Genetics inMedicine | ARTICLE



# Impact of integrated translational research on clinical exome sequencing

A full list of authors and affiliations appears at the end of the paper.

**Purpose:** Exome sequencing often identifies pathogenic genetic variants in patients with undiagnosed diseases. Nevertheless, frequent findings of variants of uncertain significance necessitate additional efforts to establish causality before reaching a conclusive diagnosis. To provide comprehensive genomic testing to patients with undiagnosed disease, we established an Individualized Medicine Clinic, which offered clinical exome testing and included a Translational Omics Program (TOP) that provided variant curation, research activities, or research exome sequencing.

**Methods:** From 2012 to 2018, 1101 unselected patients with undiagnosed diseases received exome testing. Outcomes were reviewed to assess impact of the TOP and patient characteristics on diagnostic rates through descriptive and multivariate analyses.

**Results:** The overall diagnostic yield was 24.9% (274 of 1101 patients), with 174 (15.8% of 1101) diagnosed on the basis of clinical exome sequencing alone. Four hundred twenty-three patients with nondiagnostic or without access to clinical exome sequencing were evaluated by the TOP, with 100 (9% of 1101) patients receiving a diagnosis, accounting for 36.5% of the diagnostic yield. The identification of a genetic diagnosis was influenced by the age at time of testing and the disease phenotype of the patient.

**Conclusion:** Integration of translational research activities into clinical practice of a tertiary medical center can significantly increase the diagnostic yield of patients with undiagnosed disease.

*Genetics in Medicine* (2021) 23:498–507; https://doi.org/10.1038/s41436-020-01005-9

**Keywords:** diagnostic odyssey; undiagnosed disease; variants of uncertain significance; clinical practice; genomics

## INTRODUCTION

The term "diagnostic odyssey" was coined to describe the protracted series of clinical consultations and testing that patients with undiagnosed disorders often undergo in search of a genetic diagnosis. The clinical adoption of exome sequencing, driven by increased efficiency and accuracy of data analysis and interpretation, reduced cost, and improved insurance coverage, has increased the diagnostic yield of these patients.[1]

In earlier reports of unselected patient populations (by age and disease phenotype) the diagnostic yield of exome sequencing has been reported to be 25–30%.[2,3] However, these studies were centered mainly on pediatric patients, were of short duration (2 years), or reported the exome sequencing experiences of diagnostic laboratories on rare Mendelian diseases.[2,3] Despite the transformative impact on diagnosing undiagnosed disease, a majority of patients with suspected rare genetic disease receiving clinical exome sequencing remain undiagnosed. This is often due to identification of either variants of uncertain significance (VUS) lacking enough evidence to be classified as pathogenic, or variants in candidate disease genes that do not yet have a proven disease association (genes of uncertain significance [GUS]).

Several programs have been established to address diagnostic barriers including access to testing and resolution of uncertain findings. Major efforts include the Undiagnosed Diseases Network,[4] the National Institutes of Health Undiagnosed Diseases Program,[5] and the global Undiagnosed Diseases Network International.[6] Studies by these programs and others highlight the benefits of exome sequencing as well as the utility of periodic reanalysis of genetic data,[7] and the use of model organisms.[8]

In 2012, Mayo Clinic launched an Individualized Medicine Clinic (IMC) to integrate genomic sequencing into clinical practice[9] with a primary focus on identifying the genetic mechanisms of disease for patients on a diagnostic odyssey.[10] Two essential components of the IMC are the Genomic Odyssey Board (GOB) and the Translational Omics Program (TOP). The GOB consists of clinicians, scientists, laboratorians, ethicists, and bioinformaticians who meet weekly to interpret patient data along with genomics findings and recommend clinical or research follow-up through multidisciplinary review and discussion. The TOP consists of bioinformaticians, scientists, and laboratorians who provide individualized and integrated translational research to resolve uncertain genetic findings or identify

Correspondence: Eric W. Klee (klee.eric@mayo.edu) or Konstantinos N. Lazaridis (Lazaridis.Konstantinos@mayo.edu)
These authors contributed equally: Margot A. Cousin, Filippo Pinto e Vairo, Joel A. Morales-Rosado, Erica L. Macke

Submitted 1 July 2020; revised 30 September 2020; accepted: 1 October 2020
Published online: 4 November 2020

genetic variation elusive to standard testing strategies. In this large study of 1101 unselected patients on a diagnostic odyssey from a tertiary medical center spanning over 6 years, we describe the impact of integrated translational research on clinical exome sequencing in improving the diagnostic rate of these patients.

## MATERIALS AND METHODS

### Ethics statement

The Mayo Clinic Institutional Review Board granted a waiver of consent for all the available clinical data of the electronic health records of the patients included in the study. All individuals participating in the described research activities provided written informed consent to a study approved by the Mayo Clinic Institutional Review Board.

### Patients

The reported cohort is comprised of 1101 unselected patients evaluated at the IMC of Mayo Clinic between its inception on 30 September 2012 to 31 December 2018, for whom clinical or research exome sequencing was completed. None of the patients had a clinical genetic diagnosis despite prior genetic testing, including single-gene, targeted gene panel, chromosomal microarray testing, or mitochondrial DNA sequencing. This comprehensive study includes a subset of 51 patients previously reported.[10] All patients were seen by Mayo Clinic's Department of Clinical Genomics, at all three enterprise sites (Rochester, MN; Jacksonville, FL; and Scottsdale, AZ). Following clinician review and recommendation, all patients referred to the Department of Clinical Genomics determined to be candidates for exome sequencing and for whom the relevant results were returned were included in this study without phenotypic selection and offered consent for testing and/or research.

### Exome sequencing

One thousand forty-one patients received clinical exome sequencing at CLIA-certified, College of American Pathology (CAP)-accredited laboratories. Sixty patients underwent only research exome sequencing from the Mayo Clinic Medical Genome Facility in Rochester, Minnesota, which was subsequently analyzed by TOP (outlined in Supplementary appendix). Reportable findings derived from the research exome sequencing were confirmed using targeted clinical genetic testing.

### Patient phenotyping

Phenotypic information of each patient was manually extracted from the electronic medical record and recorded by the TOP team using structured Human Phenotype Ontology (HPO) terms combined with textual description of the reason for referral. This information was categorized by two physicians by phenotype/disease into seven main phenotype categories and 23 subcategories. Phenotype stratification criteria are outlined in Table S1.

### Translational research impact

Retrospective analysis of patients with a genetic diagnosis was performed to determine the impact of the TOP. Patients were referred to the TOP when clinical exome sequencing was not revealing a diagnosis, returned VUS or GUS of interest, or the disease phenotype of the proband was questioned in relation to the reported variant. The curation and research findings of the TOP were evaluated by the GOB for a final decision on whether or not a genetic diagnosis was reached. The TOP research activities included variant curation, carrying out or facilitating functional studies, leading or participating in cohort studies, in silico protein modeling, RNA-sequencing, reanalysis of raw clinical sequencing data, or research exome sequencing.

## RESULTS

One thousand one hundred one unselected patients were referred to the IMC for suspected genetic disease eluding diagnosis through any other testing modality during the study period and underwent exome sequencing. As illustrated in Fig. 1, all individuals were evaluated by a medical geneticist and offered enrollment in the TOP.

### Cohort description

This cohort was assembled without selection for disease phenotype, age, or sex and consists of both pediatric and adult patients presenting with variable symptoms/signs or conditions across multiple clinical areas, many with diverse and complex phenotypes. The cohort is 51.2% female and 49.6% pediatric. Age at the time of the exome sequencing report ranged from 0 to 85 years, with a median age of 18 years (Fig. S1). The distribution of disease phenotypes is illustrated in Fig. 2. Of the seven primary disease categories, neurological disorders (including complex neurological, movement disorders, autism spectrum disorder [ASD] and developmental delay, seizures, and neuromuscular phenotype) was the largest (46.3% of patients). The multisystem disorders group was the second largest disease category, representing only 18.5% of patients. The five remaining categories each accounted for no more than 10% of the entire cohort.

### Sequencing results

Exome testing was completed through a clinical laboratory for 94.5% (1041 of 1101) of patients and was diagnostic, without the need for additional research study, for 15.8% (174 of 1101). Patients with negative or uncertain clinical exome sequencing findings were referred to the TOP after initial exome results review by a medical geneticist for variant curation and/or other individualized research efforts to attempt a genetic diagnosis. Sixty patients (5.4%) underwent research exome sequencing, in the absence of a clinical exome sequencing, after enrollment in the research study. These patients were analyzed by the TOP along with clinical laboratory confirmation of the diagnostic genetic findings. Patients unable to pursue clinical exome testing due to denied insurance coverage or high out-of-pocket costs were offered

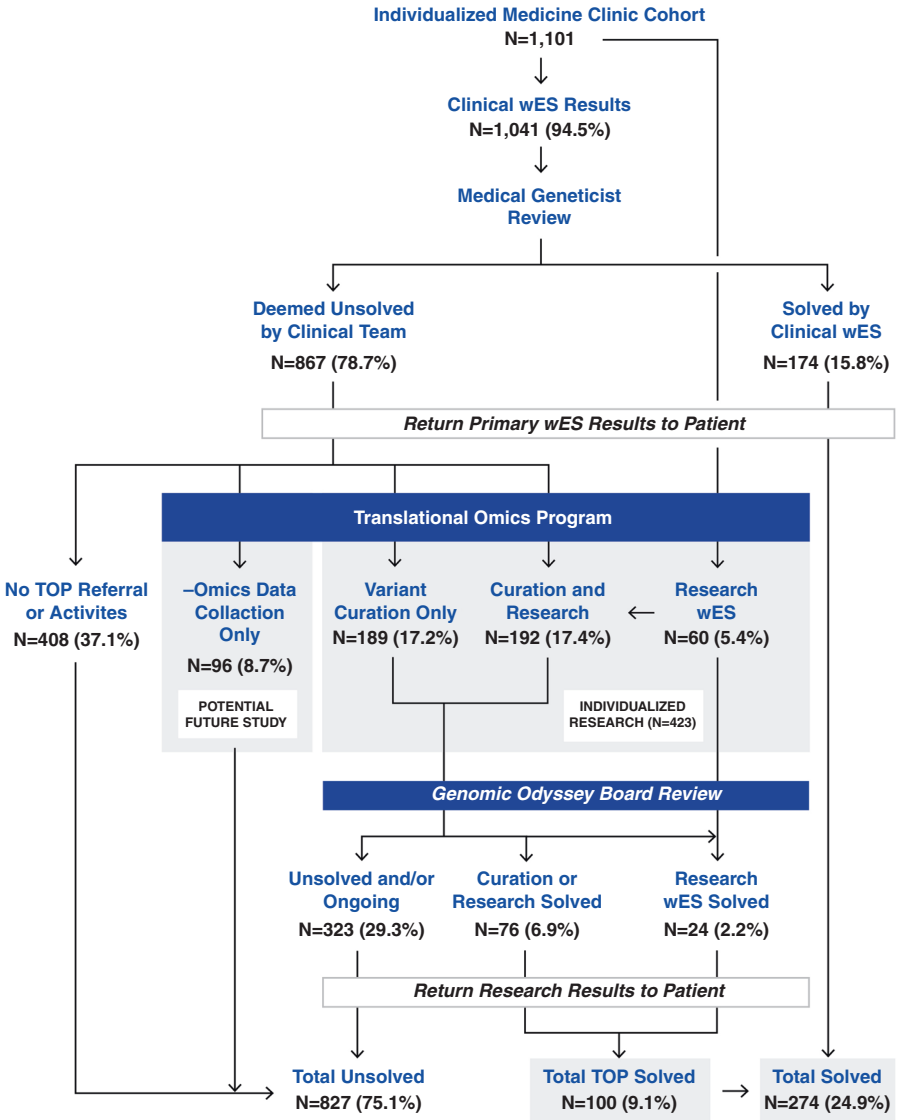


**Fig. 1 Workflow of patients of at the Individualized Medicine Clinic.** Patients were seen by a medical geneticist and genetic counselor from the Department of Clinical Genomics, where clinical exome sequencing was ordered and consent to the research program was offered. Following receipt of clinical results, research services when appropriate were pursued to identify or clarify a genetic finding. A multidisciplinary review board, the Genomic Odyssey Board, was convened to discuss the clinical and research findings for a case, leading to a final genetic diagnosis decision and disclosure of results to the patient. For patients who did not have a clinical exome sequencing completed a research exome was offered. *ES* exome sequencing, *Solved* genetic diagnosis was established for the patient, *TOP* Translational Omics Program.

research exome sequencing through a need-based program at Mayo Clinic. Of the 60 patients undergoing research exome sequencing, 24 (40.0%) received a genetic diagnosis. For 7 of these patients, the research exome sequencing findings identified VUS, variants in a GUS, or only a single variant in a gene associated with a recessive condition but additional TOP research efforts were able to resolve the diagnosis. Results from this evaluation of TOP were discussed at the GOB for expert multidisciplinary review and decision.

### TOP research

Of the 867 patients who received nondiagnostic clinical exome sequencing testing, 363 (41.9%) received research assessment from the TOP. Of the 60 patients who received research exome testing, 18 received additional research support, and for 7 of these, a diagnosis was made. In total, 423 patients received individualized TOP research evaluation. This effort, to date, has resulted in resolving genetic diagnoses for 100 patients undiagnosed based on the clinical effort alone (23.6% of TOP referrals) contributing 36.5% of the total 274 patients receiving a diagnosis in this cohort.

Figure 3a summarizes those research activities pursued in the 274 patients receiving a genetic diagnosis and the proportion where this activity positively contributed to the diagnosis. The most common research activity pursued was detailed variant curation in which literature, variant-based data, and data from publicly available databases (gnomAD, Human Gene Mutation Database, ClinVar, etc.) were

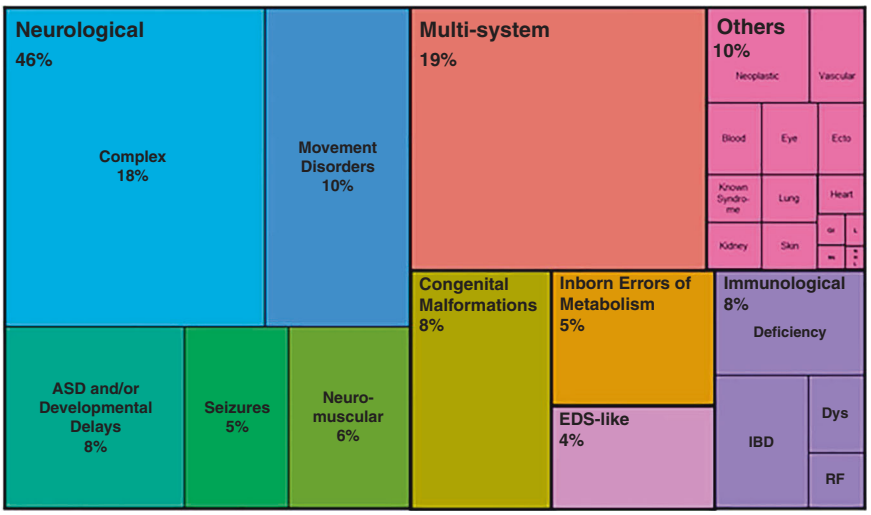


**Fig. 2 Individualized Medicine Clinic cohort disease phenotype spectrum.** The disease phenotype(s) of all patients was reviewed and assigned to one of seven high-level categories, and in some instances a more specific subcategory. Neurological disease was most representative in the study cohort (46.3%), with multisystem disorders second (18.5%). The phenotypic makeup of many individuals was relatively unique in the cohort, and given the low numbers were assigned to the "Others" major category, and then provided a more descriptive subcategory assignment. *ASD* autism spectrum disorder, *Dys* immune dysregulation, *Ecto* ectodermal dysplasia, *GI* gastrointestinal, *IBD* inflammatory bowel disease, *L* liver-related disorders, *Mc* miscellaneous, *NHL* nonsyndromic hearing loss, *RF* recurrent fevers.

summarized in the context of the patient's disease phenotype and discussed at the GOB. This was the only research activity performed for 189 patients with nondiagnostic clinical exome sequencing, of whom 26 were deemed solved after TOP evaluation and GOB review. Of these 26 individuals, literature evidence was identified supporting that the patient's phenotype can be explained by the genetic findings for 14 individuals. Often these instances involve nonclassic phenotypic presentations or newly emerging or extremely rare disorders. For eight individuals, variant or gene-level data such as functional studies or newly published studies reporting the same or similar variation were supportive of pathogenicity of the alteration(s) identified. Six individuals had variants in more than one gene reported that could have contributed to the disease presentation and additional review determined these were not likely contributory. Three individuals had clinical confirmatory tests suggested by literature, two had professional communication with external clinicians or disease experts on an emerging disorder that confirmed pathogenicity, and for one, an evaluation of a potential splice-impacting missense variant was provided supporting that the impact of the variant is consistent with the known disease mechanism.

Individualized translational research beyond only variant curation was pursued for 192 patients referred to the TOP (174 patients from clinical exome sequencing and 18 patients from research exome sequencing). RNA-sequencing was leveraged in 25 cases and 14 were subsequently solved, with RNA-sequencing being integral to the diagnosis in 10 cases (38.4%), which is similar to the reported diagnostic rates of RNA-sequencing in the literature.[11] For seven individuals, the impact of potential splice variants was resolved;[12,13] for two, gene fusions were detected;[14–16] and for one, outlier expression analysis determined a significantly reduced expression level and monoallelic expression.

The TOP also used in silico, in vitro, and in vivo functional models to clarify the nature of VUS[17,18] and to provide evidence for new disease–gene associations.[13] Our ability to pursue functional studies was limited primarily by the current understanding the candidate gene's function, the availability of appropriate samples, and the expertise to carry out the relevant experiments. Of those individuals who achieved a diagnosis, 11 individuals had functional studies conducted using patient samples (for 5, some the assays were done through internal or external collaborations); 8 had cell-based studies[10,19,20] and 4 had studies using other sample types from patients such as blood or serum.[10,21–23] Eleven patients had heterologous cell-based studies[10,18,24–27] including electrophysiology, reporter, and cellular localization studies and others, of which six were accomplished through collaborative efforts. For six individuals, animal models where used[28–30] to better understand the gene–disease relationships or the underlying disease mechanism and included the use of *Drosophila*, zebrafish, and mouse studies; two of the zebrafish studies were conducted by the TOP and the remaining were in collaboration. For three patients without both parental samples available, variant phase was determined to be *trans* using RNA for two patients or DNA for one using polymerase chain reaction (PCR)–based approaches. Additionally, for two individuals targeted PCR-based RNA evaluation determined splice variant impact, and for one individual, recombinant protein studies were pursued to determine protein folding abnormalities induced by the genetic variant.[17] In silico protein modeling

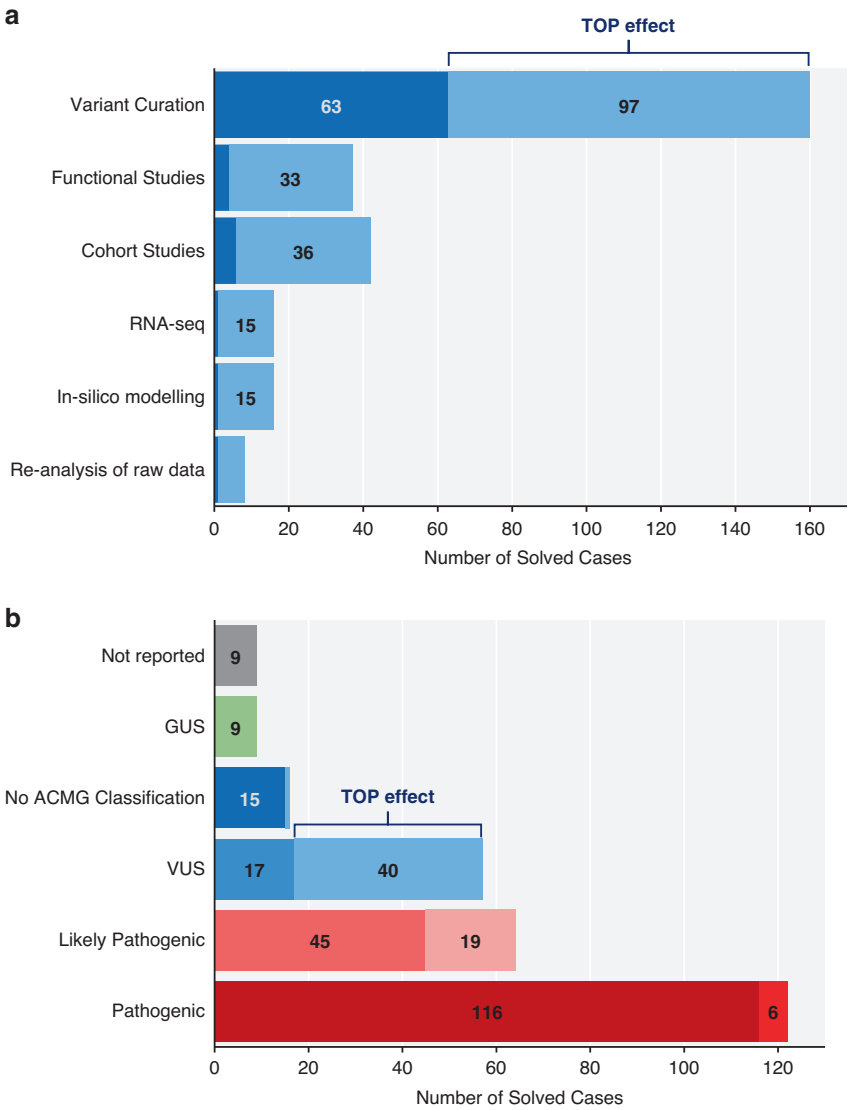


**Fig. 3 Translational Omics Program (TOP) research activities and efforts.** (**a**) The research activities were broadly assigned to one of seven categories. For the majority of patients, the translational research was critical to the final genetic diagnosis, but there were patients where this effort was pursued but not critical to the genetic finding that was ultimately reported as causal. (**b**) The TOP primarily focused on research activities related to the resolution of variants of uncertain significance (VUS) based on the clinical exome sequencing report variant classification that led to a genetic diagnosis. Furthermore, the TOP was critical in identifying previously unidentified candidates by clinical testing. *GUS* gene of uncertain significance.

studies were also pursued for 15 individuals who achieved a diagnosis.[10,18,23,31–35]

Finally, support of multi-institutional cohort studies is an important aspect of the TOP research studies, which resulted in several diagnoses.[27,30,36,37] Of these 192 patients receiving additional evaluation and analysis, 53 were subsequently solved by the TOP research efforts. As shown in Fig. 3**b**, the majority of the TOP effort was spent investigating VUS, but these efforts also spanned GUS, unreported variants, as well as likely pathogenic, or, in cases of questionable phenotypic fit with the disease, pathogenic variants. In each of these solved cases, it was the integration of clinical exome testing along with the TOP research findings that enabled a genetic diagnosis to be identified.

## Diagnostic yield

In total, 24.9% (274 of 1101) of patients included in this study received genetic diagnoses: 174 through clinical exome sequencing alone and 100 through individualized research by the TOP. Among these 274 patients, 284 genetic diagnoses were reported. One patient was diagnosed with three independent monogenic disorders,[38] and eight patients were diagnosed with two independent monogenic disorders. Inheritance patterns of the diagnosed conditions included 60.6% (172 of 284) dominant disorders, 30.6% (87 of 284) recessive disorders, and 8.8% (25 of 284) X-linked disorders. We report on 14 emerging gene–disease associations. The complete list of reported disease-causing variants and phenotypic classification is provided in Table S2.

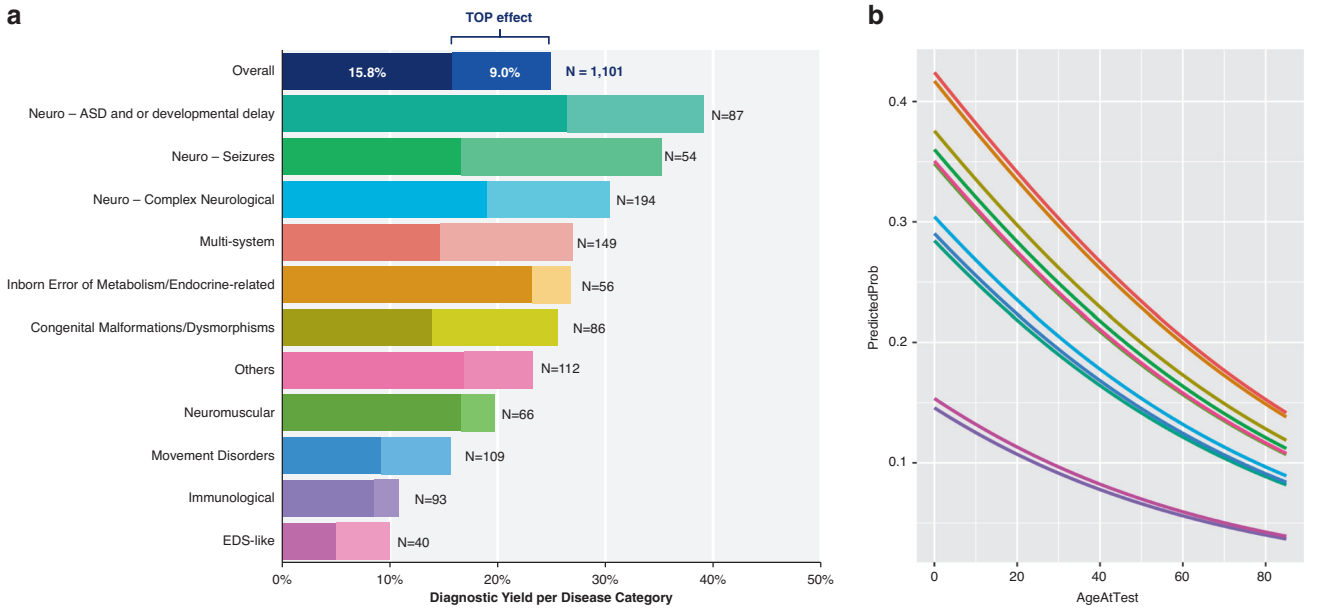


**Fig. 4 Diagnostic yield by disease category and age, and impact of the Translational Omics Program (TOP).** (**a**) Diagnostic yield per disease category. Light bar colors depict patients with diagnostic yield impacted by TOP efforts or activities. (**b**) Age at time of testing and probability of diagnostic yield across disease categories. *ASD* autism spectrum disorder, *EDS* Ehlers–Danlos syndrome.

### Confounding factors

Data review of the cohort identified several factors that may influence the ability to identify a genetic diagnosis, including category of disease, age at time of testing, and family members included in testing. Trio exome sequencing was performed for 55.6% of the probands, 18.2% received singleton exome sequencing, and 15.8% received singleton exome sequencing with additional family members used for segregation analysis. The remainder had other combinations of proband and family members tested. The diagnostic rate by disease category displays a range from 10% to 39%, as illustrated in Fig. 4**a**. Pediatric patients had a diagnostic rate of 31.7%, while adult patients had a diagnostic rate of 18.2%. To account for confounding effects between factors, logistic regression was used to evaluate the effect of testing laboratory, age, family members included in testing (pedigree), and disease category simultaneously. Age at the time of the exome testing was not different between group of solved versus group of unsolved status (i.e., whether or not a genetic diagnosis was determined [Fig. S2]). Age was statistically significant ($P < 0.01$) even after controlling for clinical lab, pedigree, and disease categories. Likewise, the overall effect of disease category was statistically significant (Wald $P = 0.016$) when controlling for age, pedigree, and clinical laboratory. The overall effect of clinical laboratory was significant (Wald $P < 0.01$) when controlling for age, pedigree, and disease category. Finally, pedigree status was insignificant (Wald $P = 0.15$) when controlling for age, disease category, and clinical laboratory, although as discussed in the statistical Supplementary appendix, this null result could be due to multicollinearity with age (Fig. S3). Figure 4**b** illustrates the effect of disease category and age on diagnostic rate.

## DISCUSSION

We report herein our approach for and diagnostic yield of 1101 patients on a diagnostic odyssey seen at a tertiary medical center. We have implemented a tactic that integrates effectively clinical exome sequencing results with translational research to improve the genetic diagnosis for these patients. Our overall diagnostic yield was 24.9%, and importantly for 36.5% of the patients the activities of translational research were instrumental in reaching a conclusive genetic diagnosis. Establishing a genetic diagnosis was influenced by the age at the time of genetic testing and the disease phenotype of the patient.

The characteristics of the patient population in our cohort have changed over the years, likely affecting the diagnostic rate. In our initial 2016 publication of 51 patients on a diagnostic odyssey, the cohort was 59% pediatric and 63% displayed neurological symptoms with a diagnostic rate of 29%.[10] In the current larger cohort, 49.6% of patients are pediatric and 46.3% are having neurological presentation with a diagnostic rate of 24.9%. This study and others demonstrate that diagnostic utility of exome sequencing is impacted significantly by disease presentation as well as by age at testing.[39,40] Early-onset epilepsy and other neurodevelopmental disorders have higher diagnostic rates (up to 49%),[41,42] whereas as adult-onset disorders (17.5%)[39] and conditions such as inflammatory bowel disease (IBD) (3.4%)[43] have lower diagnostic yields.

Our current diagnostic yield of 24.9% is comparable with other studies.[2,3] This is despite the fact that our study included a majority of adult patients (50.4%) in contrast to previous studies, which were overwhelmingly comprised of pediatric patients.[2,3] Of note, the diagnostic yield of our

pediatric patients, as a subset, is 31.7%. We also realize that our processes, referrals, and expertise have evolved since the implementation of the IMC in 2012. During the first six months of operation of the IMC, the GOB evaluated all the referrals of patients for appropriateness of exome testing and likely diagnostic yield thereof. Subsequently, this step was lifted and the clinicians no longer need GOB approval to proceed with exome testing. Further, Mayo Clinic is a tertiary medical center and therefore over the years the complexity and difficulty of the patients referred, both externally or internally, to the IMC has increased. The evolution of our cohort to encompass an older patient population and a broader as well as more challenging phenotypic spectrum of disease coupled with of increased utilization of exome sequencing in clinical practice could have reduced the overall diagnostic yield over time. Despite these challenges, our diagnostic rate remains stable largely due to the successful activities of the TOP. Integration of the combined expertise and resources of omic data scientists, bioinformaticians, and clinicians to address these diagnostic challenges significantly improves the outcomes obtained from clinical exome sequencing.

In our cohort, there is a clear effect of age after controlling for disease categories, with lower diagnostic rates as the age of the patient at the time of exome testing increases. A clear stepwise influence of disease category on diagnostic yield was also observed. The inverse correlation of diagnostic yield and age at the time of testing could reflect that early-onset disorders may be more likely to be Mendelian, the genetic etiologies may be better understood, or it could be confounded by greater access to trio sequencing in pediatric versus adult patients facilitating de novo variant discovery. Trio versus singleton testing has been previously shown to impact diagnostic rates.[44] However, in our study trio versus singleton versus singleton plus targeted segregation was not statistically significant when controlling for age, disease category, and clinical laboratory. It is possible that age has a dominant effect on its own, but it could be due to multicollinearity of age and testing pedigree, whereby causality cannot be established. From these data we do not conclude that trio sequencing will have the same yield as sequencing a singleton, but it is a factor that is also related to the effect of age and is challenging to assess with this retrospective study design. These results, however, may suggest that different pretest counseling strategies should be considered for patients at different ages to ensure appropriate expectations are set.

This study reports and supports the benefits of an integrated translational research program along with the use of clinical exome sequencing in patient care (Fig. 5). As we know, challenges in exome sequencing interpretation include that (1) the patient's clinical picture is not a perfect representation of the published phenotypic spectrum for a reported genetic finding, (2) VUS or variants in GUS are reported without enough evidence available to definitively determine pathogenicity, and (3) no relevant variation was identified and reported. Adequately addressing these challenges often requires data beyond what is available from a clinical exome sequencing report. Many of these activities of the TOP are "dry lab" efforts that bring forward knowledge to facilitate patient care without the need for "wet lab" activities. These endeavors include data rereview (i.e., patient and variant gene–level data), academic and industry collaborations through contact with published authors, outside laboratories, and tools such as GeneMatcher,[45] or in silico protein modeling. In addition, several patients could benefit from wet lab studies including functional characterization of VUS and in-depth description of GUS to prove the gene–disease phenotype relationship, or novel technologies to identify variation not found with exome sequencing. The latter approaches are often more time-consuming and resource-intensive efforts making them less amenable to broad implementation. However, the advantage of incorporating these translational research activities, under the umbrella of the TOP, into a clinical practice is the unique opportunity to speed the application of novel findings into improved patient care. Unlike the National Institutes of Health's (NIH's) Undiagnosed Disease Network, this approach was implemented such that the translational research is seamlessly embedded into the clinical practice

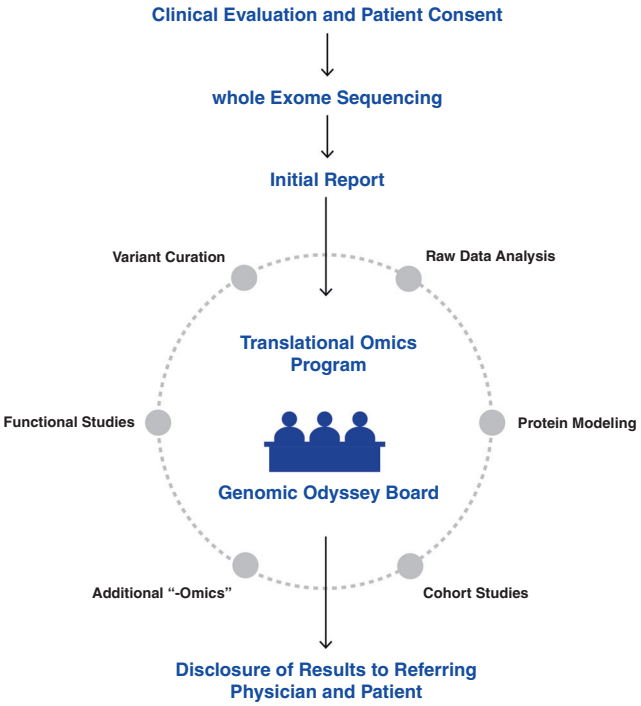


**Fig. 5 Research integration of clinical exome sequencing activities.** Following the report of the clinical exome sequencing findings and consent of the patient the TOP activities are integrated to provide further analysis of the genomic data. These actions include a variety of efforts ranging from variant curation to functional studies. Subsequently, the new scientific findings of the TOP are presented and discussed at the Genomic Odyssey Board (GOB) prior to the latter providing the final recommendation and subsequently the results are disclosed to the referring physician and the patient.

to complement clinical exome testing. This allows the continuity of patient care to be maintained as the results of the clinical sequencing report are enriched by the translational research efforts and subsequently these data are taken directly back to the clinician and patient for better delivery of care.

The TOP was able to provide a definitive diagnosis for 23.6% (100 of 423) of patients studied where clinical exome sequencing was uninformative or not performed. For the remaining 323 individuals studied by the TOP, many continue to have ongoing research conducted on their behalf, but nearly half (49.3%) of undiagnosed patients within this cohort have as yet been studied to a limited degree. Of interest, 37% of patients (408 of 1101) were not referred to the TOP for further study (Fig. 1). This could be for several reasons including lack of patient interest for participating in research, lost to clinical follow-up, or new clinical information leading to a nongenetic diagnosis.

For 96 patients, the TOP has obtained -omics data including entire raw data of clinical exome, genome, RNA-, or methyl-sequencing technologies, but patient-specific analyses have not yet been completed. With these and similar data from the 423 studied patients and those continuing to be seen in the IMC, we continue to improve our analytical pipelines by implementing novel bioinformatics methods from our own design,[15] published methods, or commercial products. Additionally, the use of metabolomics and proteomics offer interesting technologies to further characterize patient conditions and improve diagnostic capabilities. Automated data reanalysis could provide alerts for the availability of new informative data suggesting the reclassification of an identified variant. The implementation of processes like these will be key to fully benefit those patients currently without a genetic diagnosis.

In conclusion, using an integrated approach of clinical exome sequencing along with the scientific contribution of the TOP, 24.9% (274 of 1101) of patients on diagnostic odyssey seen at our tertiary medical center received a genetic diagnosis. The scientific knowledge generated by the TOP increased the diagnostic yield of sole clinical exome sequencing findings by 36.5%. Identification of a genetic diagnosis was influenced by the age at time of testing and the disease phenotype of the patient. This study highlights the value that a translational research program could bring to the clinical care of patients on a diagnostic odyssey. As the field of genomic medicine advances, integrating innovative research activities with standard clinical practice will continue to positively impact genetic diagnoses and therapeutic decision making for these patients.

## SUPPLEMENTARY INFORMATION

The online version of this article (https://doi.org/10.1038/s41436-020-01005-9) contains supplementary material, which is available to authorized users.

## ACKNOWLEDGEMENTS

This research did not receive support from any funding agency in the public, commercial, or not-for-profit sectors.

## DISCLOSURE

The authors declare no conflicts of interest.



## REFERENCES

1. Ashley EA. Towards precision medicine. Nat Rev Genet. 2016;17:507–522.
2. Farwell KD, Shahmirzadi L, El-Khechen D, et al. Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. Genet Med. 2015;17:578–586.
3. Posey JE, Harel T, Liu P, et al. Resolution of disease phenotypes resulting from multilocus genomic variation. N Engl J Med. 2017;376:21–31.
4. Gahl WA, Wise AL, Ashley EA. The Undiagnosed Diseases Network of the National Institutes of Health: a national extension. JAMA. 2015;314:1797–1798.
5. Gahl WA, Markello TC, Toro C, et al. The National Institutes of Health Undiagnosed Diseases Program: insights into rare diseases. Genet Med. 2012;14:51–59.
6. Taruscio D, Groft SC, Cederroth H, et al. Undiagnosed Diseases Network International (UDNI): white paper for global actions to meet patient needs. Mol Genet Metab. 2015;116:223–225.
7. Li J, Gao K, Yan H, et al. Reanalysis of whole exome sequencing data in patients with epilepsy and intellectual disability/mental retardation. Gene. 2019;700:168–175.
8. Splinter K, Adams DR, Bacino CA, et al. Effect of genetic diagnosis on patients with previously undiagnosed disease. N Engl J Med. 2018;379:2131–2139.
9. Lazaridis KN, McAllister TM, Babovic-Vuksanovic D, et al. Implementing individualized medicine into the medical practice. Am J Med Genet C Semin Med Genet. 2014;166C:15–23.
10. Lazaridis KN, Schahl KA, Cousin MA, et al. Outcome of whole exome sequencing for diagnostic odyssey cases of an individualized medicine clinic: the Mayo Clinic experience. Mayo Clin Proc. 2016;91:297–307.
11. Cummings BB, Marshall JL, Tukiainen T, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. Sci Transl Med. 2017;9:eaal5209.
12. Boczek NJ, Hopp K, Benoit L, et al. Characterization of three ciliopathy pedigrees expands the phenotype associated with biallelic C2CD3 variants. Eur J Hum Genet. 2018;26:1797–1809.
13. Cousin MA, Conboy E, Wang JS, et al. RINT1 bi-allelic variations cause infantile-onset recurrent acute liver failure and skeletal abnormalities. Am J Hum Genet. 2019;105:108–121.
14. Oliver GR, Blackburn PR, Ellingson MS, et al. RNA-Seq detects a SAMD12-EXT1 fusion transcript and leads to the discovery of an EXT1 deletion in a child with multiple osteochondromas. Mol Genet Genomic Med. 2019;7:e00560.
15. Oliver GR, Tang X, Schultz-Rogers LE, et al. A tailored approach to fusion transcript identification increases diagnosis of rare inherited disease. PLoS One. 2019;14:e0223337.
16. Cousin MA, Smith MJ, Sigafoos AN, et al. Utility of DNA, RNA, protein, and functional approaches to solve cryptic immunodeficiencies. J Clin Immunol. 2018;38:307–319.
17. Blackburn PR, Tischer A, Zimmermann MT, et al. A novel Kleefstra syndrome-associated variant that affects the conserved TPLX motif within the Ankyrin repeat of EHMT1 leads to abnormal protein folding. J Biol Chem. 2017;292:3866–3876.
18. Gupta A, Zimmermann MT, Wang H, et al. Molecular characterization of known and novel ACVR1 variants in phenotypes of aberrant ossification. Am J Med Genet A. 2019;179:1764–1777.
19. Perez Botero J, Chen D, Cousin MA, et al. Clinical characteristics and platelet phenotype in a family with RUNX1 mutated thrombocytopenia. Leuk Lymphoma. 2017;58:1963–1967.
20. Blackburn PR, Xu Z, Tumelty KE, et al. Bi-allelic alterations in AEBP1 lead to defective collagen assembly and connective tissue structure resulting in

a variant of Ehlers-Danlos syndrome. Am J Hum Genet. 2018;102: 696–705.
21. Pinto EVF, Kroc SA, Bertsch NL, et al. Biallelic variants in PROZ as a cause of hypercoagulability and livedo racemosa. Thromb Res. 2020;195:187–189.
22. Gonzalez Santiago TM, Zavialov A, Saarela J, et al. Dermatologic features of ADA2 deficiency in cutaneous polyarteritis nodosa. JAMA Dermatol. 2015;151:1230–1234.
23. Kotwal A, Ferrer A, Kumar R, et al. Clinical and biochemical phenotypes in a family with ENPP1 mutations. J Bone Miner Res. 2020;35:662–670.
24. Blackburn PR, Sullivan AE, Gerassimou AG. et al. Functional analysis of the SIM1 variant p.G715V in 2 patients with obesity. J Clin Endocrinol Metab. 2020;105:dgz192.
25. Salpietro V, Dixon CL, Guo H, et al. AMPA receptor GluA2 subunit defects are a cause of neurodevelopmental disorders. Nat Commun. 2019;10:3094.
26. Weyhrauch DL, Ye D, Boczek NJ, et al. Whole exome sequencing and heterologous cellular electrophysiology studies elucidate a novel loss-of-function mutation in the CACNA1A-encoded neuronal P/Q-type calcium channel in a child with congenital hypotonia and developmental delay. Pediatr Neurol. 2016;55:46–51.
27. Helbig KL, Lauerer RJ, Bahr JC, et al. De novo pathogenic variants in CACNA1E cause developmental and epileptic encephalopathy with contractures, macrocephaly, and dyskinesias. Am J Hum Genet. 2018; 103:666–678.
28. Lu HC, Tan Q, Rousseaux MW, et al. Disruption of the ATXN1-CIC complex causes a spectrum of neurobehavioral phenotypes in mice and humans. Nat Genet. 2017;49:527–536.
29. Konrad EDH, Nardini N, Caliebe A, et al. CTCF variants in 39 individuals with a variable neurodevelopmental disorder broaden the mutational and clinical spectrum. Genet Med. 2019;21:2723–2733.
30. Pant DC, Dorboz I, Schluter A, et al. Loss of the sphingolipid desaturase DEGS1 causes hypomyelinating leukodystrophy. J Clin Invest. 2019;129: 1240–1256.
31. Zimmermann MT, Urrutia R, Cousin MA, et al. Assessing human genetic variations in glucose transporter SLC2A10 and their role in altering structural and functional properties. Front Genet. 2018;9:276.
32. Kaiwar C, Zimmermann MT, Ferber MJ, et al. Novel NR2F1 variants likely disrupt DNA binding: molecular modeling in two cases, review of published cases, genotype-phenotype correlation, and phenotypic expansion of the Bosch-Boonstra-Schaaf optic atrophy syndrome. Cold Spring Harb Mol Case Stud. 2017;3:a002162.
33. Blackburn PR, Barnett SS, Zimmermann MT, et al. Novel de novo variant in EBF3 is likely to impact DNA binding in a patient with a neurodevelopmental disorder and expanded phenotypes: patient report, in silico functional assessment, and review of published cases. Cold Spring Harb Mol Case Stud. 2017;3:a001743.
34. Wiltrout K, Ferrer A, van de Laar I, et al. Variants in DOCK3 cause developmental delay and hypotonia. Eur J Hum Genet. 2019;27:1225–1234.
35. Gupta A, Dsouza NR, Zarate YA, et al. Genetic variants in DGAT1 cause diverse clinical presentations of malnutrition through a specific molecular mechanism. Eur J Med Genet. 2020;63:103817.
36. Masnada S, Hedrich UBS, Gardella E, et al. Clinical spectrum and genotype-phenotype associations of KCNA2-related encephalopathies. Brain. 2017;140:2337–2354.
37. Nambot S, Faivre L, Mirzaa G, et al. De novo TBR1 variants cause a neurocognitive phenotype with ID and autistic traits: report of 25 new individuals and review of the literature. Eur J Hum Genet. 2020;28:770–782.
38. Ferrer A, Schultz-Rogers L, Kaiwar C. et al. Three rare disease diagnoses in one patient through exome sequencing. Cold Spring Harb Mol Case Stud. 2019;5:a004390.
39. Posey JE, Rosenfeld JA, James RA, et al. Molecular diagnostic experience of whole-exome sequencing in adult patients. Genet Med. 2016;18:678–685.
40. Retterer K, Juusola J, Cho MT, et al. Clinical application of whole-exome sequencing across clinical indications. Genet Med. 2016;18:696–704.
41. Cherot E, Keren B, Dubourg C, et al. Using medical exome sequencing to identify the causes of neurodevelopmental disorders: experience of 2 clinical units and 216 patients. Clin Genet. 2018;93:567–576.
42. Tumiene B, Maver A, Writzl K, et al. Diagnostic exome sequencing of syndromic epilepsy patients in clinical practice. Clin Genet. 2018;93: 1057–1062.
43. Serra EG, Schwerd T, Moutsianas L, et al. Somatic mosaicism and common genetic variation contribute to the risk of very-early-onset inflammatory bowel disease. Nat Commun. 2020;11:995.
44. Ji J, Shen L, Bootwalla M, et al. A semiautomated whole-exome sequencing workflow leads to increased diagnostic yield and identification of novel candidate variants. Cold Spring Harb Mol Case Stud. 2019;5:a003756.
45. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat. 2015;36:928–930.

Eric W. Klee, PhD[1,2,3], Margot A. Cousin, PhD[1,2], Filippo Pinto e Vairo, MD, PhD[2,3], Joel A. Morales-Rosado, MD[1,2], Erica L. Macke, PhD[1,2], W. Garrett Jenkinson, PhD, MS[1,2], Alejandro Ferrer, PhD[1,2], Laura E. Schultz-Rogers, PhD[1,2], Rory J. Olson, PhD[1,2], Gavin R. Oliver, MS[1,2], Ashley N. Sigafoos, BS[2,4], Tanya L. Schwab, MS[2,4], Michael T. Zimmermann, PhD[5], Raul A. Urrutia, MD[6], Charu Kaiwar, MD, PhD[7], Aditi Gupta, PhD[2], Patrick R. Blackburn, PhD[2,8], Nicole J. Boczek, PhD[2,8], Carri A. Prochnow, BS[2], Rebecca J. Lowy, BS[2], Lindsay A. Mulvihill, CCRP[2], Tammy M. McAllister, MA[2], Stacy L. Aoudia, MBA[2], Teresa M. Kruisselbrink, MS, LCGC[2], Lauren B. Gunderson, MS, LCGC[3], Jennifer L. Kemppainen, MS, LCGC[2,3], Laura J. Fisher, MS, LCGC[3], Jessica M. Tarnowski, MS, LCGC[3], Megan M. Hager, MS, CGC[9], Sarah A. Kroc, MS, LCGC[3], Nicole L. Bertsch, MS, LCGC[3], Katherine E. Agre, MS, LCGC[3], Jessica L. Jackson, MS, LCGC[10], Sarah K. Macklin-Mantia, MS, LCGC[10], Marine I. Murphree, MS, LCGC[3], Laura M. Rust, MS, LCGC[3], Jolene M. Summer Bolster, MA[2], Scott A. Beck, BBA, MBA[2], Paldeep S. Atwal, MB, ChB[10,11], Marissa S. Ellingson, MS, LCGC[8], Sarah S. Barnett, MS, LCGC[8], Kristen J. Rasmussen, MS, LCGC[8], Carrie A. Lahner, MS, LCGC[8], Zhiyv Niu, PhD[3,8], Linda Hasadsri, MD, PhD[8], Matthew J. Ferber, PhD[2], Cherisse A. Marcou, PhD[8], Karl J. Clark, PhD[2,4], Pavel N. Pichurin, MD[2,3], David R. Deyle, MD[2,3], Eva Morava-Kozicz, MD, PhD[2,3], Ralitza H. Gavrilova, MD[2,3], Radhika Dhamija, MBBS[7,9], Klaas J. Wierenga, MD[10,11], Brendan C. Lanpher, MD[2,3], Dusica Babovic-Vuksanovic, MD[2,3], Gianrico Farrugia, MD[2,12], Lisa A. Schimmenti, MD[2,3], A. Keith Stewart, MB, ChB[13] and Konstantinos N. Lazaridis, MD[2,12]

[1]Department of Health Sciences Research, College of Medicine, Mayo Clinic, Rochester, MN, USA; [2]Center for Individualized Medicine, College of Medicine, Mayo Clinic, Rochester, MN, USA; [3]Department of Clinical Genomics, Mayo Clinic, Rochester, MN, USA; [4]Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN, USA; [5]Bioinformatics Research and Development Laboratory, Genomic Sciences and Precision Medicine Center, Medical College of Wisconsin, Milwaukee, WI, USA; [6]Division of Research, Department of Surgery and the Genomic Sciences and Precision Medicine Center, Medical College of Wisconsin, Milwaukee, WI, USA; [7]Center for Individualized Medicine, Mayo Clinic, Scottsdale, AZ, USA; [8]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA; [9]Department of Clinical Genomics, Mayo Clinic, Scottsdale, AZ, USA; [10]Department of Clinical Genomics, Mayo Clinic, Jacksonville, FL, USA; [11]Center for Individualized Medicine, Mayo Clinic, Jacksonville, FL, USA; [12]Division of Gastroenterology and Hepatology, College of Medicine, Mayo Clinic, Rochester, MN, USA; [13]Princess Margaret Cancer Centre, Toronto, ON, Canada.