function scores for SNVs in a single exonic region (exon 4) are plotted by position. Variants scoring similarly to nonsense variants are deemed loss-of-function variants (shaded). (B) Function scores discriminate known pathogenic variants with >95 % accuracy (top), and can be used as strong evidence for classifying VUS (middle) and variants yet to be reported in ClinVar (bottom). Data from Findlay et al. 2018.

can be both deposited on their own or cited as evidence in ClinVar [1].

While the HAP1 line has proven useful for assessing variant effects in numerous genes, in order to perform different functional assays that require other cell types to reflect biologically meaningful effects, it will be necessary to develop models of physiologically diverse diseases that are amenable to SGE. Haploidization of specific gene loci has been demonstrated as a strategy to allow editing of a single allele, permitting variant interpretation in diploid cell lines [25]. In order to understand the effects of variants on more complex disease processes, for instance developmental disorders, improved tools for generating and tracking variant libraries in organoid systems and model organisms will be required. Applying SGE and similar ap-

proaches to study variants with more subtle functional effects, such as those identified by genome-wide association studies, will require further technology development. Assays yielding highly quantitative data on gene expression in cell models which faithfully recapitulate effects of disease-associated noncoding variants will be particularly valuable for this task.

As CRISPR technology continues to improve, the ease of performing SGE in increasingly complex models of disease will inevitably increase. Indeed, newer editing reagents are already allowing more variants to be tested. CRISPR "base editor" screens [42, 43], in which tens of thousands of variants across multiple genes are engineered and assayed simultaneously, have emerged as a means of creating more variants per experiment, al-

beit with notable trade-offs. Another recent CRISPR-based technology called "prime editing" [44] has proven well suited for achieving saturation mutagenesis of endogenous regions of *NPC1* and *BRCA2* [25]. It is inevitable that larger variant effect maps will be generated in the near future using genome editing methods.

The forthcoming variant effect maps for many disease genes will bring challenges regarding how to best integrate and interpret orthologous data. For instance, different MAVEs covering variants in the same gene may produce contradictory results. In the rare instances where assay data contradict established interpretations from human genetic data, investigating the molecular basis of the discrepancies may prove necessary, and doing so will likely further our understanding of disease mechanisms. For variants that score intermediately in a functional assay, there may not be enough clinical data to calibrate disease risk. Yet over the next decade, having functional data for far more variants than ever before will surely improve efforts to establish more quantitative frameworks for genetic risk that go beyond simple "pathogenic" and "benign" distinctions. Ultimately, large amounts of unbiased functional data will also likely improve computational models of variant effect [45].

Conclusions

Experimental methods such as SGE that enable high-throughput characterization of human variants may soon bring about an era in which clinical geneticists have prospective functional data on any possible mutation likely to be seen in a disease-associated gene. While the technologies ushering in this new era are still improving, the high clinical utility of recent data sets suggest these methods will prove highly valuable for eliminating the considerable amount of uncertainty that exists in genetic testing today.

References

[1]   Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J, Jang W, Katz K, Ovetsky M, Riley G, Sethi A, Tully R, Villamarin-Salomon R, Rubinstein W, Maglott DR. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 2016;44:D862–8.

[2]   Acuna-Hidalgo R, Veltman JA, Hoischen A. New insights into the generation and role of de novo mutations in health and disease. Genome Biol. 2016;17:241.

[3]   Ndugga-Kabuye MK, Issaka RB. Inequities in multi-gene hereditary cancer testing: lower diagnostic yield and higher VUS rate in individuals who identify as Hispanic, African or Asian and Pacific Islander as compared to European. Fam Cancer. 2019;18:465–9.

[4]   Ludwig KK, Neuner J, Butler A, Geurts JL, Kong AL. Risk reduction and survival benefit of prophylactic surgery in BRCA mutation carriers, a systematic review. Am J Surg. 2016;212:660–9.

[5]   Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. Science. 2017;355:1152–8.

[6]   Boyiadzis MM, Kirkwood JM, Marshall JL, Pritchard CC, Azad NS, Gulley JL. Significance and implications of FDA approval of pembrolizumab for biomarker-defined disease. J Immunother Cancer. 2018;6:35.

[7]   Jonasch E, Donskov F, Iliopoulos O, Rathmell WK, Narayan VK, Maughan BL, Oudard S, Else T, Maranchie JK, Welsh SJ, Thamake S, Park EK, Perini RF, Linehan WM, Srinivasan R. Belzutifan for Renal Cell Carcinoma in von Hippel–Lindau Disease. N Engl J Med. 2021;385:2036–46.

[8]   Halldorsson BV, Eggertsson HP, Moore KHS, Hauswedell H, Eiriksson O, Ulfarsson MO, Palsson G, Hardarson MT, Oddsson A, Jensson BO, Kristmundsdottir S, Sigurpalsdottir BD, Stefansson OA, Beyter D, Holley G, Tragante V, Gylfason A, Olason PI, Zink F, Asgeirsdottir M, Sverrisson ST, Sigurdsson B, Gudjonsson SA, Sigurdsson GT, Halldorsson GH, Sveinbjornsson G, Norland K, Styrkarsdottir U, Magnusdottir DN, Snorradottir S, Kristinsson K, Sobech E, Jonsson H, Geirsson AJ, Olafsson I, Jonsson P, Pedersen OB, Erikstrup C, Brunak S, Ostrowski SR, DBDS Genetic Consortium, Thorleifsson G, Jonsson F, Melsted P, Jonsdottir I, Rafnar T, Holm H, Stefansson H, Saemundsdottir J, Gudbjartsson DF, Magnusson OT, Masson G, Thorsteinsdottir U, Helgason A, Jonsson H, Sulem P, Stefansson K. The sequences of 150,119 genomes in the UK Biobank. Nature. 2022;607:732–40.

[9]   Lappalainen T, MacArthur DG. From variant to function in human disease genetics. Science. 2021;373:1464–8.

[10] Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res. 2019;47:D886–94.

[11] Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D,

Cannon-Albright LA, Teerlink CC, Stanford JL, Isaacs WB, Xu J, Cooney KA, Lange EM, Schleutker J, Carpten JD, Powell IJ, Cussenot O, Cancel-Tassin G, Giles GG, MacInnis RJ, Maier C, Hsieh C-L, Wiklund F, Catalona WJ, Foulkes WD, Mandal D, Eeles RA, Kote-Jarai Z, Bustamante CD, Schaid DJ, Hastie T, Ostrander EA, Bailey-Wilson JE, Radivojac P, Thibodeau SN, Whittemore AS, Sieh W. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet. 2016;99:877–85.

[12] Muiños F, Martínez-Jiménez F, Pich O, Gonzalez-Perez A, Lopez-Bigas N. In silico saturation mutagenesis of cancer genes. Nature. 2021;596:428–32.

[13] Frazer J, Notin P, Dias M, Gomez A, Min JK, Brock K, Gal Y, Marks DS. Publisher Correction: Disease variant prediction with deep generative models of evolutionary data. Nature. 2022;601:E7.

[14] Wu Y, Liu H, Li R, Sun S, Weile J, Roth FP. Improved pathogenicity prediction for rare human missense variants. Am J Hum Genet. 2021;108:2389.

[15] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL. ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the. American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17:405–24.

[16] Gasperini M, Starita L, Shendure J. The power of multiplexed functional analysis of genetic variants. Nat Protoc. 2016;11:1782–7.

[17] Weile J, Roth FP. Multiplexed assays of variant effects contribute to a growing genotype-phenotype atlas. Hum Genet. 2018;137:665–78.

[18] Findlay GM. Linking genome variants to disease: scalable approaches to test the functional impact of human mutations. Hum Mol Genet. 2021;30:R187–97.

[19] Fowler DM, Fields S. Deep mutational scanning: a new style of protein science. Nat Methods. 2014;11:801–7.

[20] Patwardhan RP, Lee C, Litvin O, Young DL, Pe'er D, Shendure J. High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. Nat Biotechnol. 2009;27:1173–5.

[21] Inoue F, Ahituv N. Decoding enhancers using massively parallel reporter assays. Genomics. 2015;106:159–64.

[22] Findlay GM, Daza RM, Martin B, Zhang MD, Leith AP, Gasperini M, Janizek JD, Huang X, Starita LM, Shendure J. Accurate classification of BRCA1 variants with saturation genome editing. Nature. 2018;562:217–22.

[23] Jia X, Burugula BB, Chen V, Lemons RM, Jayakody S, Maksutova M, Kitzman JO. Massively parallel functional testing of MSH2 missense variants conferring Lynch syndrome risk. Am J Hum Genet. 2021;108:163–75.

[24] Mighell TL, Thacker S, Fombonne E, Eng C, O'Roak BJ. An Integrated Deep-Mutational-Scanning Approach Provides Clinical Insights on PTEN Genotype-Phenotype Relationships. Am J Hum Genet. 2020;106:818–29.

[25] Erwood S, Bily TMI, Lequyer J, Yan J, Gulati N, Brewer RA, Zhou L, Pelletier L, Ivakine EA, Cohn RD. Saturation variant interpretation using CRISPR prime editing. Nat Biotechnol. 2022;40:885–95.

[26] Findlay GM, Boyle EA, Hause RJ, Klein JC, Shendure

J. Saturation editing of genomic regions by multiplex homology-directed repair. Nature. 2014;513:120–3.

[27] Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8:2281–308.

[28] Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, Gill S, Harrington WF, Pantel S, Krill-Burger JM, Meyers RM, Ali L, Goodale A, Lee Y, Jiang G, Hsiao J, Gerath WFJ, Howell S, Merkel E, Ghandi M, Garraway LA, Root DE, Golub TR, Boehm JS, Hahn WC. Defining a Cancer Dependency Map. Cell. 2017;170:564–576.e16.

[29] Essletzbichler P, Konopka T, Santoro F, Chen D, Gapp BV, Kralovics R, Brummelkamp TR, Nijman SMB, Bürckstümmer N. Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. Genome Res. 2014;24:2059–65.

[30] Lyra PCM Jr, Nepomuceno TC, de Souza MLM, Machado GF, Veloso MF, Henriques TB, Dos Santos DZ, Ribeiro IG, Ribeiro RS Jr, Rangel LBA, Richardson M, Iversen ES, Goldgar D, Couch FJ, Carvalho MA, Monteiro ANA. Integration of functional assay data results provides strong evidence for classification of hundreds of BRCA1 variants of uncertain significance. Genet Med. 2021;23:306–15.

[31] Wan Q, Hu L, Ouyang T, Li J, Wang T, Fan Z, Fan T, Lin B Xu Y XY. Clinical phenotypes combined with saturation genome editing identifying the pathogenicity of BRCA1 variants of uncertain significance in breast cancer. Fam Cancer. 2021;20:85–95.

[32] Schiabor Barrett KM, Masnick M, Hatchell KE, Savatt JM, Banet N, Buchanan A, Willard HF. Clinical validation of genomic functional screen data: Analysis of observed BRCA1 variants in an unselected population cohort. HGG Adv. 2022;3:100086.

[33] Fayer S, Horton C, Dines JN, Rubin AF, Richardson ME, McGoldrick K, Hernandez F, Pesaran T, Karam R, Shirts BH, Fowler DM, Starita LM. Closing the gap: Systematic integration of multiplexed functional data resolves variants of uncertain significance in BRCA1, TP53, and PTEN. Am J Hum Genet. 2021. 108:2248–58.

[34] Meitlis I, Allenspach EJ, Bauman BM, Phan IQ, Dabbah G, Schmitt EG, Camp ND, Torgerson TR, Nickerson DA, Bamshad MJ, Hagin D, Luthers CR, Stinson JR, Gray J, Lundgren I, Church JA, Butte MJ, Jordan MB, Aceves SS, Schwartz DM, Milner JD, Schuval S, Skoda-Smith S, Cooper MA, Starita LM, Rawlings DJ, Snow AL, James RG. Multiplexed Functional Assessment of Genetic Variants in CARD11. Am J Hum Genet. 2020;107:1029–43.

[35] Radford EJ, Tan HK, Andersson MHL, Stephenson JD, Gardner EJ, Ironfield H, Waters AJ, Gitterman D, Lindsay S, Abascal F, Martincorena I, Kolesnik A, Ng-Cordell E, Firth HV, Baker K, Perry JRB, Adams DJ, Gerety SS, Hurles ME. Saturation genome editing of DDX3X clarifies pathogenicity of germline and somatic variation. bioRxiv. 2022. https://doi.org/10.1101/2022.06.10.22276179.

[36] Blomen VA, Májek P, Jae LT, Bigenzahn JW, Nieuwenhuis J, Staring J, Sacco R, van Diemen FR, Olk N, Stukalov A, Marceau C, Janssen H, Carette JE, Bennett KL, Colinge J, Superti-Furga G, Brummelkamp TR. Gene essentiality and synthetic lethality in haploid human cells. Science. 2015;350:1092–6.

[37] Miller DT, Lee K, Chung WK, Gordon AS, Herman GE, Klein TE, Stewart DR, Amendola LM, Adelman K, Bale SJ, Gollob MH, Harrison SM, Hershberger RE, McKelvey K, Richards CS, Vlangos CN, Watson MS, Martin CL. ACMG Secondary Findings

Working Group. ACMG SF v3.0 list for reporting of secondary findings in clinical exome and genome sequencing: a policy statement of the American College of Medical Genetics and Genomics (ACMG). Genet Med. 2021;23:1381–90.

[38] Kuang Da, Weile J, Kishore N, Rubin AF, Fields S, Fowler DM, Roth FP. MaveRegistry a collaboration platform for multiplexed assays of variant effect. Bioinformatics. Published Online First 27 March 2021. https://doi.org/10.1093/bioinformatics/btab215.

[39] Esposito D, Weile J, Shendure J, Starita LM, Papenfuss AT, Roth FP, Fowler DM, Rubin AF. MaveDB: an open-source platform to distribute and interpret data from multiplexed assays of variant effect. Genome Biol. 2019;20:223.

[40] Garrett A, Callaway A, Durkie M, Cubuk C, Alikian M, Burghel GJ, Robinson R, Izatt L, Talukdar S, Side L, Cranston T, Palmer-Smith S, Baralle D, Berry IR, Drummond J, Wallace AJ, Norbury G, Eccles DM, Ellard S, Lalloo F, Gareth Evans D, Woodward E, Tischkowitz M, Hanson H, Turnbull C. Cancer Variant Interpretation Group UK (CanVIG-UK): an exemplar national subspecialty multidisciplinary network. J Med Genet. 2020;57:829–34.

[41] Cline MS, Liao RG, Parsons MT, Paten B, Alquaddoomi F, Antoniou A, Baxter S, Brody L, Cook-Deegan R, Coffin A, Couch FJ, Craft B, Currie R, Dlott CC, Dolman L, den Dunnen JT, Dyke SOM, Domchek SM, Easton D, Fischmann Z, Foulkes WD, Garber J, Goldgar D, Goldman MJ, Goodhand P, Harrison S, Haussler D, Kato K, Knoppers B, Markello C, Nussbaum R, Offit K, Plon SE, Rashbass J, Rehm HL, Robson M, Rubinstein WS, Stoppa-Lyonnet D, Tavtigian S, Thorogood A, Zhang C, Zimmermann M, BRCA Challenge Authors, Burn J, Chanock S, Rätsch G, Spurdle AB. BRCA Challenge: BRCA Exchange as a global resource for variants in BRCA1 and BRCA2. PLoS Genet 2018; 14:e1007752.

[42] Hanna RE, Hegde M, Fagre CR, DeWeirdt PC, Sangree AK, Szegletes Z, Griffith A, Feeley MN, Sanson KR, Baidi Y, Koblan LW, Liu DR, Neal JT, Doench JG. Massively parallel assessment of human variants with base editor screens. bioRxiv. 2020;2020.05.17.100818.

[43] Cuella-Martin R, Hayward SB, Fan X, Chen X, Taglialatela A, Leuzzi G, Zhao J, Rabadan R, Lu C, Shen Y, Ciccia A. Functional interrogation of DNA damage response variants with base editing screens. Cell. 2021;184:1081–1097.e19.

[44] Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, Chen PJ, Wilson C, Newby GA, Raguram A, Liu DR. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature. 2019;576:149–57.

[45] Gray VE, Hause RJ, Luebeck J, Shendure J, Fowler DM. Quantitative Missense Variant Effect Prediction Using Large-Scale Mutagenesis Data. Cell Syst. 2018;6:116–124.e3.

Phoebe Dace
The Genome Function Laboratory, The Francis Crick Institute, 1 Midland Rd, London, United Kingdom

Gregory M. Findlay
The Genome Function Laboratory, The Francis Crick Institute, 1 Midland Rd, London, United Kingdom
greg.findlay@crick.ac.uk