Published in final edited form as:

Am J Med Genet A. 2011 December ; 155A(12): 2916–2924. doi:10.1002/ajmg.a.34357.

Genomics Really Gets Personal: How Exome and Whole Genome Sequencing Challenge the Ethical Framework of Human Genetics Research

Holly K. Tabor^{1,2,*}, Benjamin E. Berkman^{3,4}, Sara Chandros Hull^{3,4}, and Michael J. Bamshad^{1,5,**}

¹Department of Pediatrics, University of Washington, Seattle, Washington ²Treuman Katz Center for Pediatric Bioethics, Seattle Children's Hospital, Seattle, Washington ³Department of Bioethics, Clinical Center, National Institutes of Health, Bethesda, Maryland ⁴Office of the Clinical Director, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland ⁵Department of Genome Sciences, University of Washington, Seattle, Washington

Abstract

Exome sequencing (ES) and whole genome sequencing (WGS) putatively identify all adverse functional alleles of protein-coding genes. Accordingly, while ES/WGS are transformative new tools for gene discovery in human and medical genetics research, they also generate new manifestations of ethical issues related to the consent process, data sharing, and return of results. These manifestations have yet to be comprehensively framed, due in part to the rapidity with which new technologies for ES/WGS are being applied and because of a lack of empirical data to provide guidance. Accordingly, researchers, funding agencies, and policy makers have largely dealt with these issues intuitively. We explain how use of ES/WGS challenges: (i) models under which informed consent is typically obtained; (ii) how harms associated with data sharing are considered; and (iii) the nature of obligations surrounding unanticipated findings. We provide broad guidance about interim ways to contend with these issues and make broad recommendations for areas for novel resource and policy development.

Keywords

whole genome sequencing; exome sequencing; ethics; informed consent; return of results; information dissemination; privacy; data sharing; confidentiality

*Correspondence to: Holly K. Tabor, Ph.D., Treuman Katz Center for Pediatric Bioethics, Seattle Children's Hospital, 1900 Ninth Ave, M/S C9S-6, Seattle, WA 98101. holly.tabor@seattlechildrens.org. **Correspondence to: Michael J. Bamshad, M.D., Department of Pediatrics, University of Washington School of Medicine, Box 356320, 1959 NE Pacific Street, Seattle, WA 98119. mbamshad@u.washington.edu.

The opinions expressed in this article are those of the authors. No statement in this article should be construed as an official position of the National Human Genome Research Institute, National Institutes of Health, or Department of Health and Human Services

INTRODUCTION

Exome sequencing (ES) and whole genome sequencing (WGS) are powerful new tools for gene discovery in human and medical genetics research. To date, the exomes of several hundred and the genomes of several dozen unrelated individuals and families have been reported, primarily as a consequence of testing different sequencing platforms, computational approaches, and analytical methods [Levy et al., 2007; Wang et al., 2008; Wheeler et al., 2008; Ashley et al., 2010; Li et al., 2010; Pelak et al., 2010; Roach et al., 2010; Schuster et al., 2010]. However, with increasing frequency, ES and WGS have been used to identify causal variants for rare, monogenic syndromes including Miller (POADS [OMIM 263750]) [Ng et al., 2010b], Schinzel-Giedion (SGS [OMIM 269150]) [Hoischen et al., 2010], Charcot-Marie-Tooth (CMT4C [OMIM 601596]) [Lupski et al., 2010], and Kabuki (KMS [OMIM 147920]) [Ng et al., 2010a] among others [Gilissen et al., 2010], and to solve diagnostic dilemmas [Rios et al., 2010; Worthey et al., 2011]. Both ES and WGS are now being applied to search for variants that underlie a wide range of common, complex phenotypes (e.g., autism, diabetes) in large-scale projects funded by the National Institutes of Health [e.g., NHLBI Exome Sequencing Project (ESP)] and the Wellcome Trust, and multiple efforts are underway to create repositories of exome/genome data linked to phenotypic traits [e.g., database of genotypes and phenotypes (dbGaP)] [Karow, 2009].

The application of new sequencing technologies in these and other studies represents a departure from the status quo in human genetic research because of the substantially larger amount of important information generated. Whereas conventional technological approaches might generate data on hundreds of thousands, or even millions of polymorphisms, the overwhelming majority of these variants are located in non-coding regions and likely not of functional significance themselves. In contrast, both ES and WGS provide information on virtually all functional, protein-coding variants in the genome for each individual participant. This includes most variants known to influence risk of human diseases and traits.

We describe how ES/WGS studies, because of the breadth and depth of coding variant data produced, challenge the “standard” ethical framework used by genetics researchers. We focus specifically on three key parts of this framework: the informed consent process, data sharing, and management of individual research results. We appreciate that there is a wide range of opinions about how to contend with informed consent, data sharing, and return of research results in ES/WGS studies. Our goal is to consider these issues in a new way so as to highlight the opportunities for the research, ethics and policy communities to consider and resolve them. We think this is particularly important because ES/WGS studies are rapidly gaining in popularity. Moreover, it is not our intent to review or critique ethical practices of existing or published ES/WGS studies or to make specific policy recommendations for ES/WGS studies.

INFORMED CONSENT

There are several issues that should be considered in making decisions about if and to what extent consent documents need to be specific for use in ES/WGS studies. These issues include: (i) whether the aims of ES/WGS research are different from more conventional

genetic research; (ii) if the risks and benefits of ES/WGS are unique and therefore require explanation in order for participants to make informed decisions about participation; and (iii) whether participants should be told specifically about ES/WGS during informed consent in order to maintain transparency and trust in the research enterprise.

First, anecdotal reports have suggested ES/WGS approaches need to be explicitly described and explained in informed consent documents because they are so different technically from targeted genetic studies. In general, it is not common practice to detail the specific approach or methodology (e.g., specific genotyping method, linkage, GWAS, targeted sequencing) used in a research study in consent documents because such details are not relevant to the goals of the research or to the benefits and risks that the participant should consider in order to make an informed decision about participation. We think a similar standard is reasonable for ES/WGS studies as they typically share the same goals as targeted genetic studies—to identify the gene(s) and variant(s) that confer risk for a specific disease or trait. In situations in which the goals of a study are substantially different, this should be reflected in an informed consent document.

Second, researchers should consider whether the risks and benefits of ES/WGS are different enough from targeted genetic studies to require specific mention and explanation in order for people to make informed decisions about participation. Compared to traditional targeted gene discovery strategies, ES/WGS approaches have two different potential risks. First, ES/WGS studies have a higher likelihood of discovering unanticipated results with clinically utility, both related and unrelated to the primary traits/diseases under study. This could be perceived both as a risk and as a benefit to participants. Optimally, researchers should decide, at the beginning of a study when possible, if they plan to offer return of any results to participants. Arguably, if researchers do not plan to return results to participants (e.g., studies using anonymized samples), it is not necessary to describe potential risks or benefits related to results identification or return in the informed consent document. However, if researchers plan to offer return of results, or think that they may consider results return in the future, the availability of such results might influence an individual's decision about participation. In this case, informed consent documents for ES/WGS should describe the scope of potential results identified, and proposed plans for return of results, including how participants will be notified. Another risk of ES/WGS that might be different than conventional genetic studies arises because the risk of identifiability and potential loss of privacy or confidentiality is compounded by the sharing of data on rare alleles, including those that are of clinical or personal utility. The magnitude of this risk and the likelihood of harms related to it are not known, but participants might desire to know about plans for data sharing either through informed consent or through notification from investigators.

Third, researchers should consider whether participants should be told specifically about ES/WGS during informed consent in order to maintain transparency and trust in the research enterprise. While researchers may not have a specific regulatory obligation to mention ES/WGS in informed consent, participants who are not aware of the scope of potential results identified, the possible return of these results, or plans for data sharing, may be surprised and even feel misled if they think that their samples were used for a different purpose without their knowledge.

If informed consent documents describe ES/WGS specifically, it is unclear how detailed an explanation is needed, how different types of genetic variation should be presented, and to what extent the implications of possible results for health-related traits (e.g., risk of disease) should be discussed. There are few empirical data about what language or approaches are most effective for explaining such complex information to research participants. Even with carefully designed consent documents, it is likely that there will be unique barriers to comprehending complicated information about ES/WGS research. More innovative and dynamic approaches to presenting information (e.g., web-based video vignettes about testing) and clearly framing participant expectations (e.g., user-selected preferences for return of results) both at the time consent is obtained and for the duration of the study (i.e., models of consent that allow participants to interact with information related to the study over time) might be helpful.

Several studies have demonstrated that individuals, when presented with hypothetical scenarios about research in general and genetic research specifically, believe that researchers will tell them about information that is important to their health, even if no such commitment is indicated on the consent documents or if the consent documents explicitly stated that no results would be returned [Miller et al., 2008; Murphy et al., 2008]. The increased likelihood that ES/WGS will identify variants that are of clinical and/or personal utility to study participants may raise their expectations. This will require the development of consent documents and processes that explain more clearly the options and plan(s) for return of results, and perhaps more importantly, the limits of interpretation.

Given the complexities of study designs, governance structures, and IRBs, a “one-size fits all” approach to the informed consent process for prospective ES/WGS studies is likely to be both burdensome and ineffective. Increasingly, participants will be asked to consent explicitly for ES/WGS, providing the opportunity to use and collect empirical data about new approaches to consent in different research populations and contexts. Such data should, in turn, also be used as the foundation for OHRP and funding agencies to provide specific guidance and policies that will help researchers and IRBs effectively deal with issues of informed consent for ES/WGS studies.

Many ES/WGS studies use DNA samples linked to rich phenotypic data that have been collected over the past several decades. The original consent documents under which these samples were collected typically explained genetic research broadly, describing the use of approaches based on linkage, genome-wide association (GWA), and/or targeted sequencing. However, these documents are highly variable in their content [Hull et al., 2004], and almost never described ES/WGS, nor possible risks and benefits associated with strategies designed to identify potentially deleterious alleles in all protein-coding genes, much less the entire genome.

There is a broad spectrum of ways to contend with samples such as these that are exceptionally valuable but that were consented in the pre-ES/WGS era. At one end of the spectrum is a pragmatic approach. For some studies, the goals, risks and benefits of ES/WGS may be similar to those described in the original consent documents. Thus, re-consent may not be necessary. Alternatively, it might be appropriate when possible to notify

participants of the intent to use of ES/WGS approaches without obtaining re-consent, providing participants with information and the opportunity to contact researchers and withdraw if they wish to do so. It is unknown whether participants in existing studies would be concerned to learn that their samples are being used for ES/WGS, or whether this would contribute to a lack of trust or other possible harms, and further empirical research is needed on this topic. It is worth noting that re-consent will not be possible for studies using de-identified or anonymous samples.

A more conservative approach would be to obtain re-consent from participants, even if it is not required, in the spirit of transparency and in order to build and maintain trust. While this approach is likely to be both burdensome to investigators and logistically challenging, it might be appropriate in studies with an active governance structure—in which participants have higher expectations in terms of how and when they are informed about new studies. Similarly, this approach might be appropriate for researchers who have long-term longitudinal clinical relationships with participants. At a minimum, researchers should consider possibly notifying participants of the use of samples for ES/WGS, as this would allow for transparency, and would reduce the chance of surprise if results are returned in the future. We think that for many existing studies, a notification approach is acceptable and sufficient, although this will depend on the study context [Hull et al., 2008; Pulley et al., 2008].

Empirical data about participant responses to both re-consent and notification across different study and population contexts would provide both investigators and regulatory bodies [e.g., institutional review boards (IRBs)] with some needed guidance. In the meantime, researchers, together with IRBs, will need to consider these issues and possible solutions on a case-by-case basis, taking context as well as resources for re-consent into account (Table I).

DATA SHARING

The last few years have witnessed a shift toward increased sharing of human genetic and phenotypic data as a means to facilitate and accelerate genomics research. For example, investigators who receive funding from the NIH for GWAS are required to deposit genotype and phenotype data for each individual participant in dbGaP, a restricted-access data repository supported and managed by NIH [Mailman et al., 2007]. If this is not feasible because of limitations of the original consent and/or an inability to get re-consent from participants for this purpose, investigators are obliged to apply for an exception from NIH.

This approach to broad data sharing is an extension of both the Bermuda Principles adopted by the Human Genome Project [Summary of principles agreed at the International Strategy Meeting on Human Genome Sequencing, 1996], and an extension of NIH data sharing policies for federally funded research. Recently, there have been discussions about developing an even broader data sharing requirement that would include DNA sequence and phenotypic data from ES/WGS studies.

Sharing of ES data via dbGaP has been a requirement for several recently funded American Recovery and Reinvestment Act (ARRA) exome sequencing studies (e.g., NHLBI Exome

Sequencing Project), and several thousand exomes from these projects will be deposited into dbGaP, together with phenotypic information. Sequence data from the 1000 Genomes Project is also available for unrestricted public use (<http://www.1000genomes.org/data-#DataAccess>). We view such efforts positively and think that further data sharing and transparency will lead to fuller utilization of public resources and faster scientific advancements. However, it is important to explore whether the risks, benefits, and potential harms of sharing ES/WGS data, even in a restricted access database like dbGaP, differ from those associated with sharing of more conventional genetic data.

Compared to GWAS, ES/WGS will be used in a wider range of study designs such that individuals, families, and populations will be sequenced to discover genes for both rare monogenic disorders and common, complex diseases. Accordingly, it will be important to assess the specific risk of identifying study participants from ES/WGS data. Inferring the identity of a study participant could facilitate linking them to other genetic or phenotypic data that might be stigmatizing or discriminatory, regardless of whether or not they have a disease or carry a disease-causing variant. Overall concern has been heightened by a recent study that demonstrated that it is possible under certain circumstances to use aggregate frequency data from GWAS studies to determine whether an individual is a member of a study cohort [Homer et al., 2008], although this requires access to genotype data from a given individual for comparison. Both ES and WGS data contain rare as well as common variants, and such rare variants may enhance identifiability as well as provide more interpretable information about disease risk. Therefore, both the absolute risk of identifiability from ES/WGS data and the relative risk compared to GWAS data could be higher. These risks should be formally assessed, as such information will be essential to the development of robust data sharing policies and protections.

Concern about the absolute risk of identifiability from ES/WGS data is important, but it can obscure the need for a more careful analysis and characterization of the actual possible harms associated with identifiability and their likelihood. Bioethics research and consent documents frequently focus on identifiability itself as a harm, rather than on the subsequent harms themselves. Ideally, informed consent should incorporate: (i) best estimates, based on empirical data, of the risk of identification from ES/WGS data; (ii) accurate descriptions of possible harms related to identification, such as discrimination, stigmatization, or loss of privacy; and (iii) best estimates of the likelihood of those harms. The informed consent process should also address participants' concerns about identifiability, and balance these concerns with the potential for direct and indirect benefits of research participation. More complete information about possible harms could also be used to improve and tailor informatics and technological strategies, as well as regulatory and legal protections, to minimize the harms themselves, rather than minimizing identifiability per se.

Another potential risk of harm from data sharing of ES/WGS data is that of reducing researcher-participant trust by not clearly communicating the plans for data sharing. Several studies have determined that research participants across a range of populations and disease groups want to be informed about plans for data sharing. One study found that while participants were generally in favor of data sharing, they would feel deceived or angry if they found out that their data were shared without their knowledge of consent [McGuire et

al., 2008a] Another study found that while participants generally supported data sharing as beneficial and that it would not deter them from participation, 90% felt that it was important for researchers to ask their permission for data sharing, and that 70% felt that notification of dbGaP submission after the fact would be unacceptable if their data were shared without notification or permission [Trinidad et al., 2010; Trinidad et al., 2011] These studies suggest that in any genetic studies that require broad genetic data sharing in dbGaP that researchers should consider the trust relationship with participants and mechanisms for providing transparency about how genetic data are being shared, and for what purposes they are being used [Trinidad et al., 2011].

As data sharing for ES/WGS studies becomes more commonplace, it will be important to strike a balance between the necessity of data sharing and respect for individuals/populations that might oppose broad data sharing (e.g., Native Americans) [Levenson, 2010; Mello and Wolf, 2010] or make exceptions for samples from individuals who cannot be re-contacted to obtain their re-consent. In other words, in some cases the benefits of gene discovery via ES/WGS will outweigh the merits of data sharing, especially in diverse and underrepresented populations.

MANAGEMENT OF INDIVIDUAL RESEARCH RESULTS

Over the last 10 years or so, researchers and policy makers have struggled to develop a framework and guidelines for the analysis and return of results from genetic studies. In general, there has been a lack of consensus on what kinds of results to return and the scope of researcher obligations to participants [Bookman et al., 2006; Ravitsky and Wilfond, 2006; Dressler, 2009; NBAC, 1999; Fabsitz et al., 2010]. While this discussion continues at an academic and a policy level, a variety of “return of results” practices have emerged, in part represented through existing institutional (e.g., NHLBI) guidelines [Bookman et al., 2006; Fabsitz et al., 2010] and through the de facto policies of IRBs. These approaches generally minimize the need to return results unless they are identified in the course of routine research analysis, have been validated, and are determined to be of clinically utility and actionable [Bookman et al., 2006; Fabsitz et al., 2010]. The details of how “clinically utility” and “actionable” are defined, and by whom, remain under dispute, and each investigator and institution has generally approached this definition on a case-by-case basis.

ES/WGS research challenges at least three key assumptions of the existing return of results paradigm. The first assumption is that very few, if any, results of clinical utility (i.e., actionable results) will be identified in the course of genetic research. This assumption is rooted in the nature of existing technological approaches, which either focused narrowly on a few genes, as in targeted resequencing, or broadly on non-functional variants throughout the genome that might be linked to functional changes. In contrast, with ES/WGS results that have or might have clinical utility will be found at some loci in virtually all participants, including known variants in genes for autosomal dominant diseases and traits, carrier status for autosomal recessive diseases and traits, and known genetic risk factors for complex diseases and pharmacogenomics response, [Roach et al., 2010; Schuster et al., 2010]. Therefore, it is no longer a question of whether or not results with clinical utility will be found in any research participant by ES/WGS, but rather *how many* such results will be

identified in each participant. This shifts the debate away from a focus on whether meaningful results will be found, and instead focuses it on understanding what obligations, if any, researchers have to evaluate, return, and explain different kinds of results in a range of research contexts.

The second assumption challenged by ES/WGS is that a distinction exists between so-called “incidental findings” (IFs) and findings that are explicitly related to the original hypotheses (i.e., primary findings). A great deal of attention has been paid to the possibility that researchers might identify a genetic result that was not related to the primary phenotype under study but might still be considered clinically important and actionable [Cho, 2008; Wolf et al., 2008a,b; Beskow and Burke, 2010]. Possible IFs in the pre-ES/WGS era included chromosomal rearrangements, sex chromosome aneuploidy, and variants in candidate genes associated with other traits or diseases. Under the existing return of results paradigm, researchers have not necessarily been obligated to look systematically or intentionally for such findings in their data because they were considered “incidental” to the content and purpose of the original research [Wolf et al., 2008a,b]. If IFs were identified in the course of the research, investigators could consider whether they should be returned, in part based on whether they were actionable or related to a “severe” phenotype [Cho, 2008; Wolf et al., 2008a,b]. For experimental approaches based on ES/WGS, the distinction between primary and IFs becomes arbitrary, since a major aim of using these approaches is to identify all variants in an exome/genome. Accordingly, genetic variants may be found that influence phenotypes that are incidental to those of the main study aims, but the variants themselves are not incidental to the experimental strategy. Thus, in the context of ES/WGS studies, such results might be more appropriately considered “unanticipated results,” because they are not related to the primary research focus.

ES/WGS research challenges what we call the “don’t look, don’t tell” assumption. In the past, researchers did not have to look for all possible results (i.e., results which would require them to do analyses beyond what was practical). For example, a recent consensus paper on management of IFs stated, “Researchers generally have no obligation to act as clinicians and affirmatively search for IFs,” and went on to suggest that it was unrealistic to expect them to do so because of the complexity of analysis and expertise required [Wolf et al., 2008a]. With ES/WGS, the act of “looking” for all possible results becomes much more practical and indeed annotating all identified variants is a fundamental component of the analytical approach. Accordingly, it is reasonable to suggest that any investigator using ES/WGS not only would be able to look for results of clinical utility, but would do so as part of the experimental design. In this way, the nature of the inquiry, and the relative “ease” of analysis argue that there is no longer an inherent practical barrier to “looking” for functional variants that might have clinical utility.

The interpretation of annotated variants found by ES/WGS, and their translation toward clinical care will not necessarily be easy or routine. Researchers will be especially challenged by the evaluation and management of novel variants identified using ES/WGS. In each person, several hundred novel protein-coding variants are likely to be found, the majority of which will occur in one individual (i.e., singletons) [Roach et al., 2010; Ng et al., 2010b]. It is feasible to identify which variants are more likely to have clinical utility based,

for example, on their predicted effect on protein stability, function, or conservation. However, it will be difficult to predict what, if any, phenotypic consequences are associated with these variants. While such variants could still meet criteria for *potential* clinical utility and investigators may decide to return them to participants, it will be challenging to explain their *actual* utility. Clinical geneticists face this situation regularly, (e.g., novel variants in *BRCA1*), but the scale of this challenge in a research setting with data from ES/WGS is unprecedented.

It will also be necessary to consider how to evaluate the clinical utility of variants identified by ES/WGS as information about disease associations changes over time. Such dynamic interpretation of ES/WGS results will necessitate the development of novel tools for study participants to “revisit” their variant data and interpretations thereof, perhaps similar to those employed by some direct-to-consumer genetic testing companies (e.g., 23andMe). The evolving interpretation of ES/WGS data will also pose new questions about the temporal limits of the duties of researchers and institutions to participants.

Despite the challenges posed by ES/WGS research toward the existing “return of results paradigm,” it does not follow that there ought to be a mandatory review and return of all possible results to participants in all studies. Instead of regarding potentially returnable results as rare or incidental, researchers will need to assume that such results in fact will be identified. The decision about how and whether to return results must take into account other factors, such as the nature of researchers’ relationships with participants, the commitments made at the time of informed consent, and the resources available both to analyze variant data and to responsibly confirm and return results. While some of these factors have been discussed elsewhere as issues to consider related to the possible return of results [Ravitsky and Wilfond, 2006], they have not been the focus of discussion, nor have they been framed in the context of ES/WGS. Accordingly, we contend that ES/WGS increases the urgency with which researchers, policy makers and institutions need to shift the debate towards large-scale results management, rethinking and reframing plans and methods for managing research results within this new context.

Researchers who decide to review ES/WGS data for returnable results will face at least four major decisions. First, researchers will need to decide how to identify “known” disease-causing variants in ES/WGS data and interpret the clinical meaning of such variants. An important consideration is that generalized mutation databases (e.g., Human Gene Mutation Database) contain a small, but not inconsequential number of variants erroneously annotated as disease-causing. They will also need to identify approaches to distinguish variants of known clinical utility, variants of suspected or possible clinical utility and variants of unknown significance. As more ES/WGS data become available, there will be substantial merit to the formation of more inclusive annotated databases for these purposes.

Second, researchers will need to consider what kinds of results to offer to return to participants. Ideally, this would be done at the time of consent and with input from participants about their preferences, although this is often not possible with existing studies. The kinds of results researchers may have an obligation to offer to return depend on the clinical utility and personal meaning of the finding, as well as the nature of the relationship

between the researcher and participant. At a minimum, researchers should consider returning those results that might fall under the rubric of “duty to rescue” about significant health issues. Beskow and Burke define this category of results as “genetic information that clearly indicates a high probability of a serious condition for which an effective intervention is readily available.” [Beskow and Burke, 2010] Examples of such information might include known mutations that result in a high risk of early-onset colon cancer, where having the information would allow participants to pursue more aggressive screening.

Third, consent documents from previously collected samples are unlikely to have addressed the kinds of results that might be identified from ES/WGS. Even if consent documents mentioned the possibility, very few if any described the possible process of offering and/or returning results [Hull et al., 2004]. Accordingly, researchers relying on retrospectively obtained informed consent may find themselves in the position of communicating about possible results with participants who do not know anything about this kind of research and were not expecting to be contacted about these kinds of results. Even if participants have given researchers permission to re-contact them, they may be surprised and even anxious if they are contacted about receiving results for genes and/or traits that they did not know would be studied in the course of their research participation. Researchers will need to decide how to approach possible recontact of participants given these challenges.

Finally, researchers will need to make decisions about practical resources related to the return of results. The volume of variants of clinical utility identified could strain the resources of many investigators and institutions, as well as the healthcare industry in translating them into clinical care. Researchers will need to decide when and how variant data should be offered and reported to study participants, both those who may not have been told about ES/WGS research when they consented to research participation and those who were provided with more specific information. To this end, there is urgency for the communities involved in ES/WGS research to develop and empirically test policies and tools for addressing the management of results from these studies in order to develop best practices for future work.

BROAD NEXT STEPS FOR MOVING FORWARD

We have suggested that research strategies using ES/WGS rather than targeted approaches will challenge the existing ethical framework of human genetics research, largely due to the simultaneous assessment of virtually all protein coding variants. Despite the fact that ES/WGS approaches are already being used widely, several steps can and should be taken to strengthen the consideration of and response to these ethical challenges both in the immediate future and in proactive planning and policy development over the long term (Table II).

Empirical data about how research participants understand and respond to consent documents in a range of contexts would facilitate the development and dissemination of shared consent tools that could be tailored to different research settings and populations. Development of such tools would also benefit from the more widespread availability of consent documents for ES/WGS that have already passed muster with regulatory boards/

agencies. Accordingly, investigators should be encouraged to share their consent documents publicly, perhaps via a publicly available database or through publication as supplementary materials.

The principle of broad sharing of genomic data is widely accepted and followed by researchers. However, the specific goals for sharing of ES/WGS data, and the mechanisms for best accomplishing them, may be different from those used for sharing of GWAS data. An explicit discussion of the unique considerations of ES/WGS data sharing, and the range of possible data sharing approaches, would result in a greater likelihood of achieving the goals of scientific progress that data sharing facilitates. To this end, funding and regulatory agencies should take a more proactive approach toward the development of data sharing policies specifically for ES/WGS studies. These policies should be developed with input from a broad range of stakeholders including researchers, bioethicists, policy-makers, IRBs, and participants. This policy-making process would be enhanced by empirical data that characterize the statistical probability of identification from ES/WGS data, as well as the perceived and actual risks of harm from the potential identifiability of subjects from ES/WGS data.

We have outlined several ways in which ES/WGS studies create new challenges for results management. While advance planning and transparency about return of results at the time consent is obtained is ideal, the speed with which the ES/WGS strategies are developing and our knowledge of the genome is advancing makes this tricky, and in some circumstances could require dynamic interactions with participants over time. The obligations and practical realities associated with the return of results vary considerably by population, researcher, and study, requiring varied and flexible solutions. Others have addressed some of these concerns in the context of WGS, arguing for transparency around return of results and data sharing in the consent process, proactive development of plans for return of results, and engagement of research governance structures that involve participants in decision-making around the use of ES/WGS [Caulfield et al., 2008; McGuire et al., 2008b]. These approaches may be appropriate in certain situations, but as we have suggested, they may not be practical or applicable in many ES/WGS research contexts (e.g., use of existing cohorts). As this new kind of research expands beyond use in small sample collections, the nature and complexity of the ethical challenges will also expand to require novel approaches that can be applied across a wider range of situations and populations.

There is an urgent need for investment in empirical research about results management across a broad range of research studies, together with the development and dissemination of results management tools. Unlike previous discussions around return of results, this research must consider the implications of returning not one but possibly many results of clinical utility simultaneously to large numbers of participants, as well as the particular challenge of interpreting large numbers of variants of unknown significance. This research should examine how participants understand and utilize this information over time, focusing on psychological, social, and health outcomes, and ideally assess the impact on clinical care and costs.

It will be critical to find financial support for management of return of ES/WGS results. Return of results, particularly on the scale possible for ES/WGS, requires substantial financial resources for analysis, genetics counseling, and re-contact of participants. At present, these efforts are given lower priority by funding agencies and academic institutions. The likely exponential increase in studies using ES/WGS may require a substantial shift in thinking about whether and how return of results efforts should be supported.

We think that any framework that addresses the ethics of ES/WGS studies must first acknowledge that the application of new sequencing technologies represents a departure from the *status quo* in human genetic research. ES/WGS is not an incremental step forward, but rather a “magnitude of order” leap into unprecedented levels of production data that includes variants of varying utility. As with many new technologies, use of ES/WGS for gene discovery raises a range of important questions and concerns. We encourage researchers, funding agencies, IRBs, and the research ethics community, in collaboration with research participants, to proactively consider how to best to address emerging ethical issues around management of results, data sharing and informed consent. By considering the potential challenges and testing possible solutions, the research community will be able to both better maintain public trust and advance science of human genetics.

Acknowledgments

Grant sponsor: National Institutes of Health/National Human Genome Research Institute; Grant number: 5R00HG004316-04; Grant sponsor: NIH; Grant number: 1R01HG006618-01; Grant sponsor: National Institutes of Health/National Human Genome Research Institute; Grant number: 5P50HG003374-07; Grant sponsor: National Institutes of Health/National Center for Research Resources 1UL1 RR025014; Funding source: Intramural Research Program of the National Human Genome Research Institute; Funding source: National Institutes of Health.

We would like to acknowledge the helpful comments of Wylie Burke, Barbara Biesecker, David Wendler, James Mullikin, and Richard Myers. Our work was supported in part by grants from the National Institutes of Health (5R00HG004316-04 and 5P50HG003374-07 to HKT and 1R01HG006618-01 to HKT and MB), and National Institutes of Health/National Center for Research Resources (1UL1 RR025014-01 to HKT). This research was also supported in part by the Intramural Research Program of the National Human Genome Research Institute, National Institutes of Health (SCH and BEB).

References

- Ashley EA, Butte AJ, Wheeler MT, Chen R, Klein TE, Dewey FE, Dudley JT, Ormond KE, Pavlovic A, Morgan AA, Pushkarev D, Neff NF, Hudgins L, Gong L, Hodges LM, Berlin DS, Thorn CF, Sangkuhl K, Hebert JM, Woon M, Sagreiya H, Whaley R, Knowles JW, Chou MF, Thakuria JV, Rosenbaum AM, Zaranek AW, Church GM, Greely HT, Quake SR, Altman RB. Clinical assessment incorporating a personal genome. *Lancet*. 2010; 375:1525–1535. [PubMed: 20435227]
- Beskow LM, Burke W. Offering individual genetic research results: Context matters. *Sci Transl Med*. 2010; 30:38cm20.
- Bookman EB, Langehorne AA, Eckfeldt JH, Glass KC, Jarvik GP, Klag M, Koski G, Motulsky A, Wilfond B, Manolio TA, Fabsitz RR, Luepker RV. NHLBI Working Group. Reporting genetic results in research studies: A summary and recommendations of an NHLBI working group. *Am J Med Genet Part A*. 2006; 140A:1033–1040. [PubMed: 16575896]
- Caulfield T, McGuire AL, Cho M, Buchanan JA, Burgess MM, Danilczyk U, Diaz CM, Fryer-Edwards K, Green SK, Hodosh MA, Juengst ET, Kaye J, Kedes L, Knoppers BM, Lemmens T, Meslin EM, Murphy J, Nussbaum RL, Otlowski M, Pullman D, Ray PN, Sugarman J, Timmons M. Research Ethics Recommendations for Whole-Genome Research: Consensus Statement. *PLOS Biol*. 2008; 6:430–435.

- Cho MK. Understanding incidental findings in the context of genetics and genomics. *J Law Med Ethics*. 2008; 36:280–285. 212. [PubMed: 18547195]
- Dressler LG. Disclosure of research results from cancer genomic studies: The state of the science. *Clin Cancer Res*. 2009; 13:4270–4276. [PubMed: 19549775]
- Fabsitz RR, McGuire A, Sharp RR, Puggal M, Beskow LM, Biesecker LG, Bookman E, Burke W, Burchard EG, Church G, Clayton EW, Eckfeldt JH, Fernandez CV, Fisher R, Fullerton SM, Gabriel S, Gachupin F, James C, Jarvik GP, Kittles R, Leib JR, O'Donnell C, O'Rourke PP, Rodriguez LL, Schully SD, Shuldiner AR, Sze RK, Thakuria JV, Wolf SM, Burke GL. Ethical and practical guidelines for reporting genetic research results to study participants: Updated guidelines from a National Heart, Lung and Blood Institute working group. *Circ Cardiovasc Genet*. 2010; 3:574–580. [PubMed: 21156933]
- Gilissen C, Arts HH, Hoischen A, Spruijt L, Mans DA, Arts P, van Lier B, Steehouwer M, van Reeuwijk J, Kant SG, Roepman R, Knoers NV, Veltman JA, Brunner HG. Exome sequencing identifies WDR35 variants involved in Sensenbrenner syndrome. *Am J Hum Genet*. 2010; 87:418–423. [PubMed: 20817137]
- Hoischen A, van Bon BW, Gilissen C, Arts P, van Lier B, Steehouwer M, de Vries P, de Reuver R, Wieskamp N, Mortier G, Devriendt K, Amorim MZ, Revencu N, Kidd A, Barbosa M, Turner A, Smith J, Oley C, Henderson A, Hayes IM, Thompson EM, Brunner HG, de Vries BB, Veltman JA. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat Genet*. 2010; 42:483–855. [PubMed: 20436468]
- Homer N, Tembe WD, Szelinger S, Redman M, Stephan DA, Pearson JV, Nelson SF, Craig D. Multi-marker analysis and imputation of multiple platform pooling-based genome-wide association studies. *Bioinformatics*. 2008; 24:1896–1902. [PubMed: 18617537]
- Hull SC, Gooding H, Klein AP, Warshauer-Baker E, Metosky S, Wilfond BS. Genetic research involving human biological materials: A need to tailor current consent forms. *IRB*. 2004; 26:1–7. [PubMed: 15281193]
- Hull SC, Sharp RR, Botkin JR, Brown M, Hughes M, Sugarman J, Schwinn D, Sankar P, Bolcic-Jankovic D, Clarridge BR, Wilfond BS. Patients' views on identifiability of samples and informed consent for genetic research. *Am J Bioeth*. 2008; 8:62–70. [PubMed: 19003716]
- Karow, J. At Least \$145M in FY'09 NIH Stimulus Funding Goes to Sequencing; Disease-Focused Studies Dominate. Oct 20. 2009 http://recovery.nih.gov/stories/disease_focused.php
- Levenson D. After Havasupai litigation, Native Americans wary of genetic research. *Am J Med Genet*. 2010; (Part A):152A:ix.
- Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Huang J, Kirkness EF, Denisov G, Lin Y, MacDonald JR, Pang AW, Shago M, Stockwell TB, Tsiamouri A, Bafna V, Bansal V, Kravitz SA, Busam DA, Beeson KY, McIntosh TC, Remington KA, Abril JF, Gill J, Borman J, Rogers YH, Frazier ME, Scherer SW, Strausberg RL, Venter JC. The diploid genome sequence of an individual human. *PLoS Biol*. 2007; 5:e254. [PubMed: 17803354]
- Li Y, Vinckenbosch N, Tian G, Huerta-Sanchez E, Jiang T, Jiang H, Albrechtsen A, Andersen G, Cao H, Korneliussen T, Grarup N, Guo Y, Hellman I, Jin X, Li Q, Liu J, Liu X, Sparsø T, Tang M, Wu H, Wu R, Yu C, Zheng H, Astrup A, Bolund L, Holmkvist J, Jørgensen T, Kristiansen K, Schmitz O, Schwartz TW, Zhang X, Li R, Yang H, Wang J, Hansen T, Pedersen O, Nielsen R, Wang J. Resequencing of 200 exomes identifies an excess of low-frequency non-synonymous coding variants. *Nat Genet*. 2010; 42:969–972. [PubMed: 20890277]
- Lupski JR, Reid JG, Gonzaga-Jauregui C, Rio Deiros D, Chen DC, Nazareth L, Bainbridge M, Dinh H, Jing C, Wheeler DA, McGuire AL, Zhang F, Stankiewicz P, Halperin JJ, Yang C, Gehman C, Guo D, Irikat RK, Tom W, Fantin NJ, Muzny DM, Gibbs RA. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med*. 2010; 362:1181–1191. [PubMed: 20220177]
- Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, Hao L, Kiang A, Paschall J, Phan L, Popova N, Pretel S, Ziyabari L, Lee M, Shao Y, Wang ZY, Sirotkin K, Ward M, Kholodov M, Zbicz K, Beck J, Kimelman M, Shevelev S, Preuss D, Yaschenko E, Graeff A, Ostell J, Sherry ST. The NCBI dbGaP database of genotypes and phenotypes. *Nat Genet*. 2007; 39:1181–1186. [PubMed: 17898773]

- McGuire AL, Hamilton JA, Lunstroth R, McCullough LB, Goldman A. DNA data sharing: Research participants' perspectives. *Genet Med*. 2008a; 10:46–53. [PubMed: 18197056]
- McGuire AL, Caulfield T, Cho MK. Research ethics and the challenge of whole genome sequencing. *Nat Rev Genet*. 2008b; 9:152–156. [PubMed: 18087293]
- Mello MM, Wolf LE. The Havasupai Indian tribe case—Lessons for research involving stored biologic samples. *N Engl J Med*. 2010; 353:204–207. [PubMed: 20538622]
- Miller FA, Giacomini M, Ahern C, Robert JS, de Laat S. When research seems like clinical care: A qualitative study of the communication of individual cancer genetic research results. *BMC Med Ethics*. 2008; 22:4. [PubMed: 18294373]
- Murphy J, Scott J, Kaufman D, Geller G, LeRoy L, Hudson K. Public expectations for return of results from large-cohort genetic research. *Am J Bioeth*. 2008; 8:36–43. [PubMed: 19061108]
- NBAC. Research involving human biological materials: Ethical issues and policy guidance. Report and Recommendations of the National Bioethics Advisory Commission. 1999; 1 Available from: http://www.bioethics.gov/reports.past_commissions/nbac_biological1.pdf.
- Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, Beck AE, Tabor HK, Cooper GM, Mefford HC, Lee C, Turner EH, Smith JD, Rieder MJ, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson DA, Bamshad MJ, Shendure J. Exome sequencing identified MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet*. 2010; 42:790–793. [PubMed: 20711175]
- Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, Huff CD, Shannon PT, Jabs EW, Nickerson DA, Shendure J, Bamshad MJ. Exome sequencing identifies the cause of a Mendelian disorder. *Nat Genet*. 2010; 42:30–36. [PubMed: 19915526]
- Pelak K, Shianna KV, Ge D, Maia JM, Zhu M, Smith JP, Cirulli ET, Fellay J, Dickson SP, Gumbs CE, Heinzen EL, Need AC, Ruzzo EK, Singh A, Campbell CR, Hong LK, Lornsen KA, McKenzie AM, Sobreira NL, Hoover-Fong JE, Milner JD, Ottman R, Haynes BF, Goedert JJ, Goldstein DB. The characterization of twenty sequence human genomes. *PLoS Genet*. 2010; 6:e1001111. [PubMed: 20838461]
- Pulley JM, Brace MM, Bernard GR, Masys DR. Attitudes and perceptions of patients towards methods of establishing a DNA biobank. *Cell Tissue Bank*. 2008; 9:55–65. [PubMed: 17960495]
- Ravitsky V, Wilfond BS. Disclosing individual genetic results to research participants. *Am J Bioeth*. 2006; 6:8–17. [PubMed: 17085395]
- Rios J, Stein E, Shendure J, Hobbs HH, Cohen JC. Identification by whole-genome resequencing of gene defect responsible for severe hyper-cholesterolemia. *Hum Mol Genet*. 2010; 19:4313–4318. [PubMed: 20719861]
- Roach JC, Glusman G, Smit AF, Huff CD, Hubley R, Shannon PT, Rowen L, Pant KP, Goodman N, Bamshad M, Shendure J, Drmanac R, Jorde LB, Hood L, Galas DJ. Analysis of genetic inheritance in a family quartet by whole genome sequencing. *Science*. 2010; 328:636–639. [PubMed: 20220176]
- Schuster SC, Miller W, Ratan A, Tomsho LP, Giardine B, Kasson LR, Harris RS, Petersen DC, Zhao F, Qi J, Alkan C, Kidd JM, Sun Y, Drautz DI, Bouffard P, Muzny DM, Reid JG, Nazareth LV, Wang Q, Burhans R, Riemer C, Wittekindt NE, Moorjani P, Tindall EA, Danko CG, Teo WS, Buboltz AM, Zhang Z, Ma Q, Oosthuisen A, Steenkamp AW, Oostuisen H, Venter P, Gajewski J, Zhang Y, Pugh BF, Makova KD, Nekrutenko A, Mardis ER, Patterson N, Pringle TH, Chiaromonte F, Mullikin JC, Eichler EE, Hardison RC, Gibbs RA, Harkins TT, Hayes VM. Complete Khoisan and Bantu genomes from southern Africa. *Nature*. 2010; 463:943–947. [PubMed: 20164927]
- Summary of principles agreed at the International Strategy Meeting on Human Genome Sequencing; Bermuda. 25–28 February 1996; www.ornl.gov/sci/techresources/Human_Genome/research/bermuda.shtml#1
- Trinidad SB, Fullerton SM, Bares JM, Jarvik GP, Larson EB, Burke W. Genome research and wide data sharing: Views of prospective participants. *Genet Med*. 2010; 12:486–495. [PubMed: 20535021]
- Trinidad SB, Fullerton SM, Ludman EJ, Jarvik GP, Larson EB, Burke W. Research ethics. Research practice and participant preferences the growing gulf. *Science*. 2011; 331:287–288. [PubMed: 21252333]

- Wang J, Wang W, Li R, Li Y, Tian G, Goodman L, Fan W, Zhang J, Li J, Zhang J, Guo Y, Feng B, Li H, Lu Y, Fang X, Liang H, Du Z, Li D, Zhao Y, Hu Y, Yang Z, Zheng H, Hellmann I, Inouye M, Pool J, Yi X, Zhao J, Duan J, Zhou Y, Qin J, Ma L, Li G, Yang Z, Zhang G, Yang B, Yu C, Liang F, Li W, Li S, Li D, Ni P, Ruan J, Li Q, Zhu H, Liu D, Lu Z, Li N, Guo G, Zhang J, Ye J, Fang L, Hao Q, Chen Q, Liang Y, Su Y, San A, Ping C, Yang S, Chen F, Li L, Zhou K, Zheng H, Ren Y, Yang L, Gao Y, Yang G, Li Z, Feng X, Kristiansen K, Wong GK, Nielsen R, Durbin R, Bolund L, Zhang X, Li S, Yang H, Wang J. The diploid sequence of an Asian individual. *Nature*. 2008; 456:60–65. [PubMed: 18987735]
- Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song XZ, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA, Rothberg JM. The complete genome of an individual by massively parallel DNA sequencing. *Nature*. 2008; 452:872–876. [PubMed: 18421352]
- Wolf SM, Lawrenz FP, Nelson CA, Kahn JP, Cho MK, Clayton EW, Fletcher JG, Georgieff MK, Hammerschmidt D, Hudson K, Illes J, Kapur V, Keane MA, Koenig BA, Leroy BS, McFarland EG, Paradise J, Parker LS, Terry SF, Van Ness B, Wilfond BS. Managing incidental findings in human subjects research: Analysis and recommendations. *J Law Med Ethics*. 2008a; 36:219–248. 211. [PubMed: 18547191]
- Wolf SM, Paradise J, Cagaanan C. The law of incidental findings in human subjects research: Establishing researchers' duties. *J Law Med Ethics*. 2008b; 36:361–383. 214. [PubMed: 18547206]
- Worthey EA, Mayer AN, Syverson GD, Helbling D, Bonacci BB, Decker B, Serpe JM, Dasu T, Tschannen MR, Veith RL, Basehore MJ, Broeckel U, Tomita-Mitchell A, Arca MJ, Casper JT, Margolis DA, Bick DP, Hessner MJ, Routes JM, Verbsky JW, Jacob HJ, Dimmock DP. Making a definitive diagnosis: Successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genet Med*. 2011; 13:255–262. [PubMed: 21173700]

TABLE I**Areas for Review of Existing Consent Forms for Exome Sequencing/Whole Genome Sequencing Research**

Area of review	Questions
Stated purpose of the research	Is the stated research purpose constrained to the study of just a particular disease or trait? This may limit uses of the data to study other diseases or traits Does the research state that participants will be re-contacted if other uses are proposed?
Data and sample sharing	Do the consent forms place any limitations on sharing of either samples or data? Do these limitations prohibit dbGaP sharing or other broad data sharing? Do the consent forms make any statements about whether participants will be re-contacted or notified before data or samples are shared? Do the consent forms make any statements about IRB review of any other protocols/investigators that use the data?
Return of results	Do the consent forms make any statements about whether results will be returned, and what kinds of results will be returned? For example do they specifically address genetic results? Clinically significant or medically important results? Do the consent forms make any statements about how genetic results would be validated if they were going to be returned? Do the consent forms make any statements about who might return results, how and when?

TABLE II**Areas for Empirical Research and Policy Development**

Consent	Empirical research about consent documents for ES/WGS
	Development of shared consent tools for ES/WGS, tailored to different research settings and populations
	Public sharing of consent documents from ES/WGS studies
Data sharing	Discussion and consensus building around specific goals and best mechanisms for sharing of ES/WGS data
	Development of ES/WGS specific data sharing policies, involving a range of stakeholders
	Consideration of specific harms deriving from ES/WGS data, as compared to other genomic data
Return of results	Empirical research about results management across a broad range of research studies, including data about how participants understand and utilize this information over time
	Development and dissemination of results management tools that are designed to return results on a large scale
	Increased support/funding provided to researchers by funding agencies to support analysis, genetic counseling and recontact of participants by investigators as indicated