Precision medicine integrating whole-genome sequencing, comprehensive metabolomics, and advanced imaging

Ying-Chen Claire Hou, Hung-Chun Yu, Rick Martin, Elizabeth T. Cirilli, Natalie M. Schenker-Ahmed, Michael Hicks, Isaac V. Cohen, Thomas J. Jönsson, Robyn Heister, Lori Napier, Christine Leon Swisher, Saints Dominguez, Haibao Tang, Weizhong Li, Bradley A. Perkins, Jaime Barea, Christina Rybak, Emily Smith, Keegan Duchicela, Michael Doney, Pamela Bray, Nathaniel Hernandez, Ewen F. Kirkness, Andrew M. Kahn, J. Craig Venter, David S. Karow, and C. Thomas Caskey

Genome sequencing has established clinical utility for rare disease diagnosis. While increasing numbers of individuals have undergone elective genome sequencing, a comprehensive study surveying genome-wide disease-associated genes in adults with deep phenotyping has not been reported. Here we report the results of a 3-y precision medicine study with a goal to integrate whole-genome sequencing with deep phenotyping. A cohort of 1,190 adult participants (402 female [33.8%]; mean age, 54 y [range 20 to 89+]; 70.6% European) had whole-genome sequencing, and were deeply phenotyped using metabolomics, advanced imaging, and clinical laboratory tests in addition to family/medical history. Of 1,190 adults, 206 (17.3%) had at least 1 genetic variant with pathogenic (P) or likely pathogenic (LP) assessment that suggests a predisposition of genetic risk. A multidisciplinary clinical team reviewed all reportable findings for the assessment of genotype and phenotype associations, and 137 (11.5%) had genotype and phenotype associations. A high percentage of genotype and phenotype associations (>75%) was observed for dyslipidemia (n = 24), cardiomyopathy, arrhythmia, and other cardiac diseases (n = 42), and diabetes and endocrine diseases (n = 17). A lack of genotype and phenotype associations, a potential burden for patient care, was observed in 69 (5.8%) individuals with P/LP variants. Genomics and metabolomics associations identified 61 (5.1%) heterozygotes with phenotype manifestations affecting serum metabolite levels in amino acid, lipid, and cofactor, and vitamin pathways. Our descriptive analysis provides results on the integration of whole-genome sequencing and deep phenotyping for clinical assessments in adults.

Significance

To understand the value and clinical impact of surveying genome-wide disease-causing genes and variants, we used a prospective cohort study design that enrolled volunteers who agreed to have their whole genome sequenced and to participate in deep phenotyping using clinical laboratory tests, metabolomics technologies, and advanced noninvasive imaging. The genomic results are integrated with the phenotype results. Approximately 1 in 6 adult individuals (17.3%) had genetic findings and, when integrated with deep phenotyping data, including family/medical histories with genetic findings, 1 in 9 (11.5%) had genotype and phenotype associations. Genomics and metabolomics association analysis revealed 5.1% of heterozygotes with phenotype manifestations affecting serum metabolite levels. We report observations from our study in which health outcomes and benefits were not measured.

Author contributions: J.C.V. conceived the study; C.T.C. led the analyses; Y.-C.C.H., H.-C.Y., R.M., and C.T.C. performed genomics analyses and genetic-phenotypic integration analysis; N.M.S.-A., C.L.S., S.D., N.H., A.M.K., and D.S.K. evaluated clinical imaging; Y.-C.C.H., H.-C.Y., R.M., R.H., J.B., C.R., E.S., K.D., M.D., P.B., and C.T.C. contributed to clinical genetics evaluation; Y.-C.C.H., H.-C.Y., E.T.C., I.V.C., and T.J.J. performed metabolomics analysis; L.N. contributed to data collection; M.H., H.T., W.L., and E.F.K. contributed to bioinformatics; R.M. R.H., J.B., C.R., E.S., K.D., M.D., P.B., A.M.K., D.S.K., and C.T.C. contributed to clinical support; B.A.P., D.S.K., and C.T.C supervised research; and Y.-C.C.H., H.-C.Y., N.M.S.-A., and C.T.C. wrote the paper.

Competing interest statement: Y.-C.C.H., H.-C.Y., N.M.S.-A., R.H., L.N., S.D., K.D., P.B., N.H., A.M.K., J.C.V., D.S.K., and C.T.C. are employees of Human Longevity, Inc.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

To whom correspondence may be addressed. Email: tcaskey@bcm.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1909378117/-/DCSupplemental.

First published January 24, 2020.

www.pnas.org/cgi/doi/10.1073/pnas.1909378117

PNAS | February 11, 2020 | vol. 117 | no. 6 | 3053–3062
Understanding the functional consequence of genomic variation has been challenging, and numerous approaches have been employed. Molecular technologies, including metabolomics (metabolites), transcriptomics (RNA), proteomics (proteins), and epigenomics, have been employed to interpret the functional consequence of genomic variations (14–17). In particular, the diagnosis of monogenic conditions in pediatric cases has been transformed by methods that allow interrogation of biochemical and genetic data for discoveries of new associations between metabolic disorders and genes (18–20). From large-scale genome studies, the use of extensive phenotypic data in EHRs and identification of loss-of-function (LoF) variants from exome-sequencing studies, the use of extensive phenotypic data in EHRs and identification of loss-of-function (LoF) variants from exome-sequencing data have improved our understanding of previously undiscovered biological functions for genes and the development of therapeutic targets (12, 21, 22).

To understand the value and impact of surveying genome-wide disease-causing genes and variants integrated with deep phenotyping, we used a prospective cohort design enrolling volunteers under a research protocol. The deep phenotyping included family history, past and current personal medical history, clinical laboratory tests, advanced noninvasive imaging, and metabolomics technologies. The study objectives were fourfold. First, we evaluated genotype and phenotype associations in adults in various disease areas, including cancer, cardiomyopathy, arrhythmia, and other cardiac diseases, dyslipidemia, diabetes and endocrine, chronic liver, hematology, inborn errors of metabolism, and other disorders. Second, we showed cases where the lack of genotype and phenotype associations may result in possible ambiguous results for patient care from surveying genome-wide disease-causing variants in adults with elective genome sequencing. Third, we interrogated care from surveying genome-wide disease-causing variants in associations may result in possible ambiguous results for patient care and provided clinical guidelines for further management. Finally, we observed cases for autosomal recessive carriers with a phenotype manifestation in imaging or metabolome. We investigated gene associations with serum metabolite changes and cholesterol homeostasis.

Results

Phenotype Test Findings. The cohort was composed of 1,190 self-referred volunteers with a median age of 54 y (range 20 to 89 + y, 33.8% female, 70.6% European). The demographic information of the cohort is shown in Table 1, and previously identified conditions (%) included cancer (11.0%), coronary heart disease (4.8%), diabetes (3.8%), chronic liver diseases (5.1%), and neurological disorders (10.2%). Our cohort had no enrichment of frequent adult chronic diseases compared with National Health and Nutrition Examination Survey (NHANES) adults, a US population-based sample. This study is an expansion of our pilot study of 209 study participants (19). We added noninvasive computed tomography (CT) of the heart to measure the amount of calcified plaque in the coronary arteries as a means of evaluating risk of coronary artery disease. The dual-energy X-ray absorptiometry test was removed. Detailed protocols used for whole-body MRI are listed in SI Appendix, Table S1. The criteria for cardiovascular reportable findings are listed in SI Appendix, Table S2 and reference ranges for clinical tests are provided in SI Appendix, Table S3. Each test was evaluated and assessed by board-certified physicians. Reportable findings of MRI, echocardiography (ECHO), electrocardiography (ECG), continuous cardiac monitoring (CCM), and CT are listed in SI Appendix, Tables S4–S7.

A heatmap representation of reportable findings from quantitative phenotype measurements of each study participant is provided in Fig. 1B in chronological order. Except for the CT test that was added after the pilot study, study participants had choices to omit certain tests based on medical decisions or personal preference; omissions are highlighted in gray in Fig. 1B. Phenotype testing revealed 407 (34.2%) participants had insulin resistance and/or impaired glucose tolerance; 343 (29.2%) participants had elevated liver fat (>4%); 193 (16.2%) participants had cardiac structure or function abnormalities; 136 (11.4%) participants had coronary heart disease risk (Agatston score > 100, relative risk [RR] 4.3) (23); 104 (9.3%) participants had elevated liver

| Table 1. Study characteristics |
|--------------------------------|
| Characteristics               | This study | NHANES |
| Age, y                        | 54 (20 to 89+) |
| Sex, %                        |               |
| Male                          | 66.2         |
| Female                        | 33.8         |
| Measured BMI, median (25 to 75%) | 25.3 (23.0 to 28.0) |
| Measured systolic blood pressure, median (25 to 75%) | 124 (115 to 136) |
| Measured LDL, median (25 to 75%) | 114 (92 to 137) |
| Diseases                      |               |
| Neoplasms, % (ever told you had cancer or malignancy) | 11.0         |
| Cardiovascular, % (ever told you had coronary heart disease) | 4.8          |
| Chronic respiratory diseases, % (ever told you had COPD) | 2.7          |
| Diabetes, % (ever told you have diabetes) | 3.8          |
| Chronic liver diseases, % (ever told you had liver diseases) | 5.1          |
| Neurological disorders, % (blood relatives have Alzheimer's disease or dementia) | 10.2         |
| Risk factors                  |               |
| Alcohol use, %                | 49.5         |
| Tobacco smoking, %            | 19.6         |
| High cholesterol, %           |               |
| Ever told you had high cholesterol | 31.5         |
| Taking prescription for high cholesterol | 18.2         |
| High blood pressure, %        |               |
| Ever told you had high blood pressure | 19.0         |
| Taking prescription for hypertension | 14.6         |

The NHANES information is at https://www.cdc.gov/nchs/nhanes/. COPD, chronic obstructive pulmonary disease.
iron (R2* > 80); 73 (6.1%) participants had arrhythmia; 57 (4.8%) participants had cardiac conduction disorders; and 23 (2.0%) showed low hippocampal occupancy score (≤0.65). We identified 20 (1.7%) individuals with early-stage neoplasia, prostate adenocarcinoma, renal cell carcinoma, lymphoma, transitional cell carcinoma, papillary thyroid cancer, pancreatic cancer, neurofibromatosis, and mediastinal thymoma. Among these, 12 cases were confirmed with a biopsy, 4 cases were confirmed with CT, 3 cases were surgically removed, and 1 case was confirmed by the genetic finding (SI Appendix, Table S4).

The clinical laboratory tests that assessed cancer and liver, kidney, hematol, endocrin, immunol, and lipid functions are shown in SI Appendix, Fig. S1.

We also calculated the age distribution of participants with reportable findings compared with those without reportable findings in each of the tests. The age distribution of each test is provided in SI Appendix, Fig. S2. The median age of participants with genetic findings and extreme metabolome findings was similar compared with participants without findings (P > 0.05). The median ages (interquartile range) of participants with reportable findings in ECHO, CT, and CCM tests were ECHO: 62 (55 to 70); CT: 65 (57 to 70); and CCM: 64 (57 to 70). The median ages of participants with reportable findings in MIRI-body, MIRI-brain, and MRI-cancer diagnosed in this study were MIRI-body: 55 (47 to 64); MIRI-brain: 70 (52 to 76); and MRI-cancer: 64.5 (56 to 71).

Identification of Sequence Variants with Pathogenicity. To understand the value and impact of surveying genome-wide disease-causing genes and variants, we used a 2-step process described in Materials and Methods to filter and manually interpret each disease-causing variant per participant. We followed the American College of Medical Genetics and Genomics/Association for Molecular Pathology guidelines for the interpretation of sequence variants. Our 2-step process evaluates variants in 30,281 unique genes present in ClinVar, Human Gene Mutation Database (HGMD), and Online Mendelian Inheritance in Man (OMIM). In our cohort, 206 (17.3%) participants had at least one medically significant genetic finding (MSF) with a clinical significance of pathogenic (P), likely pathogenic (LP), or variant of unknown pathogenicity (VUS-R). The medically significant findings included heterozygotes in autosomal dominant or X-linked recessive conditions or homozygotes or compound heterozygotes in autosomal recessive conditions. The most commonly affected genes (number of cases) for autosomal dominant conditions were CHEK2 (17), MYBPC3 (8), BRCA2 (7), ATM (6), HOXB13 (6), and LDLR (6). For autosomal recessive conditions, the most commonly affected genes were HFE (15), BTD (5), and GJB2 (2). For X-linked recessive conditions, we identified 2 cases in the G6PD gene. In the ACMG 59 genes, 29 individuals (2.4%) had P or LP variants identified. Frequently observed risk alleles, including F2 c.97G>A, F5 Leiden c.1601G>A, and ALDH2 c.1510G>A variants, were observed in 83 individuals (7%) of this cohort.

The assertion of pathogenicity of a genomic sequence variant is not equivalent to the clinical diagnosis of genetic disorders. Our multidisciplinary team integrated the genomic findings and evaluated whether a participant’s clinical characteristics and family/personal history provided sufficient evidence to support clinical diagnosis of genetic disorders. Clinical characteristics listed in OMIM clinical synopses (24) or publications such as GeneReviews or primary studies were employed to evaluate genotype–phenotype associations. For example, for a participant with a pathogenic variant observed in the LDLR gene, the team would review his/her results in clinical laboratory tests and metabolome for profiles that are indicators of impaired glucose tolerance or impaired insulin sensitivity, and MIRI test for renal cysts. Genotype and phenotype associations were identified in 137 (11.5%) participants among 206 participants with at least 1 MSF (Fig. 2).

Dyslipidemia, cardiomyopathy and arrhythmia, and diabetes and endocrine diseases had associations between genotype and phenotype higher than 75% (24/32 [100%], 36/42 [85.7%], and 13/17 [76.5%), respectively). The details of genes, variants, and associated phenotypes are provided in Dataset S1 (MSF and Phenotype Associations) for 206 participants with MSFs.

Integration of advanced imaging and genomic sequencing led to the clinical diagnosis of genetic disorders which were not established previously (Fig. 3). In individuals with the HNF1B, PKD1, PCSK9, MYH7, NFI, and CFTR variants (Fig. 3 A–G), whole-body (WB)-MRI, CT, and genetic findings of variants with pathogenicity provided evidence to support the diagnosis of inherited genetic diseases. A diagnosis of renal cysts and diabetes syndrome was made in the participant with a likely pathogenic variant in the HNF1B gene without family/personal medical history. The imaging findings found multiple subcentimeter cysts in kidney bilaterally (Fig. S4). The metabolome-based clinical tests showed impaired glucose tolerance and insulin sensitivity. The detection of a pathogenic variant in the NFI gene (de novo) was seen in an individual with optic glioma, white matter lesion, stenosis in the middle cerebral artery, and Moyamoya syndrome (Fig. 3 E and F). A diagnosis of cystic fibrosis was made in a mid-50s male participant. He had a history of digestive symptoms and chronic sinus infections and the imaging findings found tree-in-bud nodularity at the right mid and upper lung zone and mild bronchial wall thickening (Fig. 3G). Two variants, c.3846G>A (p.Thr1282*) and c.3454G>C (p.Asp1152His), were identified in the CFTR gene. Results from genetics and clinical presentations were consistent with atypical cystic fibrosis. We selected 12 cases to illustrate the data integration of our tests for clinical assessments (Fig. 3I). The integration of genomic data assisted new clinical diagnosis of genetic diseases.

Prevalence of P/LP Variants without Disease Symptoms or Family History. The identification of a genomic variant with assertion of pathogenicity does not imply a clinical diagnosis of inherited genetic disorders. Integrating genomic data with deep phenotyping presented an opportunity to further understand the lack of genotype and phenotype associations which may result in uncertainty in patient care. We report the prevalence of the P/LP variants per disease category (Table 2). For example, after carefully examining data from family history, past medical history, and advanced imaging tests, 30 out of 69 participants with P/LP variants associated with cancer predisposition did not have corresponding family history and phenotypes at the time of evaluation. The most commonly detected cancer variants (number of cases) were HOXB13 c.251G>A (6), CHEK2 c.470T>C (5), CHEK2 c.1283C>T (2), and FH c.1431_1433dupAAA (2). Three homozygotes and 2 compound heterozygotes with P/LP variants in the BTD gene did not have expected elevated beta-hydroxyisovalerate and lactate phenotypes. Four compound heterozygotes with P/LP variants in the HFE gene also did not have expected elevated liver iron (R2*), ferritin, and iron phenotypes.

Autosomal Recessive Genetic Variants with Phenotype Manifestations. Using the genome-wide approach, 1,027 (86.3%) individuals had at least 1 autosomal recessive variant in 680 genes (Dataset S1, Autosomal Recessive Conditions and Autosomal Recessive Variants). The most commonly detected genes (number of cases) were SERPINA1 (99), FLC (95), BTD (84), HFE (84), and GJB2 (71) (SI Appendix, Table S8). Eighteen genes (>1% observed frequency) not commonly found in gene panels for prenatal carrier
screening were identified (SI Appendix, Table S8). We observed that carriers of autosomal recessive conditions had phenotype manifestations. Findings of PKDH1 (Fig. 3H) (autosomal recessive polycystic kidney disease), ALPL (hypophosphatasia), and LMBRD1 (methylmalonic aciduria and homocystinuria) are shown in Table 3. For hemochromatosis, 18% (12/68) of heterozygotes for the HFE p.Cys282Tyr variant had high R2*, a marker of liver iron content, compared with only 8% (87/1,034) of individuals with normal genotype (P = 0.0156, Fisher exact test), suggesting iron regulation is compromised. One participant with a low 10-y risk Framingham score (<5%) (25) had an Agatston score of 2,963 (RR 10.8) at the 99th percentile for people of the same age, sex, and race/ethnicity per the Multi-Ethnic Study of Atherosclerosis (25, 26). In this participant’s WGS, we identified an LoF variant, c.63dupA (p.Leu22fs), in the LMBRD1 gene in the heterozygous state with phenotypic manifestation of elevated homocysteine. Elevated homocysteine has been associated with vascular calcification (27). In our cohort, we observed an approximately threefold increased risk (OR 2.76, 95% CI [1.4001 to 5.4413]; P = 0.03) of coronary artery calcification (CAC > 1) in individuals with high homocysteine (>15 mmol/L). The same variant was detected in 3 second-degree relatives with high CAC scores (>400, 89, and 91 percentiles, respectively) found in 2 relatives, and one of them had elevated homocysteine and the other one did not complete the homocysteine test. Phenotypes of methylmalonic aciduria and homocystinuria, cblF type (MIM 277380), including failure to thrive, developmental delay, stomatitis, and skin rashes were not observed in the LMBRD1 heterozygotes. Other unknown factors may also contribute to this family’s coronary artery calcification.

Genotype and Metabolome Associations. To understand the aberrant levels of any of the 1,007 measured metabolites that are associated with physiological states of health conditions, we investigated the associations between functional genetic variants and metabolite level. We detected extreme metabolites (±6 SD) in 17.3% of individuals. A portion of the underlying mechanisms that contribute to extreme metabolites (±6 SD) is due to hormone aberrations and medication/supplement intakes. The
curated data of gene and metabolite relationships listed in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (28), Human Metabolome Database (HMDB) (29), and proprietary pathway database constructed by Metabolon, Inc. were used to derive candidate lists of genes whose biological functions have been implicated in the regulation of metabolites of Fig. 2.

**Fig. 2.** Associations between medically significant genetic findings and phenotype tests by disease group and test. (A) Numbers of cases (gene [cases]) per disease categories are listed. (B) Clinical synopses in OMIM and clinical phenotypes listed in the literature are used to establish associations between MSFs and phenotype tests. The heatmap shows the fraction of cases (percentage listed) where phenotype tests detected expected clinical features for the MSF. For example, for a participant with a pathogenic variant observed in the HNF1B gene, the team would review his/her results in clinical laboratory tests and metabolome for profiles that are indicators of impaired glucose tolerance or impaired insulin sensitivity and an MRI test for renal cysts. Thus, he/she may have reportable findings for MRI, family history, metabolome, and clinical laboratory tests; however, the phenotype and genotype association still counts as 1 event for the bar plot. Among the 206 individuals, 21 individuals have more than 1 MSF in different disease groups. The percentages of “not tested” for the participants with MSFs are 41.7% for CT, 38.8% for metabolome, 24.2% for ECG/CCM, 12.6% for clinical laboratory test, and 0% for MRI and ECHO. FHx, family history; PHx, past and current personal medical history.
interest. We also took the directional effect of enzymatic reactions into consideration. Then, genomic data were investigated to identify functional genetic variants. We identified the associations between functional genetic variation and extreme metabolite levels in 31 (2.6%) individuals, mostly enriched in amino acid, lipids, and nucleotide pathways (SI Appendix, Table S9).

| Case 1 | F, age 50s D: None |
|--------|----------------------|
| Case 2 | F, age 50s D: None |
| Case 3 | F, age 60s D: 1. Thalassemia. 2. Renal cysts and diabetes |
| Case 4 | M, age 40s D: Polycystic kidney disease |
| Case 5 | M, age 50s D: Familial hypercholesterolemia |
| Case 6 | M, age 60s D: Familial hypertrophic cardiomyopathy |
| Case 7 | M, age 50s D: Hereditary cancer-predisposing syndrome |
| Case 8 | M, age 70s D: 1. Hemochromatosis. 2. Renal cell carcinoma |
| Case 9 | M, age 60s D: Long QT |
| Case 10 | M, age 60s D: 1. Neurofibromatosis. 2. Vitamin D deficiency |
| Case 11 | M, age 40s D: 1. Thrombosis. 2. Hypercholesterolemia 3. Cardiomyopathy 4. Frontal lobe encephalomalacia |
| Case 12 | M, age 60s D: Cystic fibrosis. 2. Digestive symptoms 3. Sinus infection |

Blue: New diagnosis from this study
Abnormal/Positive  Normal/Negative  Not Tested

Fig. 3. Examples of integrated diagnoses. (A) A female (40s) heterozygous for an HNF1B likely pathogenic variant with renal cysts and diabetes syndrome. Yellow arrows point to the bilateral renal cysts. (B) A male (40s) heterozygous for a likely pathogenic PKD1 variant with polycystic kidney disease. (C) A male (50s) heterozygous for a pathogenic variant in the PCSK9 gene. The yellow arrow points to the calcified left main/anterior descending coronary artery. (D) A male (60s) compound heterozygous for CFTR pathogenic variants with digestive issues/bloating and chronic sinus infections. The green circle shows a tree-in-bud nodularity. (H) A female (60s) heterozygous for a pathogenic PKDH1 variant with a defined liver cyst (yellow arrow) and numerous scattered subcentimeter liver cysts/hemangiomas (white arrows). (I) Summary of clinical findings in represented cases. Dx, diagnosis; eGFR, estimated glomerular filtration rate; F, female; LV, left ventricular; M, male; RV, right ventricular.
guidelines. The rest of the variants listed in SI Appendix, Table S9 were classified as VUS. We also evaluated the phenotypic manifestation for individuals who carried P/LP variants associated with metabolic disorders using metabolic feature and laboratory abnormalities listed in the OMIM synopsis (24). This approach established associations in 34 (2.9%) individuals with elevated/decreased metabolites (95% CI) (Table S10) enriched in amino acid, cofactor and vitamin, and lipid pathways. The genetic basis of some metabolite variation was identified. An interesting case illustrating the value of this approach was the demonstration of an association of an LoF nonsense variant, c.13C>T (p.Arg5*), in the TTPA gene associated with ataxia with isolated vitamin E deficiency (MIM 277460). Three family members had the maternally segregated variant, and all had markedly reduced levels of vitamin E ranging from −3.6 to −6 SD. Clinical ataxia was not observed in these participants.

Genes Associated with Metabolites and the Ratio of Cholesterol and Hydroxy-3-Methylglutarate. The collection of deep phenotype data allowed us to pursue additional research avenues. To expand our sample set for these analyses, we also included metabolome data from 1,969 European ancestry twins enrolled in the TwinsUK Registry, a British national register of adult twins (30). We used a gene-based collapsing analysis to identify genes where rare functional variants were associated with a statistically significant difference in the levels of any of the 1,007 measured metabolites. We identified significant associations between the phenylketonuria gene PAH and the metabolites phenylalanine and gammaglutamylphenylalanine as well as between the glutaric acidemia gene ETFDH and octanoylcarnitine, decanoylcarnitine, and nonanoylcarnitine (Fig. 4). We identified 19 significant associations with 11 other genes (SI Appendix, Table S11). The identified associations largely reflected known conditions, such as dimethylglycine dehydrogenase deficiency and histidinemia. However, 5 associations were identified: between the 1,5-anhydroglucitol and SLC5A10; between the amino acids alpha-hydroxyisovalerate and 2-hydroxy-3-methylvalerate and HAO2; between the amino acid 5-hydroxylysine and HYYK; between N-acetyl-beta-alanine and PTER; and between alpha-keto glutarate and NIT2. NIT2 is known to break down alpha-keto glutarate, and we have identified LoF variants.

To measure the efficacy of hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors such as statins, we measured the precursor, hydroxy-3-methylgluturate (HMG), and the product, cholesterol. We calculated the ratios of cholesterol and HMG (CHO:HMG) as an indicator for HMG-CoA reductase inhibitor efficacy. Statins have been used to reduce low-density lipoproteins (LDLs), one of the major risk factors for coronary heart disease (31). Individuals on statin therapy (n = 52) and on ezetimibe therapy (n = 9) had a lower CHO:HMG ratio (t test, mean 0.85, P = 0.001 and mean 0.80, P = 0.02, respectively)
disease and cardiomyopathy of adult genetic disorders, providing early-stage physiological measurements such as atherosclerotic that advanced imaging plays an important role in the detection of ages of associations higher than 75%. Moreover, we demonstrate arrhythmia, and diabetes and endocrine diseases had percent-
ranged from 37.5 to 100% depending on disease categories.

cohort. The percentage of genotype and phenotype associations make a diagnosis of genetic disorders in our adult self-referred type and phenotype associations supporting clinicians medical and family history, 137 (11.5%) participants had geno-
ing metabolomics and advanced imaging in addition to past disease-causing genes and variants, theoretical burden based on possible ambiguous results because of the lack of genotype and phenotype associations. Genetic testing that identifies a P/LP variant is not equivalent to diagnosing the patient with the associated genetic disorder. Physicians need to properly integrate that predictive power of genetic findings with the patient medical data for differential diagnosis (34). In our cohort, 69 (5.8%) individuals had P/LP variants but did not have associated family history, medical history, or phenotypes detected in tests. There are 3 possible explanations for the lack of disease symptoms in this population. First, the probability of nonpathogenicity (1 to 10% for a likely pathogenic variant) (34, 35), reduced penetrance, variable expressivity, or late onset of disease presentation may be reasons for the lack of genotype and phenotype associations. For example, the c.470T>C and c.1283C>T variants in the CHEK2 gene have been reported to cause an increased risk of cancer from large meta-analyses but to be associated with a much lower risk (reduced penetrance) than other CHEK2 pathogenic vari-
ants (36). Second, family and medical history intake may not be comprehensive enough to be used to establish the clinical diagnosis of a genetic disease. Additional personal and family history for potential phenotype associations with the variant may need to be obtained for further assessments. None of the 6 in-
dividuals with the HOXB13 c.251G>A variant reported family or medical history of prostate cancer. Cardiovascular diseases, hyper-tension, and diabetes were the most frequent family history documentation in the health system (37). From our data, the fraction of genotype and phenotype associations from family history ranged from 24 to 57% in dyslipidemia, cardiomyopathy and arrhythmia, diabetes, and endocrine diseases, and our deep phenotype tests detected clinical and preclinical presentations,
Fig. 4. Metabolite levels in heterozygous carriers of rare coding variants in recessive inborn errors of metabolism genes. Shown is the normalized Z score of the metabolite levels of each individual who carries a coding variant in each gene (y-axis), with the amino acid position shown on the x-axis. All variants have minor allele frequency (MAF) <0.5%, and all have Combined Annotation Dependent Depletion (CADD) scores >15. Red lines indicate cutoffs for normalized Z scores above 0, indicating increased levels of phenylalanine relative to the mean, and 2, indicating outliers of interest. Known causal variants are shown in black according to http://www.biogpu.org/home/home.asp for PAH and ClinVar or HGMD for ETFDH and DMGDH (none of the DMGDH variants shown here were found in ClinVar or HGMD). Gene annotations were generated from pfam.xfam.org. (A) Phenylalanine levels of PAH variant carriers. Green, ACT; red, biopterin-dependent aromatic amino acid hydroxylase. (B) Dimethylglycine levels of DMGDH variant carriers. Green, FAD-dependent oxidoreductase; red, Flavin adenine dinucleotide (FAD)-dependent oxidoreductase central domain; blue, aminomethyltransferase folate-binding domain; yellow, glycine cleavage T-protein C-terminal barrel domain. (C and D) Octanoylcarnitine (C) and decanoylcarnitine (D) levels of ETFDH variant carriers. Green, Th4 family; red, electron transfer flavoprotein-ubiquinone oxidoreductase, 4Fe-4S.

further enhancing the clinical diagnosis of genetic disorders in adults (Fig. 2B). Third, protective (i.e., resilience) alleles may exist in the genomes of these asymptomatic individuals and are yet to be discovered.

With deep phenotyping, we found that heterozygous carriers of autosomal recessive diseases exhibited detectable phenotypic changes. Heterozygous carriers of autosomal recessive conditions are usually healthy with no related symptoms of the disease. Some carriers may have subtle or milder symptoms of the disease. We identified 8 PKDH1 (average age 53.1 y) heterozygotes with hepatic and/or renal cysts from the WB-MRI results. The remaining 5 PKDH1 carriers (average age 35.6 y) had no detectable cysts at the time of evaluation. From the genomics and metabolomics associations, we identified 61 (5.1%) heterozygotes with metabolite manifestations affecting serum metabolite levels. In particular, 10 out of 30 (33%) PAH carriers had elevated phenylalanine detected by metabolomics, similar to observations using the phenylalanine tolerance test (38), further confirming the pathogenicity of the genetic variants. Our data have not been explored extensively to evaluate if a long-term phenotype manifestation in heterozygous carriers of autosomal recessive conditions may cause an impact on health. For example, heterozygous carriers of HFE C282Y have been shown to have mild phenotypes of iron overload, resulting in an increased risk of hepatocellular carcinoma (39). Longitudinal evaluation of these individuals will be required to characterize the clinical significance of undefined findings.

A fraction of the reportable metabolomic results was associated with glucose and insulin dysregulation. These personalized measurements may enable individuals to optimize diet and lifestyle adjustment on the basis of personalized glycemic response. For example, the personalized measurement using metabolomics facilitated the clinical diagnosis of maturity-onset diabetes of the young, where a precision diagnosis can provide appropriate treatment such as low-dose sulfonylurea without the possibility of incorrect treatment using insulin. As for cholesterol homeostasis, we found that measurement of the CHO:HMG ratio tracked with the changes observed with use of cholesterol-lowering medications. Our study also identified several candidate genes associated with the CHO:HMG ratio, suggesting possible mechanistic pathways that regulate cholesterol homeostasis. Further, HMG was associated with cardiovascular events using the TwinsUK Registry. Further studies are required to determine if HMG or the CHO:HMG ratio is a useful biomarker for measuring the efficiency of cholesterol control and management. Other examples are the inhibitors of xanthine oxidase, such as allopurinol, commonly used in the treatment of diseases associated with high levels of uric acid (i.e., urate), including gout and tumor lysis syndrome. Metabolite levels of xanthine, hypoxanthine, orotic acid, orotate acid (i.e., urate), and urate may be useful to monitor drug efficacy (SI Appendix, Fig. S3C).

Medicine is traditionally practiced on individuals who are already symptomatic. Clinicians integrate silo test results for differential diagnosis and devise treatments accordingly. Our approach of integration of omics and advance imaging achieved some notable near-term successes on the identification and clinical assessment of adults with previously undiagnosed genetic disorders. Furthermore, our data also provided plausible genetic causes for abnormal physiological measurements at the individual level of analysis. Our study did not measure health outcomes, benefits, and cost-effectiveness. Repeat evaluation of these individuals is required to characterize the clinical significance of the findings. Overall, our study may be evaluated in the context of methodological considerations for precision medicine initiatives and have the potential to change clinical assessments beyond genome sequencing.

Materials and Methods

Study Population and Process. We enrolled self-referred adults ≥18 y old without acute illness, activity-limiting unexplained illness or symptoms, or known active cancer between September 2015 and March 2018. Participants underwent a verbal review of the institutional review board (IRB)-approved consent. During the consent process, we explained to individuals with known genetic disorders that our test is not suitable for clinical testing and referred them to facilities with clinical genetic tests. The study protocol was
approved by the Western IRB (WIRB), and all subjects gave informed consent. We received permission from the WIRB to collect up to $25,000 for participation with an average of $8,000 for participation in this study. Past medical and family history (the 3-generation pedigree), risk factors, medical symptoms, and medication list were collected prior to or during the visit. Data on medical history of participants in this study included self-reported, medical charts from participants’ physicians, or electronic health records. Noninvasive quantitative whole-body and brain MRI/magnetic resonance angiography, CT coronary artery calcium scoring, electrocardiogram, echocardiogram, and clinical laboratory tests were employed as phenotype tests. Phenotype tests were performed as described in Perkins et al. (19), our pilot study. For clinical laboratory tests, participants performed approximately 120 tests that covered liver, kidney, hematology, endocrine, immunology, lipid, and metal functions. Additional materials and methods have been placed in SI Appendix.

Data Availability. Detailed information of the reporting findings and medically significant variants is provided in SI Appendix.

ACKNOWLEDGMENTS. We acknowledge the individuals who participated in this precision medicine study, without whom the findings would not be possible. Julie Ellison provided medical editing assistance. We acknowledge Human Longevity and Health Nucleus staff past and present: Alexander Graff, Arny Reed, Ana Sanchez, Athena Hutchinson, Carina Sarabia, Catherine Vrona, Cheryl Buffaloing, Cheryl Greenberg, Christina Bonas, Daniel Jones, Danielle Dorris, Desirée Cheney, Diana Cardin Escobedo, Dmitry Tkach, Eric Dec, Genelle Olsen, Greg Olson, Heidi Millard, Helen Messier, Hyun-Kyung Chung, Jason Deckman, Javier Velazquez-Muriel, Jian Wu, Kathy Levine, Keisha Robinson, Krista-Lynn Baker, Kristi Erickson-Miller, Laura Edwards, Melissa Schweitzer, Nicole Boramand, Nolan Tengonciang, Patrick Jamieson, Samantha Punsalan, Tom Folan, Víctor Lavrenko, Wayne Delport, and William Herrera.