Supporting Information

Gonzalez-Garay et al. 10.1073/pnas.1315934110

SI Materials and Methods

Cohort Description. The cohort consists of members and spouses in the Houston Chapter of the Young Presidents Organization (YPO). Criteria for membership into the YPO includes corporate and community leadership (1). This cohort is well educated and of higher socioeconomic status. All 450 YPO members were invited to attend an 8-h educational program incorporating technology, human genetics, anticipated outcomes, ethical considerations, discussion groups, and technology demonstrations and printed materials. Of the 150 attendees, 81 volunteered to participate in this study: 46 men and 35 women, with an average age of 54 y. All 81 elected under the terms of the University of Texas Health Science Center at Houston's institutional review board to receive "need to know" genomic disease risk results. Each volunteer provided a detailed medical and drug use history reviewed by our physician-researcher (C.T.C.). A three-generation medical pedigree was acquired on each volunteer. One volunteer could provide no family history.

Whole exome sequencing (WES) Sequencing. Genomic DNA was extracted using a DNA kit (Promega wizard genomic DNA purification kit) following Promega's instructions (2). The cohort was sequenced twice: the first whole exome sequencing experiment (2011) was performed using Illumina's HiSeq and the Genome Analyzer IIx system (3) after enrichment with Nimblegen V2 kit (44 Mb) (4) (outsourced to the national center for genome resources). Our second WES experiment (2013) was performed using Illumina's newest machines HiSEq. 2500 (3) after enrichment with Agilent SureSelect target enrichment V5+UTRs (targeting coding regions plus UTRs) (5) (outsourced to Axeq Technologies). Genome sequencing of a small subset (24 subjects) for validation purposes was carried out by Complete Genomics Inc. (CGI) (6).

Sequencing Analysis. Our analysis pipeline consists of Novoalign (7), Samtools (8), Picard (9), and The Genome Analysis Toolkit (GATK) (10), followed by variant annotation (11–14) using multiple databases from the University of California Santa Cruz (UCSC) Genome bioinformatics site (15). Fig. 1 illustrates our pipeline. Fig. 2 describes our pipeline to detect known pathogenic variations. We detected known variants associated with human diseases using the Human Genome Mutation Database (HGMD) database from Biobase (16, 17) and genes known to be associated with human disorders from Online Mendelian Inheritance in Man (OMIM) (18, 19) and GeneTests (20). Functional effects of each nonsynonymous coding variant were evaluated using three different functional prediction algorithms [Polyphen 2.0 (21), Sift (22–27), and MutationTaster (28)] using the Database of Human Non-synonymous SNVs and their functional predictions and annotations (dbNSFP) (29). Filtration of common polymorphisms was accomplished using frequencies from the National Heart, Lung, and Blood Institute (NHLBI) exome sequencing project (ESP) (30), 1,000 Genomes (31, 32), and internally by removing any variant that appeared more than three times in our cohort. In addition, a group of candidate genes was obtained from OMIM (18, 19) for each volunteer after a careful analysis of the family and personal health history of each volunteer. Variations in those OMIM (18, 19) candidate genes were identified and submitted to the same frequency and functional effects filter as described before.

Variant Validation. Every variant identified in our pipeline was evaluated for quality control, and the variant's read alignments in the BAM file [Binary version of a SAM (Sequencing Alignment Map) file] file were visualized using Integrative Genomics Viewer (IGV) (33). The purpose of this step was to try to remove the remaining false positives.

Each genetic variant was validated using the following steps: (*i*) retrieve reads over variant sites for each individual; (*ii*) make SamTools (8) genotype calls (an alternate calling algorithm); (*iii*) retrieve quality scores for all reads; (*iv*) keep track of the directional depth and require at least two variant reads in the 5′ and 3′ orientation for a variant to be considered true; and (*v*) filter out variants if the SamTools (8) genotype call disagrees with the GATK (10) call or if the quality scores or directional depth values do not exceed minimum values.

Establishing Criteria for Highly Reliable Variant Calling from Exome Sequencing. Our first objective was to define the methods needed to identify a set of "highly reliable" variants from the Illumina sequencing and apply these methods to variant calling on all of our samples. To meet our definition of a highly reliable variant, each variant had to be detected under two independent orthogonal sequencing technologies and been considered as high quality. Because there is not a common definition of what a high-quality variant is, we decided to take advantage of the confidence category scores provided from complete genomics; variants with a score of VQHIGH are consider high quality (masterVarbeta files version 2.0) and develop an equivalent value in our illumina sequencing data. To accomplish our first objective, a dataset of variants was generated from a set of 24 samples that we sequenced using Illumina (3) and an orthogonal sequencing technology (CGI) (6). CGI has their own proprietary workflow from alignment to data annotation (34), Fig. 1 describes our analysis workflow for exome sequencing data. Fig. S2*A* shows the intersection between the nonsynonymous coding variants (NSCVs) detected by CGI (6) and Illumina (3) exome sequencing. We extracted variants from CGI with a score of VQHIGH and that were also detected in the corresponding illumina's vcf file (Fig. S2*B*). This subset of highly reliable variants represents an average of 72% of the variants detected by CGI. By using our dataset, we were able to systematically test for conditions and software setting in our pipeline that generate the majority of the highly reliable variants and reduce the probability of selecting variants not present in our dataset. We reached the conclusions that by using two variant callers tools, GATK UnifiedGenotyper and mplileup/bcftools (samtools), and selecting an overlapping set of variants, we obtained variants of the highest quality. In addition, a postcalling filter enforces that each variant has to have a mapping quality >30, a base quality >20, and a coverage ≥10, with at least a 3:7 ratio of variant to reference (Het) and the presence of the variant in reads from both orientations. By using these postcalling filters, we eliminated the majority of false-positive calls (FP).

Counseling. Genome counseling was conducted by a board-certified internist and a medical geneticist by both individual meetings and two written summaries over a period of 12 mo. The summary reports were prepared and jointly endorsed by a bioinformatician and a physician. Additional counseling was conducted by phone calls and appointments with their physician as requested by the volunteers.

Counseling of Results. Both causative and problematic alleles were reported verbally and in two written reports over an 18-mo period.

The first comprehensive report was updated ~1 y after (*i*) larger control databases downgraded some problematic alleles with more than a 1% frequency; (*ii*) private consultation with disease experts; and (*iii*) validation with original publications and small disease center databases. Several new disease–gene associations were discovered for the reported familial diseases found by pedigree and personal medical histories. Volunteers were informed that these were research results and instructed to consult with their personal physician so that they could have the results validated in a Clinical Laboratory Improvement Amendments (CLIA)-certified laboratory. Volunteers whose family members warranted genetic study were referred to the Baylor College of Medicine genetics program as a medical referral because this function was outside the institutional review board scope and Baylor College of Medicine offered both clinical genetic and CLIA Laboratory expertise. Our study preceded the publication of the incidental findings guidelines in clinical WES and whole genome sequencing (WGS) of the American College of Medical Genetics and Genomics (ACMG) (35). However, we have reviewed their list of 57 genes and 24 actionable conditions, and we found that we included all their genes in our analysis.

Poststudy Survey

We conducted an online survey to assess volunteers' experiences of participating in this project under a Baylor College of Medicine institutional review board. The survey consisted of 82 items and focused on how the volunteers felt about taking part in the research project, as well as their perspectives on genetic information in health care and genomic research in general. Study participants were told the survey was completely voluntary and that they could skip any question they preferred not to answer and could end their participation at any time.

All 81 study volunteers were invited via e-mail to participate in the anonymous online survey within 12 mo after receiving their individual genome reports. Forty-two participants responded to the online survey (response rate, 51.9%; 38 responses were complete). Of those who responded, 59% were men, 41% were women, and 95% had biological children. Ninety-seven percent described their race as white, and 5% chose "other" (participants could choose all that applied); 5% also identified themselves as Hispanic or Latino. All participants had earned a college degree, and 63% had completed at least some graduate work. All participants reported having had a routine medical check-up within the last 2 y, and when asked how they would rate their health, 58% reported excellent, 29% reported very good, 11% reported good, and 3% reported fair.

Poststudy survey results. This study had as its objective to deliver helpful medical genetic information. The mandatory education program informed volunteers that unexpected risks were to be expected. Our institutional review board required volunteers to have the options of declining this information. None chose that option.

The results of the anonymous online survey showed that, overall, participants were motivated to take part in the project to receive their genetic results and learn about their personal risk of disease. Seventy-nine percent of respondents reported that the opportunity to receive their personal genetic results was the most important factor in their decision to take part in the project, whereas another 10% cited a personal interest in genetics in general. When asked to choose which factor was most important in their decision to receive their personal genetic results, most respondents (52%) reported that their interest in finding out their personal risk for diseases was the most important factor; other important factors included the desire to get information about risk of health conditions for their children (17%), the desire to learn more about the medical conditions in their family (10%), and curiosity about their genetic makeup (10%).

Ninety-seven percent of respondents agreed or strongly agreed that they were glad that they decided to participate in this study and receive their personal results, leaving only 3% undecided. Most respondents (72%) spoke with their primary care provider about their results, and 50% reported that they spoke with other medical professionals, including cardiologists, oncologists, and obstetricians/gynecologists, among others; 22% reported that they had their twice-confirmed research results confirmed in a CLIA-certified laboratory.

Twenty-five percent of respondents reported that the test results motivated them to make changes to their health care (i.e., undergoing tests, seeing a specialist, taking vitamins or herbal supplements), exercise, medications, or insurance (Table S11).

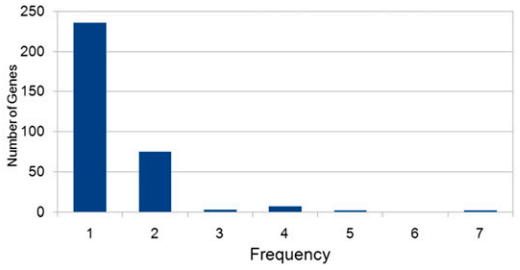
Respondents generally felt that researchers should offer personalized results to research participants: 54% felt that researchers are obligated to offer results, 22% felt that researchers are obligated to offer results only if the researcher is a physician, and the remaining 24% did not think researchers were obligated to offer results. Respondents were pleased with the methods by which they were given their results in this study, with 95% agreeing or strongly agreeing that they were glad the researchers sent them a personalized results report, and 100% agreeing or strongly agreeing that they found the in-person consultation about their results very helpful. When asked, 94% said they would also want an electronic record of their entire genome if it were available.

When asked about genetic testing in health care, 83% reported that they felt that genetic testing should be a regular part of health care and 97% agreed or strongly agreed that they felt comfortable using these results to make decisions about their health. Nevertheless, respondents were evenly split when asked if they thought these results should be part of their medical record.

In summary, our poststudy surveys indicated that volunteers were motivated to gain personal and family health knowledge, satisfied with the translation of the genetic information, and had a divided opinion about incorporating their genetic information into their medical records.

1. Anonymous Membership criteria YPO. Available at http://www.ypo.org/join-ypo/. Accessed September 19, 2013.
2. Anonymous Wizard. Available at http://www.promega.com/resources/protocols/technical-manuals/0/wizard-genomic-dna-purification-kit-protocol/. Accessed September 19, 2013.
3. Anonymous Illumina. Available at http://www.illumina.com. Accessed September 19, 2013.
4. Anonymous NimbleGen Roche. Available at http://www.nimblegen.com/products/seqcap/ez/v2/index.html. Accessed September 19, 2013.
5. Agilent Technologies Agilent SureSelect array. Available at http://www.genomics.agilent.com/en/Exome-Sequencing/SureSelect-Human-All-Exon-Kits/?cid=cat240002&tabId=AG-PR-1204. Accessed September 19, 2013.
6. Anonymous Complete Genomics Inc. Available at http://www.completegenomics.com. Accessed September 19, 2013.
7. Novocraft.com (2012) Available at http://www.novocraft.com. Accessed September 19, 2013.
8. SAMtools. Available at http://samtools.sourceforge.net/. Accessed September 19, 2013.
9. Picard. Available at http://picard.sourceforge.net/. Accessed September 19, 2013.
10. McKenna A, et al. (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20(9):1297–1303.
11. Cingolani P snpEff: SNP effect predictor. Available at http://snpeff.sourceforge.net/. Accessed September 19, 2013.
12. Cingolani P, et al. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* 6(2):80–92.
13. San Lucas FA, Wang G, Scheet P, Peng B (2012) Integrated annotation and analysis of genetic variants from next-generation sequencing studies with variant tools. *Bioinformatics* 28(3):421–422.
14. Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38(16):e164.
15. Kuhn RM, Haussler D, Kent WJ (2013) The UCSC genome browser and associated tools. *Brief Bioinform* 14(2):144–161.
16. Stenson PD, et al. (2012) The Human Gene Mutation Database (HGMD) and its exploitation in the fields of personalized genomics and molecular evolution. *Curr Protocol Bioinform.*
17. Stenson PD, et al. (2009) The Human Gene Mutation Database: 2008 update. *Genome Med* 1(1):13.

18. Anonymous NCBI OMIM Online Mendelian Inheritance in Man. Available at http://www.ncbi.nlm.nih.gov/omim. Accessed September 19, 2013.
19. Anonymous Online Mendelian Inheritance in Man OMIM. Available at http://omim.org. Accessed September 19, 2013.
20. Anonymous Genetic Testing Registry (GeneTests). Available at http://www.genetests.org. Accessed September 19, 2013.
21. Adzhubei IA, et al. (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7(4):248–249.
22. Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 4(7):1073–1081.
23. Sim NL, et al. (2012) SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res* 40(Web Server issue):W452–W457.
24. Hu J, Ng PC (2012) Predicting the effects of frameshifting indels. *Genome Biol* 13(2):R9.
25. Ng PC, Henikoff S (2001) Predicting deleterious amino acid substitutions. *Genome Res* 11(5):863–874.
26. Ng PC, Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* 31(13):3812–3814.
27. Ng PC, Henikoff S (2006) Predicting the effects of amino acid substitutions on protein function. *Annu Rev Genomics Hum Genet* 7:61–80.
28. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D (2010) MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods* 7(8):575–576.
29. Liu X, Jian X, Boerwinkle E (2011) dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum Mutat* 32(8):894–899.
30. Anonymous NHLBI Exome Sequencing Project (ESP) exome variant server. Available at http://evs.gs.washington.edu/EVS/. Accessed September 19, 2013.
31. Clarke L, Zheng-Bradley X, et al. (2012) The 1000 Genomes Project: Data management and community access. *Nat Methods* 9(5):459–462.
32. Abecasis GR, et al.; 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. *Nature* 467(7319):1061–1073.
33. Robinson JT, et al. (2011) Integrative genomics viewer. *Nat Biotechnol* 29(1):24–26.
34. Complete genomics (data file format standard pipeline version 2.0). Available at http://www.completegenomics.com/customer-support/documentation/100357139.html. Accessed September 19, 2013.
35. Green RC, Berg JS, et al. (2013) ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* 15(7):565–574.



7	LRRK2	BRCA2									
5	CFTR	CDH23									
4	GIGYF2	LDLR	LRP2	MEFV	MYOM1	PAH	SCN5A				
3	APC	EIF4G1	TTN								
2	ABCA1	ABCA3	ABCA4	ABCB4	ABCC8	ACADM	ACAN	ACTN2	ACY1	APOB	...
1	ABCG5	ACAD9	ACAT1	ACE	ACVRL1	AGL	AGT	AKAP9	ALMS1	ALOXE3	...

Fig. S1. Grouping genes by occurrence. Frequency of genes with nonsynonymous coding mutations in our cohort. This graphic provides a summary of the number of times alleles were observed for an individual gene. In each of these cases, the allele was either part of HGMD or OMIM, rare, and carried a high polyphen2 score. An example of a gene with frequent risk alleles include *Titin*, the largest genes in our genome and recently reported to be causative of dilated cardiomyopathy. A second example of a smaller gene with a large number of variations is *CFTR*, where the disease database is deep, and it is known to be one of the most common autosomal recessive diseases in whites. This graphic supports that we did not select polymorphic genes but unique mutations in each volunteer.

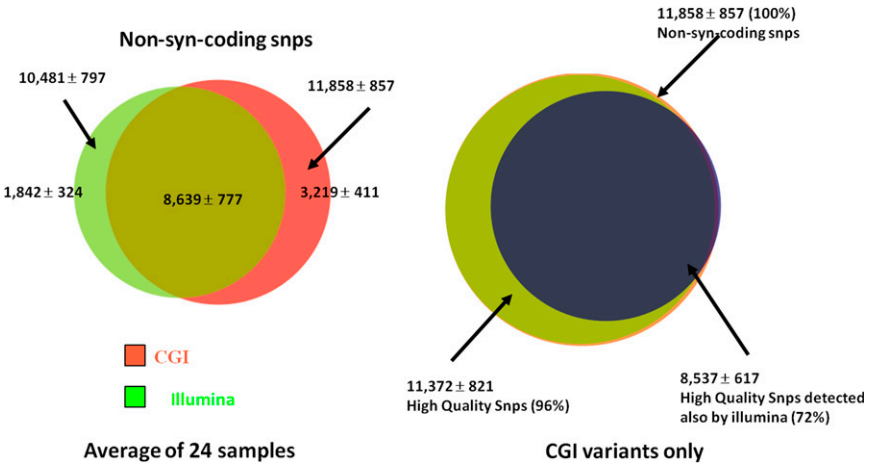


Fig. S2. Variants detected using Complete Genomics Inc. (CGI) and Illumina. (*Left*) Comparison of nonsynonymous coding SNPs (NSCS) obtained from Complete Genomics (red) and Illumina (green). Twenty-four human samples were sequenced using both technologies, and NSCS were compared in each sample. The average results were calculated and graphed as a venn diagram. The intersection represents the set of NSCS detected by both technologies. On average, 73% of the NSCS detected by CGI were also detected by Illumina, while 82% of the NSCS detected by Illumina were also detected by CGI. (*Right*) Using the same samples we calculated that 96% of all the CGI NSCS are considered "High Quality" according to the CGI proprietary quality matrix. An average of 72% of all the NSCS detected by CGI was also detected by Illumina (blue). Since two orthogonal sequence technologies detected the same set of NSCS, this group of variants most likely represents a set of real variants which we refer to as "Highly reliable NSCS." The set of "Highly reliable NSCS" were used to establish quality criteria in our Illumina's variant detection pipeline.

Table S1. Disease associations with alleles

Case	Disease	Risk gene	Allele	HGMD	OMIM gene ID
3937	Hypercholesterolaemia	*LDLR*	p.P526H	CM100938	606945
3890	Hypercholesterolaemia	*LDLR*	p.T726I	CM920469	606945
3910	Hypercholesterolaemia	*LDLR*	p.T726I	CM920469	606945
3900	Hypercholesterolaemia	*LDLR*	p.V827I	CM920471	606945
3915	Hypercholesterolaemia	*LDLR*	p.V827I	CM920471	606945
3923	Obesity	*MC4R*	p.I251L	CM030483	155541
3923	Diabetes mellitus, type II	*MAPK8IP1*	p.D386E	NA	604641
3973	Obesity	*MC4R*	p.C326R	CM070992	155541
3937	Diabetes mellitus type 2 (MODY)	*FN3K*	p.H146R	NA	608425
3937	Diabetes mellitus type 2 (MODY)	*PASK*	p.P1256L	NA	607505
3923	Macular degeneration, age related	*ABCA4*	p.G863A	CM970003	601691
3898	Brittle cornea syndrome type 1 (BCS1) keratoconus	*ZNF469*	p.D2902Y	NA	612078
3889	Male infertility	*USP26*	p.T123_Q124insT	NA	300309
3942	Melanoma	*BAG4*	p.W103X	NA	603884
3959	Melanoma	*GRIN2A*	p.N1076K	NA	138253
3896	Breast or ovarian cancer	*BRCA2*	p.I505T	CM010167	600185
3959	Breast or ovarian cancer	*BRCA2*	p.S384F	CM065036	600185
3897	Breast or ovarian cancer	*BRCA2*	p.T2515I	CM994287	600185
3950	Follicular thyroid cancer (age 41)	*TPR*	p.R105C	NA	189940
3960	Prostate cancer	*LRP2*	P.N479H	NA	600073
3960	Prostate cancer	*LRP2*	P.G4417D	NA	600073
3934	Nonsyndromic deafness	*MYH14*	p.M161I	NA	608568
3934	Nonsyndromic deafness	*SLC17A8*	p.R75C	NA	607557

NA, not available.

Table S2. Familial diseases and associations

				Association	
Case	Disorder	Gene	Volunteer relatedness	Volunteer	Affected relative
3949	Praeder Willie	*MAGEL2*	2°	—	+
3947	Paraganglioma	*SDHB*	1°	—	+
3930	Ankylosing spondylitis	*HLA-B27*	1°	—	+
3930	Tourettes	*TBD*	1°(3)	IP	IP
3928	Parkinson	*LRRK2*	1°	—	+

—, negative; IP, research in progress.

Table S3. Recessive disorders

Cases	Disease	Risk gene	Allele	HGMD	OMIM
3958	Niemann-Pick type C2 disease	*NPC2*	p.N111K	CM081368	601015
3896, 3900, 3915, 3895	Antitrypsin α1 deficiency	*SERPINA1*	p.R247C, p.E366K (3)	CM910298, CM830003	107400
3894	Glycogen storage disease 0	*GYS2*	p.Q183X	CM023388	138571
3889	Glycogen storage disease 1a	*G6PC*	p.R83C	CM930261	613742
3901	Glycogen storage disease 3	*AGL*	p.R477H	CM104343	610860
3945	Glycogen storage disease 4	*GBE1*	p.Y329S	CM960705	607839
3898	Glycogen storage disease 6	*PYGL*	p.D634H	CM078418	613741
3941, 3952	Glycogen storage disease 9B	*PHKB*	p.Q650K	CM031327	172490
3915, 3919, 3943, 3954	Fanconi anemia	*FANCA*	p.T126R, p.S858R (3)	CM043494, CM992317	607139
3936, 3934	Familial Mediterranean fever	*MEFV*	p.E148Q, p.P369S, p.R408Q	CM981240, CM990837, CM990838	608107
395, 439, 243, 953	Cystic fibrosis	*CFTR*	p.D1152H, p.S1235R,	CM950256, CM930133	602421
3933	Sandhoff disease	*HEXB*	p.A543T	CM970723	606873
3940	Fuchs endothelial dystrophy	*ZEB1*	p.Q824P	CM100242	189909
3908	Factor V deficiency	*F5*	p.P1816S	CM095204	612309
3952	Hepatic lipase deficiency	*LIPC*	p.T405M	CM910258	151670
3962	Krabbe disease	*GALC*	p.T112A	CM960678	606890
3954	Macular corneal dystrophy, type 2	*CHST6*	p.Q331H	CM055930	605294
3891, 3947, 3959, 3924, 3895, 3897	Usher syndrome 1d	*CDH23*	p.A366, p.D1806E, p.R1060W	CM050545, CM105104, CM021537	605516
3900, 3910	Phenylketonuria	*PAH*	p.A300S, p.R53H	CM920555, CM981427	612349
3933, 3946	MCAD (medium-chain acyl-coA dehydrogenase deficiency)	*ACADM*	p.K329E (2)	CM900001	607008
3914	Adrenal hyperplasia	*HSD3B2*	p.R249X	CM950655	613890
3926	17-α-hydroxylase/17,20-lyase deficiency	*CYP17A1*	p.R449C	HM0669	609300

Table S4. X-linked recessive

Case	Disorder	Risk gene	Allele	Sex	HGMD	OMIM
3891	ATRX syndrome	*ATRX*	p.N1860S	Female	CM950125	300032
3930	Fabry disease	*GLA*	p.A143T	Female	CM972773	300644
3901	Mucopolysaccharidosis II	*IDS*	p.D252N	Female	CM960865	300823

Table S5. Breast cancer risk

Case	Disease	Risk gene	Allele	Family history	Sex	Age (y)	HGMD	OMIM gene ID
3959	Breast cancer	*BRCA2*	p.S384F	Affected (44)	Female	44	CM065036	600185
3896	Breast cancer	*BRCA2*	p.I505T	Affected	Female	49	CM010167	600185
3955	Breast cancer	*BRCA2*	p.E1625fs	Negative	Female	42	CD011121	600185
3962	Breast cancer	*PALB2*	p.V1103M	First, second, third degree (2) (49–60s)	Female	51	CM118272	610355
3936	Breast cancer	*BRCA1*	p.Y856H	First degree (sister 40s)	Male	62	CM042673	113705
3936	Breast cancer	*BRCA2*	p.K2729N	First degree (sister 40s)	Male	62	CM021957	600185
3963	Breast cancer	*BRCA2*	p.R2034C	First degree (60s)	Male	48	CM994286	600185
3897	Breast cancer	*BRCA2*	p.T2515I	First degree (80)	Female	51	CM994287	600185
3934	Breast cancer	*RAD51C*	pT287A	First degree (uterine)	Female	50	NA	602774
3939	Breast cancer	*RAD50*	p.R1069X	First degree breast (60s)/second colon (60s)	Male	56	NA	604040
3912	Breast cancer	*RAD51C*	p.A126T	Negative	Male	77	CM1010201	602774
3923	Breast cancer	*RAD51C*	pT287A	Negative	Male	60	CM1010198	602774
3956	Breast cancer	*RAD51C*	pT287A	Negative	Male	59	CM1010198	602774

NA, not available.

Table S6. Colon cancer risk

Case	Disease	Risk gene	Allele	Family history	Sex	Age (y)	HGMD	OMIM gene ID
3896	Colon cancer	*MLH1*	p.K618A	First degree	Female	49	CM973729, CM950808	120436
3891	Colon cancer	*MLH3*	p.E1451K	First degree (70s)	Female	62	CM013011	604395
3897	Colon cancer	*APC*	p.A2690T	First and second degree cancer	Female	51	CM045404	611731
3904	Colon cancer	*MSH2*	p.G315V	Second degree	Male	49	CM995220	609309
3897	Colon cancer	*MSH2*	p.G12D	Negative	Female	51	CM950813	609309
3962	Colon cancer	*APC*	p.S2621C	Negative	Female	51	CM921028	611731
3955	Colon cancer	*APC*	p.R2505Q	Negative	Female	42	NA	611731
3933	Colon cancer	*MUTYH*	p.G382D	Negative	Female	69	CM020287	604933

NA, not available.

Table S7. Other cancer risk

Case	Disease	Risk gene	Allele	Family history	Sex	Age (y)	HGMD	OMIM gene ID
3959	Melanoma	*GRIN2A*	p.N1076K	Affected	Female	44	NA	138253
3942	Melanoma	*BAG4*	p.W103X	Affected	Male	70	NA	603884
3950	Follicular thyroid cancer	*TPR*	p.R105C	Affected	Male	48	NA	189940
3960	Prostate cancer	*LRP2*	p.N479H	Affected	Male	65	NA	600073
3946	Prostate cancer	*LRP2*	p.M4601I	Negative	Female	59	NA	600073
3957	Prostate cancer	*LRP2*	p.N1797S	First degree (father)	Male	44	NA	600073
3957	Prostate cancer	*DLC1*	p.D89N	First degree (father)	Male	44	NA	604258
3932	Prostate cancer	*CHEK2*	p.E64K	Negative	Male	47	CM030414	604373
3935	Prostate cancer	*ELAC2*	p.R781H	Negative	Female	70	CM010221	605367
3902	Prostate cancer	*MSR1*	p.H441R	Negative	Female	46	CM023581	153622
3900	Prostate cancer	*MSR1*	p.R293X	Negative	Male	45	CM023579	153622
3954	Prostate cancer	*RNASEL*	p.E265X	Negative	Male	72	CM020300	180435
3954	Prostate cancer	*RNASEL*	p.G59S	Negative	Male	72	CM031342	180435
3963	Retinoblastoma	*RB1*	p.R656W	Negative	Male	48	CM030511	614041
3896	Pituitary cancer	*ACVRL1*	p.A482V	Negative	Female	46	CM994582	601284
3896	Pituitary cancer	*ACVRL1*	p.A482V	Negative	Female	46	CM994582	601284
3930	Esophageal cancer	*WWOX*	p.G178S	Negative	Female	52	NA	605131
3973	Esophageal cancer	*WWOX*	p.R120W	Negative	Male	71	CM016224	605131
3916	Esophageal cancer	*WWOX*	p.R120W	Negative	Male	70	CM016224	605131
3941	Gastric cancer	*MET*	p.A347T	Negative	Male	46	NA	164860

NA, not available.

Table S8. Cardiomyopathy-affected volunteers

Case	Disease	Risk gene	Allele	Clinical	Age (y)	HGMD	OMIM gene ID
3925	Dilated cardiomyopathy	*MYH6*	p.A1443D	Atrial fibrillation	65	CM107536	160710
3926	Cardiomyopathy arrhythmogenic right ventricular	*DSG2*	p.V158G	Arrhythmia	65	CM070921	125671
3935	Dilated cardiomyopathy	*MYH6*	p.R1398Q	Cardiac dysrhythmia	70	NA	160710
3935	Cardiomyopathy, dilated, 1EE	*MYH6*	p.R1398Q	Cardiac dysrhythmia	70	NA	160710
3935	Arrhythmogenic right ventricular cardiomyopathy	*TTN*	p.P3751R	Cardiac dysrhythmia	70	NA	188840
3955	Dilated cardiomyopathy	*ACTN2*	p.Q349L	1° pacemaker	53	NA	102573
3955	Familial hypertrophic cardiomyopathy 12	*CSRP3*	p.R100H	1° pacemaker	53	CM091458	600824
3916	Dilated cardiomyopathy type 1A	*LAMA2*	p.T821M	Stent placement	71	NA	156225
3887	Cardiomyopathy, hypertrophic	*MYBPC3*	p.R326Q	Stent placement (3)	73	CM020155	600958
3887	Cardiomyopathy familial hypertrophic (CMH)	*MYLK2*	p.V402F	Stent placement (3)	73	NA	606566
3953	Brugada syndrome (arrhythmia)	*KCNE3*	p.M65T	Two bypass, stent, and familial history of CAD	71	NA	604433
3953	Arrhythmogenic right ventricular cardiomyopathy	*TTN*	p.P5237T	Two bypass, stent, and familial history of CAD	71	NA	188840
3937	Hypercholesterolaemia	*LDLR*	p.P526H	Three generations of early MI, elevated LDL, cholesterol, triglycerides, and treated with statins	53	CM100938	606945
3890	Hypercholesterolaemia	*LDLR*	p.T726I	1° early MI	57	CM920469	606945
3910	Hypercholesterolaemia	*LDLR*	p.T726I	1° aortic occlusion, elevated cholesterol	51	CM920469	606945
3900	Hypercholesterolaemia	*LDLR*	p.V827I	1° early MI	45	CM920471	606945
3915	Hypercholesterolaemia	*LDLR*	p.V827I	Three generations of elevated cholesterol, treated with statins	70	CM920471	606945

CAD, coronary artery disease; MI, myocardial infarction; NA, not available.

Table S9. Cardiomyopathy unaffected but family history

Case	Disease	Risk gene	Allele	Clinical	Age (y)	HGMD	OMIM gene ID
3943	Arrhythmogenic right ventricular cardiomyopathy	*TTN*	p.G1345D	Familial history of arrhythmia	44	NA	188840
3896	Dilated cardiomyopathy	*SYNE1*	p.L3057V	Familial history	45	NA	608441
3896	Arrhythmogenic right ventricular dysplasia/cardiomyopathy	*JUP*	p.V648I	Familial history	45	NA	173325
3944	Hypertrophic cardiomyopathy	*OBSCN*	p.K1671N	Father	45	NA	608616
3931	Dilated cardiomyopathy	*MYH6*	p.R1398Q	Familial history	46	NA	160710
3907	Cardiomyopathy, hypertrophic	*ACTN2*	p.T495M	Father	47	CM101366	102573
3950	Cardiomyopathy	*MYOM1*	p.G1162S	Familial history	48	NA	603508
3919	Romano-Ward syndrome (arrhythmia)	*SCN5A*	p.S1769N	Familial history	51	CM002391	600163
3889	Romano-Ward syndrome (arrhythmia)	*SCN5A*	p.S1769N	Mother	51	CM002391	600163
3917	Cardiomyopathy	*MYOM1*	p.R1573Q	Familial history + father	51	NA	603508
3960	Dilated cardiomyopathy	*NEBL*	p.K60N	Son CAD	66	CM106905	605491
3976	Cardiomyopathy	*MYOM1*	p.E704K	Older brother	72	NA	603508
3976	Early onset myopathy	*MYH2*	p.V970I	Older brother	72	CM051560	160740

Table S10. Neurodegenerative risk

Case	Disease	Risk gene	Allele	Family history	Age (y)	HGMD	OMIM
3908	Alzheimer's disease	*APOE*	p.C130R	Negative	44	CM900020	107741
3916	Alzheimer's disease	*APOE*	p.L46P	Parkinson 1° (72)	71	CM990167	107741
3954	Alzheimer's disease	*APP*	p.R469H	Negative	72	NA	104760
3942	Frontotemporal dementia	*MAPT*	p.S427F	Negative	71	NA	157140
3954	Frontotemporal dementia	*MAPT*	p.V224G	Negative	72	NA	157140
3895	Parkinson disease	*EIF4G1*	p.G686C	Negative	49	CM117028	600495
3916	Parkinson disease	*EIF4G1*	p.R1205H	Parkinson 1° (78)	64	CM117009	600495
3951	Parkinson disease	*EIF4G1*	p.S1596T	Negative	64	NA	600495
3931	Parkinson disease 11	*GIGYF2*	p.P1222fs	Negative	44	NA	612003
3946	Parkinson disease 11	*GIGYF2*	p.H1171R	Negative	59	NA	612003
3957	Parkinson disease 11	*GIGYF2*	p.M48I	Negative	44	NA	612003
3930	Parkinson disease 11	*GIGYF2*	p.S1035C	Negative	52	NA	612003
3933	Parkinson disease 11	*GIGYF2*	p.S1035C	Negative	68	NA	612003
3928	Parkinson disease	*LRRK2*	p.A419V	Tremor 1° Parkinson 2°	68	CM125746	609007
3903	Parkinson disease	*LRRK2*	p.D972G	Negative	54	NA	609007
3919	Parkinson disease	*LRRK2*	p.D972G	Negative	51	NA	609007
3889	Parkinson disease	*LRRK2*	p.G2019S	Negative	51	CM050659	609007
3951	Parkinson disease	*LRRK2*	p.L119P	Negative	50	NA	609007
3918	Parkinson disease	*LRRK2*	p.L286V	Negative	64	NA	609007
3907	Parkinson disease	*LRRK2*	p.P1542S	Alzheimer's 2°	47	NA	609007
3935	Parkinson disease	*LRRK2*	p.P1542S	Negative	70	NA	609007
3893	Parkinson disease	*LRRK2*	p.R1514Q	Negative	45	CM057190	609007
3943	Parkinson disease	*LRRK2*	p.R1514Q	Negative	50	CM057190	609007
3949	Parkinsonism, juvenile, autosomal recessive	*PARK2*	p.R275W	2° three siblings	52	CM991007	602544
3924	Parkinsonism, juvenile, autosomal recessive	*PARK2*	p.R334C	Negative	54	CM003865	602544
3927	Parkinson	*PM20D1*	p.A332V	Negative	73	NA	613164
3886	Parkinson	*PM20D1*	p.P281Q	Negative	62	NA	613164

Table S11. Percentage of survey respondents reporting having made behavioral changes specifically motivated by their test results

Type of behavior change	Yes	No
Changes to diet	4 (10%)	36 (90%)
Changes to health care (such as undergoing tests or seeing a specialist)	4 (10%)	36 (90%)
Changes to use of vitamins/herbal supplements	4 (10%)	36 (90%)
Changes to exercise	3 (8%)	37 (92%)
Changes to medications	1 (2%)	39 (98%)
Changes to insurance coverage	1 (2%)	39 (98%)
Number of respondents making at least one of the above behavior changes	10 (25%)	