The transformation of medical genetics by clinical genomics: hubris meets humility

Wayne W. Grody, MD, PhD¹

There is no question that the advent of massively parallel (“next-generation”) DNA sequencing has thrust Medical Genetics and Molecular Diagnostics into a new era, availing practitioners and patients of a form of genetic testing unprecedented in its scope and comprehensiveness. It has produced impressive diagnostic yield, ended the “diagnostic odyssey” for many patients and families, expanded the known phenotypes of countless disorders, and led to almost weekly new disease gene discoveries. Nevertheless, it still fails to identify the molecular cause of many patients who clearly exhibit genetic/syndromic conditions, while at the same time unmasking other sequence changes of uncertain significance or unexpected consequences. With over six years’ experience in the clinical application of NGS, this seems an opportune time to take stock and face up honestly to how much we still do not know about

genome action and, indeed, the DNA molecule itself. This review and assessment examines a number of residual deficiencies and misconceptions in clinical genomics, while daring to predict its future incorporation of other “-omics” approaches and even quantum phenomena in our unending quest to understand the heredity of *Homo sapiens*.

Genetics in Medicine (2019) 21:1916–1926; <https://doi.org/10.1038/s41436-019-0450-0>

Keywords: exome/genome sequencing; variant interpretation; secondary findings; quantum biology; clinical genomics

“We shall assume the structure of a gene to be that of a huge molecule...which has to be a masterpiece of differentiated order, safeguarded by the conjuring rod of quantum theory.”

Erwin Schrödinger, *What Is Life?*, Chapter 6 (1944)

“There are more things in heaven and earth, Horatio, than are dreamt of in your philosophy.”

William Shakespeare, *Hamlet*, Act I, Scene v (1599)

INTRODUCTION

The human genome is often compared to a book or, more grandiloquently, a series of books like an encyclopedia of 23 or 46 volumes, with the individual letters in the text analogous to the four nucleotides of DNA, the sentences analogous to genes, and the volumes analogous to chromosomes. Of course, we all know this is overly simplistic; the text of a book is nothing more than countless black marks on a white surface, absolutely meaningless until interpreted by a reader who is proficient in the given language, possesses visual reception within a specific

wavelength range of the electromagnetic spectrum, and has the intellectual capacity to convert the abstract black marks into mental images with content, context, and meaning. Likewise, the sequence of the genome comprises a few billion nucleotides held in tandem by phosphodiester bonds, meaningless until interpreted by the cellular machinery—primarily transcription and translation—into useful instructions for the synthesis of proteins. And at the level of human perception, they are entirely invisible and unintelligible to us until detected, translated, and interpreted by sophisticated sequencing instruments and powerful computers. But unlike a book, in which most readers of a certain intelligence and reading skill would be expected to arrive at a reasonable approximation of the author’s intent, when it comes to reading the genome, we are little better than a two-year-old trying to make sense of the *Encyclopedia Britannica*.

Modern technology is indeed extremely impressive, and the instruments at our disposal for performing genomic sequencing are miracles of 21st century engineering. But, like previous paradigm-shifting advances in technology, such as the attainment of controlled nuclear fission, our technological abilities often exceed our understanding and, sometimes, our

¹Divisions of Medical Genetics and Molecular Diagnostics, Departments of Path. & Lab. Medicine, Pediatrics, and Human Genetics, UCLA School of Medicine, UCLA Institute for Society and Genetics, Molecular Diagnostic Laboratories and Clinical Genomics Center, UCLA Medical Center, Los Angeles, CA, USA.

Correspondence: Wayne W. Grody (wgrody@mednet.ucla.edu)

Submitted 1 October 2018; accepted: 22 January 2019

Published online: 7 March 2019

wisdom. The typical result is a feeling of hubris, accompanied by a desire to deploy the technology as quickly as possible. To the credit of the specialty, our progress has been guided by the careful development of detailed technical and clinical guidelines, sometimes in collaboration with other professional societies and stakeholders, many of them published in this journal.¹

As we just completed a year-long celebration of the 20th anniversary of *Genetics in Medicine*, much of which had been devoted to the astonishing advances in genomic technology, it is worth keeping this perspective in mind. Encompassed within those 20 years were two groundbreaking events: the completion of the Human Genome Project in 2003, and the advent of massively parallel or next-generation sequencing (NGS) in the mid-2000s. (One could also cite a third: the exponential expansion of computer power, without which neither of the other two accomplishments would have had quite the impact. The first issues of *GIM* were processed using Windows 98 and communicated over Netscape!) The Genome Project provided the tools and reference information to analyze the genes of patients with genetic disorders, and NGS enabled medical geneticists and molecular diagnostic laboratories to analyze all of the genes (and more) in individual patients at acceptable cost and turnaround time for clinical use. The impact was truly transformative, especially for patients with rare, unknown, or undiagnosed disorders. So impactful was it on our specialty that the parent organization of the journal changed its name to the American College of Medical Genetics and Genomics (ACMG)—followed in short order by its sister organization, the American Board of Medical Genetics and Genomics (ABMG). Likewise, the next edition of the most comprehensive textbook in the field will be entitled *Emery & Rimoin's Principles and Practice of Medical Genetics and Genomics*.

Is genomics, and more specifically genomic sequencing, a new subspecialty of medical genetics, or a new component of the specialty as a whole? Or is it merely a powerful new technology or clinical tool that will revolutionize the field, not unlike the way in which the development of tandem mass spectrometry revolutionized newborn screening? Is it exclusive to medical genetics, or is it also being incorporated into other specialties such as pathology, oncology, microbiology, and neurology? As even newer tools become available in the future—epigenomics, proteomics, metabolomics, glycomics—will they too need to be recognized by further appending the name of the specialty? This could quickly become a mouthful. Also, official adoption implies that we actually understand how the genome works, when in fact our current knowledge is fairly superficial. The primary DNA sequence of the human genome—which for now and the immediately foreseeable future is the fundamental basis of genomic medicine practice—is but one dimension of a four-dimensional structure (if Einstein's concept of time as a dimension [for our purposes, the lifespan of the patient] is included)—or, for all we know, an 11-dimensional structure, if string theory is correct. Later in this article these concepts will be proposed for consideration, for the time being within the context of genomic

diagnostics, but likely applicable as well to genome-based gene therapy and gene editing.

THE MOST POWERFUL GENETIC TEST?

One would need to hark all the way back to the early 1980s and the invention of the polymerase chain reaction (PCR) for the last technological development that was as transformative on clinical molecular genetics as is NGS. And even PCR did not carry with it the intellectual impact of NGS. The former technology was essentially a more efficient way to obtain the same knowledge as was previously done with the expensive and laborious Southern blot. There is no question that it drastically expanded the convenience and practicality of routine genetic testing, but it did not in and of itself shed new light on fundamental human genetic questions. NGS is not only more efficient in its output; the output itself is far broader and more comprehensive than anything that has come before, including the prior methods of DNA sequencing (most prominently Sanger sequencing). By providing a genome-wide view of the patient's genetic makeup, it sees things that were formerly, for all practical purposes, hidden from our view.

In contrast to the former workhorse of DNA sequencing, the Sanger method, which utilized primers specific to one small region of a single gene to produce about 150–200 bp of readable sequence, NGS—which made its first commercial appearance with the Roche-454 instrument in 2005², followed by a few short-lived competitor platforms and then the Illumina instrument in 2008³ and the Ion Torrent in 2011⁴—takes a “shotgun” approach, employing random priming of many millions of sheared DNA fragments to produce “massively parallel” sequence across the entire genome (Fig. 1). At present, most clinical genomics laboratories include an exon-capture step to limit the sequence output to the coding sequences of all the genes (the exome), reducing the cost and effort it would take to sequence and interpret the other 99% of the genome, about which far less is known. However, genome sequencing (GS) is now becoming more common (at least on a clinical research basis as in the Undiagnosed Disease Network,⁵ especially for patients highly suspected to have a genetic syndrome despite a negative exome sequencing result. In either case, the new platforms demanded some significant reorientation of molecular diagnosticians who were used to methods with relatively simple visual corroboration; now they needed to resign their trust to the inscrutable output of these instruments and the complex computer analytics required to make biologic sense of it (Fig. 2a, b).

Once these platforms further refined their core technologies and reduced their initially unsatisfactory error rates, they became impressive new fixtures in advanced clinical molecular genetics laboratories, and the impact on genetic disease diagnosis has been nothing less than transformative. Perhaps the strongest evidence for this is the impressive diagnostic yield of exome sequencing which surprised everyone within the first 1–2 years of its introduction. Clinical genomics laboratories around the world began

Fig. 1 Schematic of library preparation for next-generation sequencing, illustrating production of genomic fragments by shearing, followed by ligation to adaptors, amplification, capture of desired target sequences, bridge amplification to generate clonal clusters, denaturation, and purification. Sequencing is initiated by adding primers, polymerase, and labeled nucleotide substrates. (Courtesy of Dr. Hane Lee, UCLA).

reporting diagnostic yields of 25–50% (depending upon whether pathogenic and/or likely pathogenic variants and/or new disease gene discoveries are included). This is far higher than any other existing genetic test, whether molecular, cytogenetic, or biochemical. As such, it has been a boon to patients and families with ultrarare or unknown disorders who have been adrift, sometimes for years or even decades, in the “diagnostic odyssey.” Even when the discovered genetic cause has no treatment (as is sadly the case more often than not), families are grateful for the closure the test has provided; and, at the very least, the revealed pathogenic variant(s) can be used for prenatal diagnosis in future pregnancies.

Still, we should not let this relative success go to our heads, as there are still many situations in which clinical genomic

sequencing falls short, or even creates new problems for the tested families. There are ways to address these issues, but they require patience, a rejection of complacency, and an openness to continuing scientific advances over the long haul (which could well mean 50 years or more). The remainder of this article discusses some of the more prominent challenges, both short- and long-term.

VARIANTS OF UNCERTAIN SIGNIFICANCE AND SECONDARY FINDINGS: THE TWO THINGS THAT KEEP CLINICAL GENOMICISTS UP AT NIGHT

If there are two terms that have entered the daily vernacular of medical geneticists more than any other with the advent of genetic testing by NGS, they are these: variants of uncertain significance (VUS) and secondary findings (sometimes, as

Fig. 2 Contrasting data output of Sanger and next-generation DNA sequencing. (a) Contrasting raw data output of traditional Sanger sequencing (Applied Biosystems) and (b) next-generation sequencing (Illumina): the same four colors, but which can you interpret with the naked eye? (Courtesy of Dr. Josh Deignan, UCLA).

borrowed from other specialties, also called incidental findings). The first was not unknown in the days of single-gene sequencing, most prominently encountered in the sequencing of *BRCA1* and *BRCA2* for individuals at risk of familial breast/ovarian cancer. But the latter is something new, since prior single-gene and phenotype-specific gene panel tests would not normally encounter “off-target” findings. Furthermore, the sheer number of VUS and secondary findings revealed in an exome or genome sequence are many orders of magnitude beyond anything geneticists have had to deal with before. For these reasons, the challenge is completely unprecedented and, more than any other barrier to routine clinical NGS (cost, technical errors, insurance reimbursement), threatens to derail our best efforts.

ACMG produced guidelines for interpretation of novel sequence variants (i.e., those not previously documented to produce disease) back in the single-gene testing days.⁶ But it was clear that a thorough update was needed to encompass the challenge of interpreting the tens of thousands of variants not matching the reference human genome sequence in exome sequencing, and the millions of unmatched variants revealed in genome sequencing. The guideline produced (in collaboration with the Association for Molecular Pathology [AMP])⁷, representing hundreds of hours of deliberation and comment, is thorough, thoughtful, and serves to codify the decision-making that clinical genomicists were already doing in a gestalt sort of way to arrive at classification of variants according to a 5-tiered scale (Pathogenic, Likely Pathogenic,

Uncertain, Likely Benign, Benign). One can debate the relative significance weightings imparted to such factors as database matches, evolutionary conservation of the affected amino acid residue in the gene product, class of sequence change (missense, loss-of-function, splice junction alteration), degree of phenotypic match with the patient, etc.—but in general the scheme has proven quite useful and has been adopted widely. Still, there is a residual element of subjectivity that cannot be eliminated, as documented by discordant interpretations of the same variants by different expert laboratories all using the ACMG/AMP guideline.⁸ Many of these discrepancies were perhaps subtle, such as Pathogenic versus Likely Pathogenic, others potentially impactful on clinical management, such as Likely Pathogenic versus Likely Benign. But perhaps that is a good thing, as it demonstrates the need for input by seasoned laboratory directors and clinicians, who are not likely to be replaced (or *should* not) by a fully automated algorithm or artificial intelligence system anytime soon. Those of us who do this sort of work every day have learned that every patient is different, and any rules are made to be broken.

The term VUS has emerged as a particular source of annoyance, both for laboratories and the clinicians who receive the reports. It is both too broad and poorly understood. Many clinicians seem to interpret “VUS” as “Likely Benign” or at least “not actionable.” This may be reminiscent of our experience in the single-gene testing days, when most patients receiving a result of VUS in a *BRCA* sequencing report chose not to act on it—at least not with an irreversible intervention such as prophylactic mastectomy.⁹ But a VUS may well be pathogenic (or benign)—we just don’t have enough knowledge at the time of reporting to make that distinction. And a VUS may well be actionable, if only in directing a new line of ancillary testing (such as plasma amino acid quantitation if the gene in question is part of an amino acid metabolism pathway). At its core, the concept of VUS is a thoroughly rhetorical or artificial philosophical construct; it has no meaning to the cell or the organism, which presumably already knows (or will in the future) whether it will prove to be harmful or not. For that reason, it is of limited immediate practical use to the ordering clinician, aside from providing an impetus for orthogonal testing in some cases, or the potential future value of retaining it in the medical record (for it would otherwise be lost, barring future reanalysis of the entire exome or genome) to await new knowledge of its clinical implications or the further evolution of the patient’s phenotype with age. This is not to say that the term should be discarded—we do still need it—but it should be used sparingly in the reports issued and should always be accompanied by an explanation of the rationale for suspecting possible relevance to the patient’s phenotype, either now or in the future.

That clinical interpretation of VUS is of more than merely academic interest has been brought home in the real world in a most painful way, as the first (and to date only) lawsuit based on alleged misclassification of a variant. The case, *Williams et al. v. Quest Diagnostics*, concerns a boy with congenital epilepsy who died at age 3 in status epilepticus.

His seizures were refractory to all standard anticonvulsant medications. An NGS panel of epilepsy genes at a commercial laboratory (since purchased by Quest) revealed a variant in the *SCN1A* gene, Y413N, that was classified in the report as a VUS. *SCN1A* is the causative gene of Dravet syndrome, a severe childhood form of epilepsy that is typically not responsive to standard antiepileptic drugs but may be at least partially responsive to more specific and multitiered medications. Because the variant was called a VUS, these other drugs were never tried, and the boy’s mother sued the laboratory, claiming that if the variant had been classified at least as Likely Pathogenic, the other drugs would have been employed and her son might not have died (the variant is now considered Pathogenic). There are a number of legal technicalities in the case regarding statute of limitations, the timing of reclassification in relation to the issuing and transmittal of the report, etc. But the case highlights a more prevalent misunderstanding of the meaning of “VUS,” particularly the misconception that a VUS is not actionable, and what some have felt may represent underweighting of patient phenotype in the variant classification guidelines, a kind of “*a priori* risk” that would supersede the likelihood of a rare variant appearing in a clinically compatible gene simply by coincidence.¹⁰ This is one reason why our own clinical genomics operation at the University of California–Los Angeles (UCLA) routinely invites the ordering physicians to attend and participate in our weekly exome and genome interpretation conferences, since they best understand the patient’s phenotype¹¹ (see Table 1). Yet on the other hand, there is potential danger in “overcalling” variants, as they may be used for prenatal diagnosis in subsequent pregnancies of the family or otherwise push management in the wrong direction; a recent commentary in this journal warned that the *a priori* evidence noted here could actually be a result of simple ascertainment bias, and implored that VUS should in fact be “considered uncertain until proven guilty.”^{12,13}

In contrast to VUS, which were already familiar to laboratories performing all-exon sequencing of single genes or small panels, secondary findings are something new that only emerged in a big way when NGS was applied to nontargeted sequencing of the entire exome or genome. The only real secondary finding at risk in earlier genetic testing was revelation of false/misattributed paternity (when parents were tested along with the child by any technique capable of demonstrating transmission or nontransmission of a variant or polymorphism). Revelation of false paternity (or maternity) is seen much more often with NGS, since exome or genome sequencing of a trio (child, mother, father) is the ultimate parentage test, and it has led to further discussions of how such information should be handled and reported¹⁴ (keeping in mind that it will often compromise the informativeness of the test, such as determining whether or not a pathogenic variant in the proband is *de novo*). In addition, NGS testing even of the proband alone can reveal unexpected consanguinity and even incest by the observation

Table 1 Clinical exome/genome sequencing: lessons learned in the first six years

- Exome sequencing has better diagnostic yield than gene panels
- Communication with the ordering clinician is essential
- Diagnostic yield is quite high (assuming careful patient selection)
- Sequencing of trios (when parents are available) greatly aids interpretation and improves conclusiveness of the results
- Insurance coverage has been spotty, much as it was for chromosomal microarrays when they were introduced
- The pace of new disease gene discovery continues to be impressive
- Despite the earlier concerns about mandatory reporting of secondary/incidental findings, the vast majority of patients and parents decline the opportunity to “opt out”
- Reportable secondary/incidental findings (using the ACMG list) are seen in about 2% of cases
- The diagnostic yield of exome-negative cases reflexed to genome sequencing has been modest, mostly in the form of structural variants and cryptic splice sites that then need to be confirmed by RNA-Seq

of long runs of homozygosity (which then becomes a medicolegal or reportable child abuse incident).

In contrast to unexpected parentage discoveries, which can have adverse social and psychological outcomes, secondary findings discovered in certain familial cancer or cardiomyopathy genes can prove lethal. Others may not be so immediately threatening but could be information that the tested patient would wish to know so that preventive, surveillance, and/or life-planning interventions can be initiated. On the other hand, the secondary findings—by definition—is not related or even relevant to the indication for doing the genomic sequencing in the first place; rather, it was done for a much more urgently distressing concern of the patient or parents. They may be entirely focused on that condition (which itself may be immediately life-threatening) and not in any mood to hear about or understand an secondary finding conferring a variable (not 100%) risk of breast cancer 40 years from now. Could we be doing more harm than good by complicating the present stressful situation of these families? Or, as physicians, do we have a “duty to warn,” just as we do if we find something unexpected and worrisome during a routine physical exam?

That question, so difficult to answer or to reach consensus on, became the crux of a long and vigorous debate in the medical genetics community and the issuance of an ACMG guideline for reporting of incidental/secondary findings that proved quite controversial. A commentary in this journal by Jarvik and Evans¹⁵ bemoaned the continuing misuse of genomic nomenclature (such as “mutation” instead of “variant,” of which this author remains guilty), so it should come as no surprise that at the start, there was even widespread disagreement over what these findings should be called (Table 2). As is now well known, a specially appointed committee of the College deliberated intensively about which sorts of secondary findings, including which genes, disorders, and classes of variants, should or must be reported. The resulting guideline, published in this journal and widely cited,¹⁶ listed 24 high-penetrance, potentially lethal, and actionable disorders, caused by 56 genes, that must be sought out and reported in every clinical exome or genome sequencing test. Almost immediately, objections were raised

Table 2 Initial suggestions for nomenclature of incidental findings

- Incidental findings
- Secondary findings
- Unexpected findings
- Unanticipated findings
- Accidental findings
- Serendipitous findings
- Off-target findings
- Significant findings
- Insignificant findings
- Actionable findings
- Panic values

that the mandatory nature of the guideline, without an explicit opportunity for the patient (or parents) to opt out, overrode patient autonomy and transformed a diagnostic sequencing test into a screening test (for the 24 later-onset disorders).¹⁷ For these reasons and others, a modification of the guideline was later issued, describing the conditions for opt-out more explicitly, and more recently the number of reportable diseases and genes was also modified (four genes added, one removed).¹⁸ The debate has since settled down, though it is always simmering in the background. It raises profound philosophical questions that have no easy answers, and cogent arguments for either side have been made by well-meaning individuals. It is by far the thorniest ethical issue this author has ever confronted.

DIAGNOSIS IS NOT THE SAME AS SCREENING

As just described, one of the objections raised to the ACMG guideline on reporting of secondary findings was that mandatory (or even voluntary) reporting of secondary findings changes exome sequencing from a diagnostic test (aimed at identifying the genetic cause of the patient's symptoms) to a screening test (aimed at identifying predisposition to unrelated future illnesses). In addition to the discovery of an unexpected finding on routine physical exam, another analogy cited in support of the recommendation to

report a selected list of secondary findings was the incidental finding of a lung mass on a chest X-ray that had been ordered to rule out rib fracture. Indeed, radiology is the prior medical specialty with the most attention to incidental findings, nonreporting of which has resulted in lawsuits.¹⁹ But it has never been entirely clear that these analogies, drawn from far older, more routine and straightforward areas of medicine, are valid. Most of the secondary findings in the ACMG list of reportable findings constitute risk alleles of uncertain penetrance, even when we follow the recommendation to report only known pathogenic variants and not VUS. This is in contrast to incidental findings on chest X-ray or physical exam, which are by definition signs of present disease (though the nature and severity may not be known until further investigation).

For example, the penetrance of “pathogenic” variants for sudden cardiac death in the long-QT syndromes is likely no higher than 50%²⁰; is it really desirable to burden all patients who show up with a secondary finding in these genes with such a frightening prospect, considering that they did not pursue the sequencing test for anything related to this purpose? Another counterargument could be raised regarding disparities of access. If it is considered so important to report these findings when observed in a patient undergoing genomic sequencing for an unrelated phenotype, then why is the test not being offered as a risk screen to the entire population? Must one first get sick with an unknown syndromic condition to qualify for general genomic risk screening? The fact remains that there are no established guidelines promoting genetic screening for familial cancers, cardiomyopathies, malignant hyperthermia, tuberous sclerosis, and the other listed conditions in the general population, whether pediatric or adult. Some centers have attempted to market such “wellness screening” to healthy adults, but the uptake has not been sizable, in part because it is not considered standard practice and no health insurance plans currently will cover it. Meanwhile, pilot studies are underway to assess the feasibility and acceptance of genomic newborn screening.²¹ It seems probable that as sequencing costs continue to diminish, some form of newborn screening by NGS will become standard in the next 5–10 years, though it is unlikely to replace biochemical/tandem mass spectrometry screening anytime soon, at least for serious inborn errors of metabolism.

In fact, there is a screening test using NGS that is now well accepted and, in terms of sheer volume, constitutes the largest application of the technology in all of medicine: noninvasive prenatal screening (NIPS). The uptake of this new technology, as a supplement to maternal serum screening and an alternative to invasive amniocentesis or chorionic villus sampling (CVS), has been remarkable and has transformed the world of prenatal diagnosis. Obstetricians like it because it is far more predictive than maternal serum screening alone, and patients like it because it can be done earlier in pregnancy and at no risk to the fetus, as compared with amniocentesis. At present it is used primarily to screen for fetal aneuploidies;

initially it was used in high-risk pregnancies but has been more recently expanded as an option to all pregnant women. It is important to remember that it is only a *screening* test, and a positive result needs to be confirmed by subsequent amniocentesis; for that reason, NIPS is the preferred term for the procedure, replacing the earlier terms noninvasive prenatal *diagnosis* (NIPD) and noninvasive prenatal *testing* (NIPT). Concerns persist over the low level of false negatives,²² and there is even some question whether the universal implementation of NIPS has really resulted in fewer iatrogenic miscarriages, at least among the higher-risk pregnancies,²³ especially because it has decreased the training volume obstetric residents are now receiving in amniocentesis and CVS.²⁴ Overall, however, the innovation is seen as a benefit to most pregnant patients.

Another type of genetic screening using NGS has similarly found a home in prenatal clinics: expanded carrier screening (ECS). The ease of large-scale sequencing brought about by NGS has enabled prenatal carrier screening to expand dramatically beyond the handful of genes and diseases formerly offered to couples of specific ethnic backgrounds (Ashkenazi Jewish, African American, Mediterranean, Asian, etc.), along with cystic fibrosis and sometimes spinal muscular atrophy offered to all couples. However, this may represent yet another example of our technology exceeding our knowledge, for in the rush to add ever more and ever rarer disorders (some panels now contain almost 300—predominantly driven, it must be said, by the commercial sector), obstetricians and their patients and genetic counselors are being asked to face and render irreversible reproductive decisions about risks we know little about, based on ultrarare variants that are not well characterized, in disorders whose natural history is not well delineated, and including some that are likely asymptomatic.²⁵ As stated above in the context of secondary findings and NIPS, we should avoid conflating screening tests with diagnostic tests. Yet ECS, despite having “screening” in its name, will ultimately lead to a *diagnostic* test in the fetus, based on the same sequencing results and no additional biologic knowledge.

A CELL IS NOT AN ORGANISM (EXCEPT FOR PROTOZOA AND FUNGI)

Some of the uncertainty in our attempts to interpret the clinical significance of novel sequence variants—whether manually using the ACMG/AMP scheme or employing computer algorithms (PolyPhen, SIFT, etc., which have a tendency to reach mutually conflicting conclusions)—may owe to semantics: exactly what do we mean by the terms “pathogenic” and “benign”? For clinical utilization, it is assumed that we are referring to the potential of the given variant to cause a disease phenotype—or not—in the patient being tested. However, some of the elements of evidence we use to make that distinction are not necessarily or reliably tied to disease. A nonsense codon in the middle of a gene is certainly deleterious at the cellular level, in that a full-length protein product will not be produced, and any messenger

RNA (mRNA) transcribed will likely be destroyed by nonsense-mediated decay. But that does not always mean that it is disease-causing at the whole-organism (i.e., patient) level. There could be another gene or biochemical pathway that can substitute, even if only partially, for the damaged primary gene. We already know that the effects of many—perhaps most—genes can be influenced by so-called “modifier genes,” most of which remain to be identified. As one prominent example, the most common variant of *CFTR*, $\Delta F508$, typically causes severe, classic, pancreatic-insufficient cystic fibrosis; yet there are examples, albeit rare, of $\Delta F508$ homozygotes surviving well into adulthood with no discernible lung disease.²⁶ Certainly $\Delta F508$ would be deemed a “Pathogenic” variant using the ACMG/AMP scoring system, or simply by common knowledge, yet it cannot always be predicted to act as such. Likewise, we typically label as VUS any detected heterozygous variants in the genes of autosomal recessive disorders, even if they are known to be deleterious, in the absence of a “second hit” in the opposite allele (and assuming there was adequate sequencing coverage, detectability of large deletions, etc.). Yet there are well-known examples of double-heterozygous variants in two or more distinct genes resulting in a full-blown recessive phenotype (the case of *GJB2/GJB6* digenic heterozygous inheritance in congenital sensorineural hearing loss comes to mind²⁷). To adequately account for such phenomena across the genome would require so-called “network analysis,”²⁸ which is still in its infancy and not a component of routine clinical genomics.

For all these reasons, perhaps we should consider eliminating the absolute categories of Pathogenic and Benign, since we can never be absolutely sure, and use only the “Likely” categories. That would be consistent with the recognition of statistical probabilities we deal with throughout human genetics and genetic counseling and, as proposed below, almost certainly extend into the molecular and atomic realms. On the other hand, clinical microbiology and infectious disease specialists use terms like “(entero)pathogenic *E. coli*,” even though not everyone who ingests that organism will get sick. In that light, this suggestion would apply mostly to patients being tested in the asymptomatic or presymptomatic phase; it stands to reason that the finding of a deleterious variant in a gene that perfectly explains the patient’s current disease phenotype can be presumed “Pathogenic.”

THE “DARK MATTER” OF THE GENOME

As remarkably fruitful as exome sequencing has proven to be in ending the diagnostic odyssey, every genomic laboratory and every clinical geneticist using these services has experienced many cases in which the phenotype of the patient was so convincingly syndromic—sometimes even with a similarly affected parent or sibling—that the resulting negative exome (or genome) result came as a shock. In some sense these mysterious disappointments should not have been unexpected. As noted earlier, exome and genome sequencing interrogate only the first, linear dimension of the genome; we now know there are many higher orders of structure and

regulation, and undoubtedly countless more that we cannot yet even imagine. Just as astrophysicists cannot explain the structure of galaxies and the cosmos without invoking a yet-undetected dark matter, there must be components of genome action that we are just not seeing.

We should have been sufficiently humbled in the 1970s when eukaryotic genes were found to be interrupted by long stretches of noncoding sequence, dubbed “introns.” Why and how did such a thing evolve? Whether by the action of transposons, the duplication of useful sequences specifying multifunctional domains of different proteins, insertion of viral sequences, or random breakage, they did happen and are evidently here to stay, at least for the predicted future existence of *Homo sapiens*. And with them come splice-site variants, alternative transcripts, exon skipping, and other phenomena that may render an otherwise inert synonymous variant pathogenic or rescue the deleterious effect of a pathogenic variant. Taking advantage of these aspects of molecular physiology has yielded targets for a new kind of gene therapy, different from traditional gene replacement, but they also raise the level of complexity of sequence variant interpretation. Add to that the other influences on gene expression that existing sequencing technology cannot adequately address—micro- and other noncoding RNAs; promoters, enhancers, and other noncoding DNAs; alternative RNA splice variants beyond the canonical transcript;²⁹ modifier genes; epigenetic modification of nucleotides and DNA-binding proteins; relative abundance of transfer RNA (tRNA) isoforms specific for particular codons (another reason why synonymous nucleotide substitutions can be deleterious³⁰); topological conformation of chromosomes in the cell that bring unlinked genes and regulatory regions into three-dimensional proximity;³¹ post-transcriptional editing of RNA;³² interspecies exchange of metabolites and signaling molecules with our microbiome; dietary and environmental effects on gene expression—and it should not be at all surprising that in many cases the one-dimensional nucleotide sequence by itself is not fully informative or even interpretable at all. Some of these mysteries are starting to be solved through newer techniques like GS and RNA-Seq, but many of them are likely to remain beyond our reach with current and near-future technology.

TUMORS ARE SOMATIC MOSAICS AND SO ARE WE

Along with medical genetics and obstetrics, there is another locus of vigorous clinical application of NGS: the sequencing of tumor DNA in oncology and surgical pathology. As in expanded carrier screening, the technology has enabled the expansion of interrogated genes beyond the few formerly queried one at a time (*KRAS*, *EGFR*, *FLT3*, etc.); many laboratories now routinely sequence from 50 to several hundred genes. Even at those numbers, the testing is far more targeted than the genetic applications discussed so far. In many cases, tumor DNA sequencing is not needed for diagnosis, because the type of malignancy is already known from histopathologic and other analysis. Rather, the

sequencing is intended to identify variants in oncogenes and cell signaling pathways that are accepted targets for the new classes of biologic drugs. Unfortunately, the results, though used for clinical management, are often no more predictive of drug response than the low-penetrance disease-risk alleles we detect in germline exome sequencing. Still, it is certainly a move in the right direction, away from the older, nonspecific and toxic chemotherapeutic agents. And it represents a legitimate form of “personalized medicine” (a term some find objectionable), just as do the mutation-specific drugs now being used for cystic fibrosis and Duchenne muscular dystrophy.^{33,34}

Almost since the birth of molecular diagnostics, there has been a kind of feigned dividing line between testing for germline/inherited and somatic/acquired variants, the latter referring essentially to variants in malignant tumors and leukemias. This dichotomy was in part intellectual but also political, involving turf and scope of practice specific to various specialties. With the enhanced power and depth of NGS, we are now coming to appreciate that this boundary is indeed rather artificial. Exome sequencing of a tumor will reveal inherited variants along with the acquired ones, and NIPS has detected aneuploidies in the maternal blood that would be incompatible with fetal life and in fact derived from an occult maternal malignancy.³⁵ Moreover, the molecular heterogeneity of tumors, with multiple variants at low allele frequency that present such a challenge in analysis, is now coming to be mirrored on the germline side, as low levels of somatic mosaicism seem to be characteristic of the human body, and not just in those patients with particular genetic diseases like Proteus syndrome that are themselves often mosaic—raising new questions about the threshold of variant allele frequency necessary to define a genetic disease.³⁶ It may be an opportune time, given the wide adoption of NGS in both areas, for this artificial division to go away.

FROM MOLECULAR GENETICS TO ATOMIC GENETICS AND BEYOND

The selection of the quote from Erwin Schrödinger at the top of this article was deliberate and meaningful on several levels. The famous quantum physicist was the first to turn his attention to speculation on the molecular basis of life and, more pointedly, heredity. He deduced that chromosomes must be composed of a complex, linear molecule of some kind, decades before the discoveries of Watson and Crick. He was followed by many other prominent nuclear physicists, some of whom had worked on the Manhattan Project, transitioning “from the science of death to the science of life”: Delbrück, Szilard, Bragg, Wilkins, Benzer, and others; Crick himself started out in physics. It makes sense beyond just personal interest; if Schrödinger was correct that heredity is based on an invisible molecule, then who better to characterize it than experts on the smallest particles in the universe?

Of course, DNA is not truly invisible, but it is so small that none of our sophisticated techniques allow us to visualize it in

its natural state. To do that would require entering the cell nucleus and assuming the size of a molecule or atom ourselves, as in the book and movie by Richard Matheson, *The Incredible Shrinking Man* (1957). Short of that impossibility, we should admit that *all* of our representations of DNA and the genome are inadequate. From Rosalind Franklin’s famous “photograph 51” showing an X-ray diffraction pattern suggestive of a helix, to Watson and Crick’s rod-and-ball model derived from it, to autoradiographic exposures on a Southern blot, to fluorescent bands on a polyacrylamide sequencing gel, to peaks on capillary electrophoresis of a Sanger sequencing reaction, to the cluster arrays of massively parallel DNA sequencers, to the largest databases of catalogued sequence variants—all of these are just representations limited to our current technologies and understanding, and in fact provide only limited insight into what is actually going on at the molecular, cellular, and organismal levels. Like Franklin’s photograph taken in 1952, they are merely snapshots of our best guesses at a certain time-point in scientific progress.

The actual, “living” DNA molecule, if we could apprehend it on its own terms, would look far different from all of these artificial representations, in fact quite alien. The atoms would appear as quantum electron clouds, indiscernible at the individual particle level. The all-important atomic nuclei, so essential to the elemental makeup of the molecule, would be all but invisible even at the size of our shrunken man, with a diameter 60,000 times smaller than the cloud itself. We all know that Einsteinian relativity operates noticeably only at the scale of the very large, and quantum mechanics becomes significant only at the very small, while classical Newtonian physics seems to work just fine at the size range we deal with in our daily lives, including that of the human body. But at the molecular level of DNA and the atomic level of its components and interacting metabolites, might quantum phenomena play a role?

To greatly oversimplify the field, quantum physics introduces uncertainty into all things; no individual particle or behavior can be described or predicted; one can only posit a statistical probability based on very large numbers of particles. Such a concept should resonate with geneticists, because our field uses statistical probability to arrive at, for example, empirical recurrence risks in genetic counseling for a disorder whose molecular causation is not known. Can there be such a thing as quantum genetics, or more globally, quantum biology, even though the size of our bodies falls within the “sweet spot” of Newtonian physics? As it happens, quantum phenomena are just now being incriminated in several biological processes, most convincingly in some steps of photosynthesis in plants, but also in the ability of migrating birds and insects to sense the direction of the earth’s magnetic field, and in certain aspects of enzyme-substrate reactions.³⁷ Is it conceivable that such quantum paradoxes as uncertainty, entanglement, decoherence, superposition, and tunneling might be active in human genetics? Could quantum mechanics help to explain at least some of the “dark matter”

of the genome: why patients who almost certainly have a genetic condition are negative on genomic sequencing, or why patients with a documented pathogenic variant that causes disease in almost everyone else who carries it remain healthy?

Perhaps we are now on the cusp of entering the quantum realm in clinical genomics and all the other classes of “-omics.” If so, we will have come full circle, back to the physicists who founded molecular biology in its earliest days. Whether this transition from molecular to atomic biology and atomic genetics occurs within the lifetime of anyone presently celebrating *GIM*'s 20th anniversary is anyone's guess. But if and when it comes, it would seem that medical geneticists would be the most apt practitioners to adapt to it, since we are the most facile and comfortable of any medical specialists in dealing with probabilities and uncertainty. We should not be cowed by such a proposition, but welcome and embrace it, for it will ensure that our specialty remains vital, exciting, and ever-changing. After all, it would not be much fun if we could already predict the contents of the 40th anniversary issues of *Genetics in Medicine*.

ACKNOWLEDGMENTS

The author thanks the other faculty, staff and patients of the UCLA Clinical Genomics Center for their boundless knowledge, empathy and inspiration.

DISCLOSURE

W.W.G. declares no conflicts of interest.

NOTE

Some of the concepts put forth in this article were first presented in the author's Presidential Address at the 2012 ACMG Annual Meeting.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- ACMG Board of Directors. Clinical utility of genetic and genomic services: a position statement of the American College of Medical Genetics and Genomics. *Genet Med*. 2015;17:505–507.
- Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005;437:376–380.
- Bentley DR, Balasubramanian S, Swerdlow HP, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 2008;456:53–59.
- Rothberg JM, Hinz W, Rearick TM, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. 2011;475:348–352.
- Gahl WA, Mulvihill JJ, Toro C, et al. The NIH Undiagnosed Diseases Program and Network: applications to modern medicine. *Mol Genet Metab*. 2016;117:393–400.
- Richards CS, Bale S, Bellissimo DB, et al. ACMG recommendations for standards for interpretation and reporting of sequence variations: revisions 2007. *Genet Med*. 2008;10:294–300.
- Richards S, Aziz N, Bale S, et al. 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;17:405–424.
- Amendola LM, Jarvik GP, Leo MC, et al. Performance of ACMG-AMP variant-interpretation guidelines among nine laboratories in the Clinical Sequencing Exploratory Research Consortium. *Am J Hum Genet*. 2016;98:1067–1076.
- Murray ML, Cerrato F, Bennett RL, Jarvik GP. Follow-up of carriers of BRCA1 and BRCA2 variants of unknown significance: variant reclassification and surgical decisions. *Genet Med*. 2011;13:998–1005.
- Cederbaum S. Interpreting sequence variants in a clinical context. (Letter). *Genet Med*. 2015;17:1012.
- Lee H, Deignan JL, Dorrani N, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA*. 2014;312:1880–1887.
- Weck KE. Interpretation of genomic sequencing: variants should be considered uncertain until proven guilty. *Genet Med*. 2018;20:291.
- Evans JP. Variant classification: high stakes and hubris. *Genet Med*. 2018;20:283.
- Eno C, Bayrak-Toydemir P, Bean L et al. Misattributed parentage as an unanticipated finding during exome/genome sequencing: current clinical laboratory practices and an opportunity for standardization. *Genet Med*. 2018 Sep 14; <https://doi.org/10.1038/s41436-018-0265-4> [Epub ahead of print].
- Jarvik GP, Evans JP. Mastering genomic terminology. *Genet Med*. 2017;19:491–492.
- Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med*. 2013;15:565–574.
- Allyse M, Michie M. Not-so-incidental findings: the ACMG recommendations on the reporting of incidental findings in clinical whole genome and whole exome sequencing. *Trends Biotechnol*. 2013;31:439–441.
- Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SFv2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med*. 2017;19:249–255.
- Berlin L. The incidentaloma: a medicolegal dilemma. *Radiol Clin N Am*. 2011;49:245–255.
- Priori SG, Napolitano C, Schwartz PJ. Low penetrance in the long-QT syndrome: clinical impact. *Circulation*. 1999;99:529–533.
- Berg JS, Agrawal PB, Bailey DB. Newborn sequencing in genomic medicine and public health. *Pediatrics*. 2017;139:e20162252.
- Zhang H, Gao Y, Jiang F, et al. Non-invasive prenatal testing for trisomies 21, 18 and 13: clinical experience from 146,958 pregnancies. *Ultrasound Obstet Gynecol*. 2015;45:530–538.
- Malan V, Bussieres L, Winer N, et al. Effect of cell-free DNA screening vs direct invasive diagnosis on miscarriage rates in women with pregnancies at high risk of trisomy 21: a randomized clinical trial. *JAMA*. 2018;320:557–565.
- Gupta N, Dragovic K, Trester R, Blankstein J. The changing scenario of obstetrics and gynecology residency training. *J Grad Med Educ*. 2015;7:401–406.
- Grody WW. Where to draw the boundaries for prenatal carrier screening. *JAMA*. 2016;316:717–719.
- Burke W, Aitken JL, Chen S-H, Scott DR. Variable severity of pulmonary disease in adults with identical cystic fibrosis mutations. *Chest*. 1992;102:506.
- Cama E, Melchionda S, Palladino T, et al. Hearing loss features in GJB2 biallelic mutations and GJB2/GJB6 digenic inheritance in a large Italian cohort. *Int J Audiol*. 2009;48:12–17.
- Ha MJ, Veerabhadran B, Do K-A. DINGO: differential network analysis in genomics. *Bioinformatics*. 2015;31:3413–3420.
- Park E, Pan Z, Zhang Z, Lin L, Xing Y. The expanding landscape of alternative splicing variation in human populations. *Am J Hum Genet*. 2018;102:11–26.
- Worthington R, Ball E, Wolf B, Takacs G. Method to identify silent codon mutations that may alter peptide elongation kinetics and co-translational protein folding. *Methods Mol Biol*. 2017;1647:237–243.
- Oudelaar AM, Hanss LP, Hardison RC, Kassouf JT, Hughes JR, Higgs DR. Between form and function: the complexity of genome folding. *Hum Mol Genet*. 2017;26:R208–2515.
- Maas S. Gene regulation through RNA editing. *Discov Med*. 2010;10:379–386.
- Wainwright CE, Elborn JS, Ramsey BW, et al. Lumacaftor-ivacaftor in patients with cystic fibrosis homozygous for Phe508del CFTR. *N Engl J Med*. 2015;373:220–221.

34. Echevarría L, Aupy P, Goyenvallé A. Exon-skipping advances for Duchenne muscular dystrophy. *Hum Mol Genet.* 2018;27:R163–72.
35. Bianchi DW, Chudova D, Sehnert AJ, et al. Noninvasive prenatal testing and incidental detection of occult maternal malignancies. *JAMA.* 2015;314:162–169.
36. Biesecker LG. Mosaic disorders and the taxonomy of human disease. *Genet Med.* 2018;20:800–801.
37. Brookes JC. Quantum effects in biology: golden rule in enzymes, olfaction, photosynthesis and magnetodetection. *Proc Math Phys Eng Sci.* 2017;473:20160822.