Published in final edited form as:
Hum Mutat. 2012 May ; 33(5): 884–886. doi:10.1002/humu.22048.

Next Generation Sequencing Demands Next Generation Phenotyping

Raoul CM Hennekam[1] and **Leslie G Biesecker**[2]

[1]Department of Pediatrics and Translational Genetics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands [2]Genetic Disease Research Branch, National Human Genome Research Institute, NIH, Bethesda, Maryland, USA

Abstract

Next-Generation Sequencing (NGS) is the most powerful diagnostic tool since the roentgenogram. NGS will facilitate diagnosis on a massive scale –allowing interrogation of all genes in a single assay. It has been suggested that NGS will decrease the need for phenotyping in general, and medical geneticists in particular. We argue that NGS will shift focus and approach of phenotyping. We predict that NGS performed for diagnostic purposes will yield variants in several genes, and consequences of these variants will need to be analyzed and integrated with clinical findings to make a diagnosis. Diagnostic skills of medical specialists will shift from a pre-NGS-test differential diagnostic mode to a post-NGS-test diagnostic assessment mode. In research phenotyping and medical genetic assessments will remain essential as well. NGS can identify primary causative variants in phenotypes inherited in a Mendelian pattern, but biology is much more complex. Phenotypes are caused by the actions of several genes, and epigenetic and environmental influences. Dissecting all influences necessitates ongoing and detailed phenotyping, refinement of clinical diagnostic assignments, and iterative analyses of NGS data. We conclude that there will be a critical need for phenotyping and clinical analysis and that medical geneticists are uniquely positioned to address this need.

Keywords

NGS; whole exome sequencing; whole genome sequencing; phenotype; dysmorphology; Mendelian; monogenic

Introduction

Next Generation Sequencing (NGS) presents a paradigm shift for medicine and has the potential to allow more tailored (personal) medical care based on individual risk. As NGS will not only be useful in disorders with a Mendelian pattern of inheritance but also in polygenic and multifactorial etiologies, it will be become a component of medical practice for all disorders, and all medical specialties will benefit from these developments. However, it has been suggested that one medical discipline may instead be a victim of this technology: Medical Genetics, specifically Dysmorphology. During a plenary session of the 2011 meeting of the European Society of Human Genetics an expert and well-respected speaker made this prediction, and one of the topics of a public debate in the 2011 International Congress of Human Genetics in Montreal is '…medical genetics will disappear as a separate

Correspondence to Dr. R.C. Hennekam, Department of Paediatrics, Floor H7-236, Academic Medical Centre, Meibergdreef 9, 1105AZ Amsterdam, The Netherlands. r.c.hennekam@amc.uva.nl.

specialty.' To paraphrase Mark Twain, the rumors of the death of dysmorphology and medical genetics are greatly exaggerated. So if the rumored death is not imminent, what is afoot?

This rumor of demise reflects not an error of judgment, but a lack of imagination and foresight. We predict that instead of a demise there will be a transformation – and this transformation will be dramatic for medical genetics researchers and clinicians, but most importantly, for the patients. A main change will be in the approach to diagnosis. Until now, diagnostic molecular tests in clinical medicine have been low-throughput and expensive. The consequence is that clinicians spend enormous amounts of time (so money) gathering data (medical and family history, physical findings, imaging, etc.), lumping and splitting patients into precise categories, and refining and debating differential diagnoses *to allow them to select a single test or a small set of tests* to determine the most likely clinical diagnosis. The diagnostic process has not been organized in this way because it is intrinsically superior but because this has been the best way to match the throughput of testing to diagnostic realities. Now that testing has become high-throughput, clinicians will ask whether that approach could be completely re-engineered. We predict that re-engineering of the diagnostic process will be driven by medical geneticists and will create a new, rewarding practice of medical genetics, and invigorate, not destroy the profession.

Medical genetics practice in a NGS-driven paradigm

Imagine a patient with ataxia. Currently, the clinician performs numerous tests such as electromyographies, cranial MRI's, etc. to gather evidence on the most likely type of ataxia and to help select the gene test that should be ordered. This consumes significant time and health care resources. Instead, with NGS, a minimal level of clinical assessment (a concise history, physical examination and discussion with the patient to obtain permission to interrogate the genome) will allow the clinician to access (of course securely) the genome server in which this patient's genome sequence is banked[1]. With a few keystrokes, the clinician can check all genes known to cause ataxia and find the one (or perhaps a few) with a mutation. If there is more than one 'hit' the clinician uses a knowledgebase regarding phenotypes associated with mutations to get recommendations on further clinical testing useful to determine which 'hit' is real and allowing identification of the causative mutation.

Imagine next a preschool-aged patient with an intellectual disability and unusual facial morphology. Currently, the primary care physician has little to contribute to the diagnostic process and would refer the patient to a medical geneticist. The geneticist would evaluate the child for additional malformations, search databases, consult with experienced colleagues, and start a process of ordering a series of individual tests, with an ultimate diagnostic yield of about 50% [Van Karnebeek et al., 2005]. Instead, with NGS, one might imagine that the primary care physician orders a single blood test (whole genome sequencing with copy number and structural variation assessment), perhaps combined with a sample from the parents, and will receive an analysis based on an input query of "intellectual disability and dysmorphic features".

If the NGS would identify a single, high-penetrance causative mutation, known to be associated with a recognized clinical entity, the family can be counseled by reviewing a knowledgebase of that disorder. If instead there were to be multiple candidate causative mutations (a more probable outcome) the primary care clinician will refer the family to a

[1]Note here we assume in an adult that the sequence was previously generated for other reasons and available to be re-used for this new purpose (at little marginal cost). The assumption is that genome interrogation will be inexpensive and likely to be used early in life (some say for newborn screening) and serve as a life-long resource to improve the health care of an individual.

medical geneticist or other appropriate specialist. Further analyses regarding the various possible diagnoses will be started, based on the genomic hits. The specialists will then use their clinical diagnostic skills to distinguish among the candidate mutations, potentially order focused follow-up tests (e.g., imaging), and make a diagnosis. This approach to diagnostics allows specialists to accelerate the diagnostic process by using the NGS test results to focus their evaluations on manifestations as indicated by the NGS results. The identification of a mutation in a causative gene will facilitate counseling, lead to treatment recommendations, allow accurate reproductive recurrence risk assessments, and offer families the opportunity for further testing within families and prenatal studies.

However, it is important to recognize that NGS and computers will not magically make diagnoses in all, or perhaps even most, patients. If a diagnosis is suggested, that will be wonderful. Typically however, NGS and informatics will provide a handful of possibilities and allow clinicians to perform focused and efficient further assessments of the patient. So, clinical skills and acumen of medical geneticists will remain essential, and only will shift from a pre-test differential diagnosis generation mode to a post-test diagnostic assessment mode. Medical geneticists will use their skills, deployed in a new way, to make more, more accurate and faster diagnoses at lower cost. These potential improvements in the diagnostic process should engender optimism and excitement among medical genetics and other specialties as well.

What do patients want to know?

Patients (or parents of young patients) ask simple, direct questions and clinicians try to answer them to the best of their abilities. These questions are; 1) ‘what do I have?’ 2) ‘what causes it?’ 3) ‘what are the consequences for my health?’ and finally, 4) ‘what can be done about it?’. NGS will markedly increase the clinician's ability to answer the first two questions – although the order in which these questions are answered will change. Patients are only rarely interested in the exact nature of a mutation they have but are interested in the consequences for their health (and for their relatives) like the likelihood to develop a particular manifestations of a disorder, how these manifestations evolve over time (natural history), and the need for specific health surveillance. Phenotypic research will be essential here: a detailed phenotype will allow researchers to compare and pool clinical findings in their patients and correlate these findings with the results of NGS. The improvement afforded by NGS is that the molecular diagnosis will be more readily identified and clinical evaluations can be performed in a directed way, avoiding stressful additional studies like biopsies or other invasive procedures.

Medical genetics in NGS research

Research techniques for the elucidation of the primary molecular etiology of Mendelian disease is at present more successful than ever and NGS technology is the main factor that has accelerated this process. One may expect that in just a few years' time the primary molecular basis for most currently recognized Mendelian disorders will have been solved. In a research context, phenotyping will remain important to identify mutations causing disorders. For disorders limited to a single organ system, most phenotyping should be performed by the relevant organ-specific specialists (e.g., cardiologists for non-syndromic congenital heart disease). However, for multisystem, pleiotropic disorders, the expertise of the medical geneticist will be essential.

A primary diagnostic label (e.g., “Marfan syndrome”) is not a sufficient level of clinical delineation to answer the patient's questions. Patients want to understand the full range of phenotypic manifestations that they may manifest, and the clinician will want to know where the patient is likely to fall on the spectrum of clinical severity. After all, the mild and

severe end of the Marfan spectrum are very different clinical scenarios. Most Mendelian disorders have significant phenotypic heterogeneity with inter-individual, intra-familial and inter-familial variability in expressivity and even penetrance. Patients (or parents) will ask whether they will have a mild or severe form of the disorder – and we cannot answer most of these questions given the current state of knowledge.

Most implementations of NGS focus on Mendelian disorders. 'Mendelian' indicates that a mutation (or two mutations for autosomal recessive disorders) explains a major part of the phenotype in an individual. However, there is no protein that does not interact with other proteins, and interaction means that the structure and function of both proteins influence the resulting function and, thus, phenotype. In addition, most proteins undergo post-translational processing and will not function in their primary, unprocessed form or must be transported to a specific subcellular or extracellular location. These processes are mediated by other proteins, which are themselves subject to inter-individual variation: they are termed modifiers. Thus, it will be essential to identify modifier loci and correlate variation at those modifying loci to predict phenotype. Every clinician has evaluated families in which a Mendelian disorder segregates, with substantial inter-individual phenotypic variation. These differences in manifestations will not be clear from the early results of NGS.

Beyond Mendelian disorders and their modifiers are the oligogenic and subsequently polygenic or multifactorial disorders. While it may be argued that all disorders are polygenic [Barabási et al, 2011] (as the so-called 'single gene' disorders have modifiers), we use the term here to distinguish disorders for which there are no variations at single loci that explain a large portion of the phenotypic variability of that disorder, yet heritability is high, and epidemiology suggests that a substantial number of genes influence the phenotype. The challenge here is largely unexplored – NGS may theoretically be used to discover variants associated with these phenotypes but new analyses will be necessary. For rare disorders inherited in a dominant pattern, we currently ignore variants that are present in databases like the 1000Genomes Project, as we can safely assume variants in these databases do not cause these disorders. However, these variants are prime candidates for the molecular basis of polygenic and multifactorial traits and their modifiers. New bio-informatics and systems biology methods will be paramount in such analyses. An important way to recognize these is by grouping patients based on their phenotype and selecting appropriate controls. The more carefully this phenotyping is performed, the higher the chance a variant influencing the presence of this manifestation will be found. New approaches to hypothesis-generating studies may also be needed. For example, one might analyze a large cohort of subjects with NGS sequence results to identify a subset who share a particular set of variants. Then, those subjects can be clinically analyzed to see if they share phenotypic attributes. Especially for disorders in which phenotyping is extremely difficult and heterogeneity is large (e.g., psychiatric disorders) [Xu etal., 2011] such strategies may be extremely useful. This hypothesis-generating mode of medical genetics research will necessitate a high level of multisystem clinical evaluation, something that can only be provided by the medical geneticist.

The future of Medical Genetics

Based on the above it is clear the future of medical genetics is bright. For the foreseeable future, medical geneticists will be essential in research to identify genes that are mutated in human disease using NGS, dissect the role of modifiers, develop new approaches to NGS clinical research, and act as teachers for other specialties to disseminate these new diagnostic tools. As well, medical geneticists will be necessary to develop databases and precise clinical language to characterize phenotypes. Activities like the recent series of papers to define all terms describing the external human phenotype [Allanson et al., 2009]

and other medical language ontology initiatives are essential to allow reliable comparisons of patients. As well, medical geneticists will need to adapt and broaden their skill sets. No longer will it be sufficient to be only a phenotypic expert – the most successful medical geneticists will combine and complement their clinical skills with expertise in molecular genetics (including NGS) and bioinformatics. Bioinformatic skills will be multifaceted – they will be as much about phenotypic and medical informatics as they will be about molecular data. Biomedicine is irrevocably moving toward large-scale datasets and the geneticist who can interrogate and analyze these data sets will have an enormous advantage over those who do not have these skills.

Summary

As translational researchers operating at the boundaries of the research laboratory and the clinic, we could not be more excited and optimistic about the changes that are underway. We foresee an increase in the need and demand for the medical geneticist, especially those with additional molecular and bioinformatic skills. Our discipline is rapidly changing and we believe that these changes will lead to dramatic improvements in our ability to diagnose and provide treatments for our patients. How we carry out these tasks will be quite different from how we currently approach both research and patient care and we are eager to jump in and lead the way in the creation of these new approaches. Next-Generation Sequencing is the most powerful diagnostic tool developed in medicine since the roentgenogram, and we predict that its value and utility in clinical medicine will be enormous. We also conclude that the value of NGS can only be optimally exploited if novel approaches to detailed phenotyping are integrated with NGS results. Medical geneticists do not yet need to put phone numbers of employment agencies in their mobiles. Instead, hospitals, academic medical centers, and research institutes will need to attract and support medical geneticists if they want to participate in these exciting developments, or be left on the sidelines.

Acknowledgments

Funding for LGB was provided by the Intramural Research Program of the National Human Genome Research Institute of the National Institutes of Health, USA.

References

Allanson JE, Biesecker LG, Carey JC, Hennekam RC. Elements of morphology: Introduction. Am J Med Genet. 2009; 149A:2–5. [PubMed: 19127575]

Barabási AL, Gulbahce N, Loscalzo J. Network Medicine: A network-based approach to human disease. Nat Rev Genet. 2011; 12:56–68. [PubMed: 21164525]

Van Karnebeek CD, Scheper FY, Abeling NG, Alders M, Barth PG, Hoovers JM, Koevoets C, Wanders RJ, Hennekam RC. Etiology of mental retardation in children referred to a tertiary care center: a prospective study. Am J Ment Retard. 2005; 110:253–267. [PubMed: 15941363]

Xu B, Roos JL, Dexheimer P, Boone B, Plummer B, Levy S, Gogos JA, Karayiorgou M. Exome sequencing supports a de novo mutational paradigm for schizophrenia. Nat Genet. 2011; 43:864–868. [PubMed: 21822266]