2022-10

Comment on: Disease gene identification strategies for exome sequencing by Gilissen et al. 2012

Baptista, Julia

10.1038/s41431-022-01164-0
European Journal of Human Genetics
Springer Nature [academic journals on nature.com]

Comment on: Disease gene identification strategies for exome Sequencing by Gilissen et al 2012

Author information:
Julia Baptista
Peninsula Medical School, Faculty of Health, University of Plymouth
Julia.baptista@plymouth.ac.uk

A decade ago, Gilissen and colleagues provided a detailed review of disease gene identification strategies for exome sequencing, a technology part of a suite of Next Generation Sequencing tools that represented a paradigm shift to the way the genetics and genomics community investigated rare genetic conditions [1]. The prediction was that not only exome sequencing would become the most utilised tool for Mendelian disease gene identification in the years to follow, but that together with genome sequencing, it would pave the way to a future where all clinically relevant variants in an individual's genome would be identified leading to personalised medicine [2]. Their review [1] stands as a useful summary and a reference point for healthcare professionals with varying levels of genomic expertise and students from different medical disciplines aiming to start their learning journey in the field of genomics.

Exome sequencing has since taken a prominent role as a first-line diagnostic test to interrogate: the exome collection of the over 20,000 human genes; large gene panels in highly heterogeneous disorders (e.g., epilepsy, intellectual disabilities); smaller gene panels; and even single genes as laboratories embrace the choice of relying upon a single streamlined laboratory workflow to benefit from higher throughputs, faster turnaround times and the flexibility of expanding the initial analysis to include additional genes if no diagnosis is

identified and the clinical presentation suggests that “opening the exome” is indicated.

Efforts to solve rare conditions have focused on understanding and applying adequate strategies for disease gene and variant identification following exome sequencing, this is vital as the recognition that data prioritisation and interpretation are considerable challenges to overcome is indisputable. Six strategies are discussed by the authors: linkage, homozygosity, double-hit, overlap, *de novo,* and candidate; and these are based on traditional and common approaches for analysing genetic and genomic data. Factors taken into consideration for strategy design include the sporadic nature of the disorder *versus* the observation of familial cases, the number of affected individuals available for testing, the availability of a single *versus* multiple families, the availability of both parental samples, the assumed genotype of the causative variant in the context of the family/population structure, and the biological characteristics of the disorder or any prioritised candidate causative gene(s) or variant(s) [1]. The complexity of the challenge often calls for combining different approaches and adapting current methods in a bid to harness the data’s power and identify a diagnosis for patients and their families [3]. Whilst devising tailored strategies, to solve individual projects, is likely to increase the chances of success, the systematic use of *ad hoc* approaches is likely to apply only to a lesser extent in diagnostic laboratories. In these settings, where constrained workforce resources and the need for standardised methods are limiting factors, there might be a reliance on two main strategies, typically, gene-agnostic inheritance-based approaches when both parental samples are available or when the family structure is highly suggestive of a homozygous or X-linked variant as the cause, and gene panel-based approaches where the

analysis is restricted to a pre-selected list of genes that are known to be causative of the phenotype being investigated. Testing of prenatal samples deserves special consideration with current practice favouring the latter strategy, with a focus on well-characterised genes to reduce unwanted uncertainty, however, it cannot completely exclude findings unrelated to the reason for referral, particularly given the evolving nature of the phenotype [4].

The need for the delineated analysis strategies to be underpinned by robust bioinformatic pipelines and access to databases of genomic variants that contain well-curated and reliable data was also discussed alongside some of the reasons for not identifying a genetic diagnosis. Limits in the technology, the bioinformatic pipelines used to process the data and the current knowledge to interpret uncertain or unknown variants were amongst the culprits. Nonetheless, the authors reported a success rate of 60% which is likely to also reflect a selection of patients with a higher prior probability of having a monogenic condition [1]. Variable success rates of ~30- 40% were typically reported in other exome sequencing studies [5, 6].

As the implementation of whole genome sequencing is now well underway, for example within the UK NHS Genomic Medicine Service, it would be of interest to evaluate the impact of the transition from exome sequencing to genome sequencing, especially concerning the diagnostic rates for patients with rare conditions [7]. Whole genome sequencing offers the potential to overcome some of the limitations of exome sequencing. For example, it allows for better sequence coverage, the ability to interrogate well more than the 1-2% of exonic sequences, and has increased power to solve complex genomic regions and to detect structural variants. Issues with interpreting the data in the context of the

clinical presentation and current knowledge available are technology-independent and will require alternative solutions. Interestingly, Whole Genome Sequencing might be perceived by healthcare professionals, patients and their families, as a comprehensive test that interrogates the full genome. It is important to highlight that also here there will be a strategy needed for analysing the data, the most commonly applied being a virtual gene-panel analysis with additional limited analysis of variants outside the selected panels, therefore, the review by Gilissen *et al.* remains a pertinent reference material.

Irrespective of the technology applied to identify diagnoses in patients with suspected rare conditions, healthcare professionals should aim to gain an understanding of the strategies for data analysis, their limitations, and what are the implications for patients when a genetic diagnosis is not identified. This understanding will underpin meaningful conversations within the multidisciplinary teams to ensure that where appropriate, patients can benefit from tailored analysis strategies and data re-analysis.

1. Gilissen C, Hoischen A, Brunner HG, Veltman JA. Disease gene identification strategies for exome sequencing. Eur J Hum Genet. 2012 May;20(5):490-7. doi: 10.1038/ejhg.2011.258. Epub 2012 Jan 18. PMID: 22258526

2. Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? J Med Genet. 2011 Sep;48(9):580-9. doi: 10.1136/jmedgenet-2011-100223. Epub 2011 Jul 5. PMID: 21730106.

3. Kaplanis J, Samocha KE, Wiel L, Zhang Z, Arvai KJ, Eberhardt RY wt al. Evidence for 28 genetic disorders discovered by combining healthcare and research data. Nature. 2020 Oct;586(7831):757-762. doi: 10.1038/s41586-020-2832-5. Epub 2020 Oct 14. PMID: 33057194

4. Lord J, McMullan DJ, Eberhardt RY, Rinck G, Hamilton SJ, Quinlan-Jones E et al. Prenatal Assessment of Genomes and Exomes Consortium. Prenatal exome sequencing analysis in fetal structural anomalies detected by ultrasonography (PAGE): a cohort study. Lancet. 2019 Feb 23;393(10173):747-757. doi: 10.1016/S0140-6736(18)31940-8. Epub 2019 Jan 31. PMID: 30712880

5. Vissers LELM, van Nimwegen KJM, Schieving JH, Kamsteeg EJ, Kleefstra T, Yntema HG, Pfundt R et al. A clinical utility study of exome sequencing versus conventional genetic testing in pediatric neurology. Genet Med. 2017 Sep;19(9):1055-1063. doi: 10.1038/gim.2017.1. Epub 2017 Mar 23. PMID: 28333917

6. Wright CF, McRae JF, Clayton S, Gallone G, Aitken S, FitzGerald TW et al. Making new genetic diagnoses with old data: iterative reanalysis and reporting from genome-wide data in 1,133 families with developmental disorders. Genet Med. 2018 Oct;20(10):1216-1223. doi: 10.1038/gim.2017.246. Epub 2018 Jan 11. PMID: 29323667

7. Turro E, Astle WJ, Megy K, Gräf S, Greene D, Shamardina O et al. Whole-genome sequencing of patients with rare diseases in a national health

system. Nature. 2020 Jul;583(7814):96-102. doi: 10.1038/s41586-020-2434-2. Epub 2020 Jun 24. PMID: 32581362

Competing interests

The author declares no competing interests.

Funding

None.