PLOS | ONE

RESEARCH ARTICLE

Pathogenic Mutations in Cancer-Predisposing Genes: A Survey of 300 Patients with Whole-Genome Sequencing and Lifetime Electronic Health Records

Karen Y. He[1], Yiqing Zhao[2], Elizabeth W. McPherson[2], Quan Li[3], Fan Xia[4], Chunhua Weng[5], Kai Wang[3,5,6]*, Max M. He[2,7,8]*

1 Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio, United States of America, **2** Center for Human Genetics, Marshfield Clinic Research Foundation, Marshfield, Wisconsin, United States of America, **3** Zilkha Neurogenetic Institute, University of Southern California, Los Angeles, California, United States of America, **4** Deparment of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, United States of America, **5** Department of Biomedical Informatics, Columbia University Medical Center, New York, New York, United States of America, **6** Institute for Genomic Medicine, Columbia University Medical Center, New York, New York, United States of America, **7** Biomedical Informatics Research Center, Marshfield Clinic Research Foundation, Marshfield, Wisconsin, United States of America, **8** Computation and Informatics in Biology and Medicine, University of Wisconsin-Madison, Madison, Wisconsin, United States of America

* kw2701@cumc.columbia.edu (KW); he.max@mcrf.mfldclin.edu (MMH)

Abstract

Background

It is unclear whether and how whole-genome sequencing (WGS) data can be used to implement genomic medicine. Our objective is to retrospectively evaluate whether WGS can facilitate improving prevention and care for patients with susceptibility to cancer syndromes.

Methods and Findings

We analyzed genetic mutations in 60 autosomal dominant cancer-predisposition genes in 300 deceased patients with WGS data and nearly complete long-term (over 30 years) medical records. To infer biological insights from massive amounts of WGS data and comprehensive clinical data in a short period of time, we developed an in-house analysis pipeline within the SeqHBase software framework to quickly identify pathogenic or likely pathogenic variants. The clinical data of the patients who carried pathogenic and/or likely pathogenic variants were further reviewed to assess their clinical conditions using their lifetime EHRs. Among the 300 participants, 5 (1.7%) carried pathogenic or likely pathogenic variants in 5 cancer-predisposing genes: one in *APC*, *BRCA1*, *BRCA2*, *NF1*, and *TP53* each. When assessing the clinical data, each of the 5 patients had one or more different types of cancers, fully consistent with their genetic profiles. Among these 5 patients, 2 died due to cancer while the others had multiple disorders later in their lifetimes; however, they may have benefited from early diagnosis and treatment for healthier lives, had the patients had genetic testing in their earlier lifetimes.

data: Murray H. Brilliant, Ph.D., Director of
Center for Human Genetics, Marshfield Clinic
Research Foundation, 1000 North Oak Ave.,
Marshfield, WI 54449, Phone: 715-221-6469,
Email: brilliant.murray@mcrf.mfldclin.edu.

Abbreviations: ACMG, American College of
Medical Genetics and Genomics; EHRs, electronic
health records; NGS, next-generation sequencing;
WES, whole-exome sequencing; WGS, whole-
genome sequencing.

Conclusions

We demonstrated a case study where the discovery of pathogenic or likely pathogenic germline mutations from population-wide WGS correlates with clinical outcome. The use of WGS may have clinical impacts to improve healthcare delivery.

Introduction

Next-generation sequencing (NGS) technologies are increasingly used in biomedical research and clinical practice to identify disease-associated genetic variants for advancing precision medicine [1]. Precision medicine allows researchers and physicians to predict more accurately which treatment and prevention strategies for a particular disease will work in which groups of people based on their genetic differences [2]. More than 4,000 Mendelian disorders have been studied at the genetic level [3]. Assessment of genetic pathogenicity leveraging whole-genome sequencing (WGS), whole-exome sequencing (WES), or other sequencing data and electronic health records (EHRs) has recently become feasible as EHRs have been implemented widely in healthcare systems [4, 5].

Even though tens of millions of genetic variants are uncovered in the human genomes, we do not have a clear understanding of the majority of their roles in health and disease [6]. The American College of Medical Genetics and Genomics (ACMG) has recommended identification and return of incidental findings (IFs) in a set of 56 actionable genes [7, 8]. A study on the NHLBI Exome Sequencing Project (ESP) cohorts has reported actionable exomic IFs from 112 genes in 6,503 participants [9]. The spectrum of pathogenic genetic variations across a diverse set of genes spanning dominant and recessively inherited disorders in the Exome Aggregation Consortium (ExAC) population has been assessed [10]. A recent study that focuses on two genes suggests that up to 3% of individuals may be at risk for heart arrhythmias [5]. Germline mutations in 565 cancer-predisposition genes with an emphasis on the analysis of 60 autosomal dominant cancer-predisposing genes have been studied by Zhang et al. [11]. Other studies suggest that 1%-3% of the population may carry clinically actionable variants linked to Mendelian diseases [12–14]. Although these studies have provided significant insights into the human genome and have been able to identify individuals who carry clinically actionable genetic variants, their clinical impact remains unknown.

Our aims were to assess how often cancer gene screening identifies actionable cancer risk genes and to retrospectively evaluate whether the combination of WGS and EHR can facilitate improving prevention and care for patients with susceptibility to cancer syndromes. In this study, we firstly classified genetic pathogenicity of germline mutations in 60 autosomal dominant cancer-predisposition genes (S1 Table) among 300 deceased patients at Marshfield Clinic with WGS data using an in-house analysis pipeline called SeqHBase [15] based on the latest ACMG guidelines issued by the ACMG and the Association of for Molecular Pathology (AMP) [16]. Then we assessed clinical conditions for the patients who carried pathogenic or likely pathogenic variants using clinical data derived from their lifetime EHRs, followed by manual review of medical charts on selected patients by a MD clinical geneticist.

Materials and Methods

Sample Selection

In this study, 300 deceased patients were recruited from a Personalized Medicine Research Project (PMRP) [17–19] launched at Marshfield Clinic in 2002. The PMRP is a unique biorepository resource that includes serum, plasma, and DNA from over 20,000 patients with links to their EHRs, including diagnosis, treatment, procedure codes, laboratory values, prescriptions, pharmacy, and physician notes. This cohort represents a very stable population that are primarily non-Hispanic whites with over 70% claiming German ancestry [20]. Each of them has an average of over 30 years of dynamic, continuous, virtually comprehensive, and extractable EHR data as well as diet and activity data linked to participant biospecimens. Since their enrollment, about 2,000 PMRP participants have passed away, and over 1,000 of them have more than 30 years of longitudinal and nearly comprehensive EHR data. 300 of the 1,000 deceased patients were randomly selected, resulting in 161 females and 139 males. In addition, EHR is not being used to identify those patients. It could be biased towards including more patients with late-onset diseases (e.g. cancer) for collecting patients with having about 30-year EHRs. All participants in the PMRP had previously consented for research in written and this project was approved by Marshfield Clinic's Institutional Review Board.

Generation of 300 WGS Data

The blood samples of the 300 participants were sequenced by Complete Genomics (Mountain View, CA) according to manufacturer's guidelines. The sequencing data were aligned to human reference (hg19). A minimum read-depth of 10 was used for variant calling. The VCF files, including single nucleotide variants (SNVs) and small insertions and/or deletions (INDELs), of the 300 genomes were provided by Complete Genomics. After quality control, over 27 million unique SNV and/or INDEL variants were identified across the 300 genomes.

Computational Methods

NGS technology is an essential component supporting genomic medicine, but the volume and complexity of the data pose challenges for its use in biomedical research [21]. Sequencing a single human genome generates about 200 gigabytes of data. Therefore, enormous challenges for analyzing large-scale NGS and clinical data still exist including data storage, processing, scaling, quality control management, and interpretation [22]. It is critical to develop an efficient computational framework and tools to analyze large-scale sequencing and clinical data. To infer biological insights from massive amounts of NGS data and comprehensive clinical data in a short period of time, we developed an in-house analysis pipeline within a software framework called SeqHBase to quickly catalogue, retrieve and query genetic variants, and to help classify genetic pathogenicity based on the latest ACMG guidelines [16]. We used ANNOVAR [23] to annotate the 300 WGS data, then the variation and annotation information were managed and analyzed by the in-house system SeqHBase.

Variant Classification

Variants in the 60 cancer-predisposition genes (S1 Table) were classified in the 300 genomes. In SeqHBase pipeline, there are a number of data quality filters, including minimum read-depth (e.g. reads > = 30X), maximum variant minor allele frequency (MAF; e.g., MAF < = 0.05%) in the 1000 Genomes Project [13], the ESP [24], and the ExAC [25], variant classification by ClinVar [26], and biological functions interested (e.g., splicing, nonsynonymous, stop-gain, stop-loss, and frameshift). We collected genetic variants of reads > = 30X, including (i)
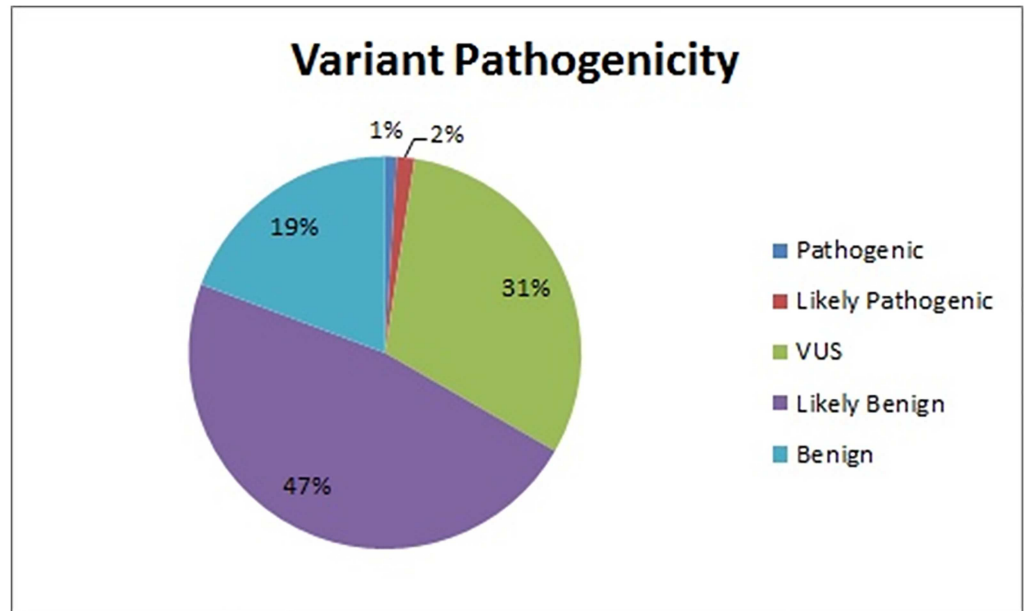
Fig 1. Pie chart of variant categories (pathogenicity) in 60 autosomal dominant cancer-predisposing genes. Provided are the percent of total variants for each category shown in S2 Table.

doi:10.1371/journal.pone.0167847.g001

any variant present in the ClinVar database and annotated in one of the biological functions interested and (ii) new variants absent in the ClinVar database and annotated in the biological functions interested with MAF $<$ = 0.5%, in the 60 cancer-predisposing genes. Variants departing from Hardy-Weinberg equilibrium (exact test P $\leq$ 1E-6) [27] were further filtered. All variants collected can then be classified as "pathogenic," "likely pathogenic," "uncertain significance," "likely benign," and "benign" using a combination of automated assessment and manual review, by following the latest ACMG guidelines [16].

Results

Germline Mutations in the 60 Cancer-Predisposing Genes

In the 300 whole genomes, the data presented in those biological functions interested across the 60 cancer-predisposing genes spanning diverse autosomal dominant cancers encompass 207 classified variants (S2 Table). Of the 207 variants in the 60 genes, 5 variants were classified as "pathogenic" or "likely pathogenic" (Fig 1). The 5 variants are shown in Table 1, and all of them are absent in the 1000 Genomes Project, the ESP, and the ExAC cohorts. To further

Table 1. Pathogenic/likely pathogenic variants in the 60 cancer-predisposing genes identified by WGS of 300 deceased patients*.

Gene	Chr	Position	AA Change	rs #	Function	Homs/Hets
APC	5	112102976	c.311C>G	rs74953290	p.Ser104Ter	0/1
BRCA1	17	41246018	c.1530del	rs80357735	p.Gly511Alafs	0/1
BRCA2	13	32913729	c.5238dupT	rs80359499	p.Ser1746Alafs	0/1
NF1	17	29679366	c.7486C>T	rs866445127	p.Arg2496Ter	0/1
TP53	17	7578389	c.541C>T	rs587782596	p.Arg181Cys	0/1

* Chr denotes chromosome, AA denotes amino acid, and rs # denotes rs number.

doi:10.1371/journal.pone.0167847.t001

Table 2. Clinically Significant Variants Associated with Cancers.

Gene	Function	Observed Phenotypes (age-onset)	Manual interpretation by a certified medical geneticist
APC	p.Ser104Ter	Prostate cancer (early-70s); many colon polyps, developed colorectal cancer (mid-60s); family history of colon cancer (mid-50s)	This variant is located within the exon 3 of the APC gene, and will cause a truncated gene product. This variant has been found in a patient with familial adenomatous polyposis (FAP), PMID 21142386. Truncating variants in APC are the cause of FAP and are high risk factor for colorectal cancer (CRC), which is consistent with this individual's phenotype and family history. In addition, other truncating variants in exon 3 of APC have been found in patients with FAP/CRC (PMID 9664575, 25559809, 21646762 and 20649969). Therefore, this variant is interpreted as pathogenic.
BRCA1	p.Gly511Alafs	Breast and ovarian cancers (early-50s)	This variant is located within the exon 11b of the BRCA1 gene, and will cause a truncated gene product. This variant has been found in a patient from a cohort of breast/ovarian cancers, PMID 12698193. Truncating variants in BRCA1 are high risk factor for breast/ovarian cancers, which is consistent with this individual's phenotype. In addition, other truncating variants in exon 11b of BRCA1 have been found in patients with breast/ovarian cancers (PMID 18489799, 15117986, and 8880569). Therefore, this variant is interpreted as pathogenic.
BRCA2	p.Ser1746Alafs	Prostate cancer (early-70s)	This variant will cause a truncated gene product of BRCA2. This variant has been found in multiple patients with breast/ovarian cancers, PMID 11802209, 23110154 and 24504028. Truncating variants in BRCA2 are high risk factor for breast/ovarian cancers and prostate cancer, which is consistent with this individual's phenotype. Therefore, this variant is interpreted as pathogenic.
NF1	p.Arg2496Ter	Multiple skin (both basal cell and squamous) cancers (mid-70s)	This variant is located within the exon 51 of the NF1 gene, and will cause a truncated gene product. This variant has been found in at least two patients with neurofibromatosis 1, PMID 7981679 and 22965773. Truncating variants in NF1 cause neurofibromatosis 1, and are high risk factor for skin cancers, which is consistent with this individual's phenotype. In addition, other truncating variants in exon 51 of NF1 have been found in patients with neurofibromatosis 1 (PMID 23656349 and 7981679). Therefore, this variant is interpreted as pathogenic.
TP53	p.Arg181Cys	Pancreatic cancer (mid-80s and mid-80s), the patient's child died from melanoma in 30's	This variant will cause a missense change from Arg to Cys at the codon 181 of the TP53 gene product. This variant has been found in the germline of patients with different types of cancers from multiple unrelated families (PMID 7981679, 27157322, 27501770, 23484829 and 22965773). This variant was not found control databases including ExAC and 1000 genome. Defects in TP53 cause Li-Fraumeni Syndrome, and are high risk factor for multiple types of cancers including pancreatic cancer and melanoma, which is consistent with this individual's phenotype and family history. In addition, studies suggest that the Arg181Cys change caused deficiency of TP53 function (PMID 10229196, 12909720 and 21343334). Therefore, this variant is interpreted as pathogenic.

doi:10.1371/journal.pone.0167847.t002

confirm that our automated analysis is reliable, a certified medical geneticist manually reviewed the variant data and provided clinical interpretation on these variants (Table 2).

Clinical Impact

5 of the 300 participants carried presumed "pathogenic" or "likely pathogenic" variants in the 60 autosomal dominant cancer-predisposing genes. The 5 "pathogenic" or "likely pathogenic" variants identified in this study are located in 5 genes including *APC*, *BRCA1*, *BRCA2*, *NF1*, and *TP53*. The *APC* gene encodes a multi-domain protein that plays an essential role in tumor suppression by antagonizing the WNT signaling pathway [28]. Inappropriate activation of this pathway through loss of *APC* function contributes to cancer progression, as in familial

adenomatous polyposis [29]. Mutations in *APC* may result in colorectal cancer [30], prostate cancer [31], and other cancers [32]. The *BRCA1* gene is a protein product and is responsible for DNA repair [33]. It forms several distinct complexes through association with different adaptor proteins, and each complex forms in a mutually exclusive manner [34]. Mutations in *BRCA1* may result in breast and/or ovarian cancer [35] and pancreatic cancer [36]. The *BRCA2* gene is also a protein product responsible for DNA repair [37]. It is a key mediator of homologous recombination [38]. Mutations in *BRCA2* may result in breast and/or ovarian cancer [39], pancreatic cancer [40, 41], and prostate cancer [42] as well. The *NF1* gene encodes neurofibromin, a cytoplasmic protein that is predominantly expressed in neurons, Schwann cells, oligodendrocytes, and leukocytes. Mutations in *NF1* may result in juvenile myelomonocytic leukemia [43], neurofibromatosis [44], Neurofibromatosis-Noonan syndrome [45], and Watson syndrome [46]. Interestingly, a multidisciplinary team at Yale University, led by Yale Cancer Center members, has confirmed that *NF1* is a "major player" in the development of skin cancer [47], which is also observed in this study. The *TP53* gene responds to diverse cellular stress to regulate target genes that induce cell cycle arrest, apoptosis, senescence, DNA repair, and/or changes in metabolism [48]. Mutations in *TP53* may result in adrenal cortical carcinoma [49], breast cancer [50], choroid plexus papilloma [51], colorectal cancer [52], hepatocellular carcinoma [53], Li-Fraumeni syndrome [54], nasopharyngeal carcinoma [55], osteosarcoma [56], pancreatic cancer [57], basal cell carcinoma [58], and glioma susceptibility [59].

The 5 patients, who carried "pathogenic" or "likely pathogenic" variants, were expected to express autosomal dominant cancer-predisposing syndromes based on their genetic profiles. We reviewed the 5 patients' lifetime EHRs and found the following results: (i) One male patient who carried a stop-gain mutation (rs72953290) in *APC* had prostate cancer, many colon polyps, and colorectal cancer. In addition, he had family history of colon cancer. (ii) One female carried a frameshift deletion (rs80357735) in *BRCA1* that is predicted to result in a significantly increased risk for breast and ovarian cancer. In fact, this patient had no known family history of breast cancer and did not receive yearly breast exams or mammograms. She was diagnosed with breast and ovarian cancer in her early-50s and died 6 years later. (iii) One male patient, who carried a frameshift insertion (rs80359499) in *BRCA2*, had prostate cancer in his early-70's and died a few years later. (iv) One male patient, who carried a stop-gain mutation (rs866445127) in *NF1*, had multiple skin (both basal cell and squamous) cancers. (v) One female patient, who carried a missense mutation (rs587782596) in *TP53*, had pancreatic cancer. Additionally her child died from melanoma in his 30's. The summary information is presented in Table 2. Namely, each of the 5 patients had one or more different types of cancers, demonstrating consistency with their genetic profiles.

Even though there are no family members in the 300 deceased patients, it is conceivable that family history recorded in EHRs can help disease prevention. As mentioned in the 2th patient who carried the cancer-predisposing mutation in *BRCA1*, she died from breast and ovarian cancer in her later 50's. This case implied that it could be important to have aggressive screening and prophylactic surgery for patients with *BRCA1* mutations. In addition, it is very likely that more aggressive surveillance or preventative measures could have extended the lives of those patients if genetic testing had been done in their earlier ages. Therefore, combining WGS and EHRs could potentially improve personalized healthcare.

Discussion

Combining the functional characterization of identified genetic variants with comprehensive clinical data available in EHRs has the potential to provide compelling evidence to implicate

novel disease-associated variants in phenotypically well-characterized patients. In this study, we analyzed germline mutations in the 60 autosomal dominant cancer-predisposition genes in 300 deceased patients with WGS data and nearly complete long-term medical records. To infer biological insights from massive amounts of WGS data and comprehensive clinical data in a short period of time, we developed an in-house analysis pipeline within a software framework called SeqHBase to quickly classify genetic pathogenicity based on the latest ACMG guidelines [16]. The pathogenic and/or likely pathogenic variants identified in this study were further reviewed using the carriers' lifetime EHRs. Of the 300 participants, each of the 5 (1.7%) carried a presumed "pathogenic" or "likely pathogenic" variant in one of the 60 cancer-predisposing genes. When assessing extensive clinical data, each of the 5 patients had one or more different cancers, exhibiting fully consistency with their genetic profiles. The results generated in this study demonstrated that genetic mutations in autosomal dominant cancer-predisposing genes could be potentially used in clinical diagnosis, prevention, and personalized treatments, showcasing the power of combining WGS and EHR to accelerate biomedical discoveries. It also showed potential impacts of clinically actionable genetic variants over a lifetime and demonstrated that genomic sequencing could be helpful in precise disease diagnoses and risk prediction. Meanwhile, we have realized that the 300 patients selected in this study should not be regarded as a representative population at Marshfield Clinic. That is, younger people may need to be recruited with WGS study in the future.

We acknowledge that WES or targeted sequencing may provide similar results at a reduce cost. However, WES or targeted sequencing technologies may not capture the whole genome coding regions comprehensively while WGS generates more complete coverage for the whole genome regions [60]. Although sequencing costs have dropped substantially in the past a few years, the cost for data analysis and interpretation remain very high. Further comprehensive studies are needed as true impacts on clinical outcomes may be much more complex. It may not be feasible to use WGS for screening general population in clinical practice now.

In addition, more efforts are needed to distinguish genetic variants that are truly clinically actionable, that is, the variants are useful for guiding clinical decisions regarding interventions to improve health outcomes. As multiple independent evaluations might have classified variant pathogenicity differently [5], more stringent criteria and the latest ACMG guidelines should be compiled prior to reporting pathogenic variants [61].

In summary, clinical research studies of the implementation of genomic data in healthcare can provide valuable lessons on how genomic data should be managed, and patient privacy should be protected, when incorporating genomic data into clinical practice on a larger scale. These lessons can alert healthcare institutions of the scientific and technical challenges of using genomic data in precision medicine. NGS technological advances in clinical genome sequencing and adoption of EHRs will pave the way to create patient-centered precision medicine in clinical practice. The rise of *Big Data* in NGS and clinical data will contribute to better treatment paradigms, leading to improvements in diagnosis and personalized treatments that may ultimately lead to an overall reduction in healthcare cost. This study portrayed a promising method for assessing genetic pathogenicity by using WGS data.

Supporting Information

S1 Table. 60 Autosomal Dominant Cancer-Predisposition Genes.
(XLS)

S2 Table. Genetic Variants of the 300 Patients Across the 60 Cancer-Predisposition Genes.
(XLS)

Acknowledgments

Author Contributions

Conceptualization: MMH.

Data curation: MMH YZ.

Formal analysis: MMH QL.

Funding acquisition: KYH KW MMH.

Investigation: MMH QL KW EWM FX.

Methodology: MMH KW QL FX.

Project administration: MMH.

Resources: MMH.

Software: MMH QL.

Supervision: MMH KW.

Validation: EWM FX.

Visualization: MMH QL.

Writing – original draft: KYH KW MMH.

Writing – review & editing: KYH CW KW MMH.

References

1. Collins FS, Varmus H. A new initiative on precision medicine. The New England journal of medicine. 2015; 372(9):793–5. Epub 2015/01/31. doi: 10.1056/NEJMp1500523 PMID: 25635347

2. Vassy JL, Korf BR, Green RC. How to know when physicians are ready for genomic medicine. Science translational medicine. 2015; 7(287):287fs19. Epub 2015/05/15. doi: 10.1126/scitranslmed.aaa2401 PMID: 25971999

3. Brunham LR, Hayden MR. Hunting human disease genes: lessons from the past, challenges for the future. Human genetics. 2013; 132(6):603–17. Epub 2013/03/19. doi: 10.1007/s00439-013-1286-3 PMID: 23504071

4. Gottesman O, Kuivaniemi H, Tromp G, Faucett WA, Li R, Manolio TA, et al. The Electronic Medical Records and Genomics (eMERGE) Network: past, present, and future. Genetics in medicine: official journal of the American College of Medical Genetics. 2013; 15(10):761–71. Epub 2013/06/08.

5. Van Driest SL, Wells QS, Stallings S, Bush WS, Gordon A, Nickerson DA, et al. Association of Arrhythmia-Related Genetic Variants With Phenotypes Documented in Electronic Medical Records. Jama. 2016; 315(1):47–57. Epub 2016/01/10. doi: 10.1001/jama.2015.17701 PMID: 26746457

6. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, et al. ClinGen—the Clinical Genome Resource. The New England journal of medicine. 2015; 372(23):2235–42. Epub 2015/05/28. doi: 10.1056/NEJMsr1406261 PMID: 26014595

7. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genetics in medicine: official journal of the American College of Medical Genetics. 2013; 15(7):565–74. Epub 2013/06/22.

8. Hampel H, Bennett RL, Buchanan A, Pearlman R, Wiesner GL. A practice guideline from the American College of Medical Genetics and Genomics and the National Society of Genetic Counselors: referral

indications for cancer predisposition assessment. Genetics in medicine: official journal of the American College of Medical Genetics. 2015; 17(1):70–87. Epub 2014/11/14.

9. Amendola LM, Dorschner MO, Robertson PD, Salama JS, Hart R, Shirts BH, et al. Actionable exomic incidental findings in 6503 participants: challenges of variant classification. Genome research. 2015; 25 (3):305–15. Epub 2015/02/01. doi: 10.1101/gr.183483.114 PMID: 25637381

10. Song W, Gardner SA, Hovhannisyan H, Natalizio A, Weymouth KS, Chen W, et al. Exploring the landscape of pathogenic genetic variation in the ExAC population database: insights of relevance to variant classification. Genetics in medicine: official journal of the American College of Medical Genetics. 2015. Epub 2015/12/19.

11. Zhang J, Walsh MF, Wu G, Edmonson MN, Gruber TA, Easton J, et al. Germline Mutations in Predisposition Genes in Pediatric Cancer. The New England journal of medicine. 2015; 373(24):2336–46. Epub 2015/11/19. doi: 10.1056/NEJMoa1508054 PMID: 26580448

12. Yang Y, Muzny DM, Xia F, Niu Z, Person R, Ding Y, et al. Molecular findings among patients referred for clinical whole-exome sequencing. Jama. 2014; 312(18):1870–9. Epub 2014/10/19. doi: 10.1001/jama.2014.14601 PMID: 25326635

13. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. A global reference for human genetic variation. Nature. 2015; 526(7571):68–74. Epub 2015/10/04. doi: 10.1038/nature15393 PMID: 26432245

14. Roche MI, Berg JS. Incidental Findings with Genomic Testing: Implications for Genetic Counseling Practice. Current genetic medicine reports. 2015; 3(4):166–76. Epub 2015/11/14. doi: 10.1007/s40142-015-0075-9 PMID: 26566463

15. He M, Person TN, Hebbring SJ, Heinzen E, Ye Z, Schrodi SJ, et al. SeqHBase: a big data toolset for family based sequencing data analysis. Journal of medical genetics. 2015; 52(4):282–8. Epub 2015/01/15. doi: 10.1136/jmedgenet-2014-102907 PMID: 25587064

16. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine: official journal of the American College of Medical Genetics. 2015; 17(5):405–24. Epub 2015/03/06.

17. McCarty CA, Nair A, Austin DM, Giampietro PF. Informed consent and subject motivation to participate in a large, population-based genomics study: the Marshfield Clinic Personalized Medicine Research Project. Community genetics. 2007; 10(1):2–9. Epub 2006/12/15. doi: 10.1159/000096274 PMID: 17167244

18. McCarty CA, Chapman-Stone D, Derfus T, Giampietro PF, Fost N. Community consultation and communication for a population-based DNA biobank: the Marshfield clinic personalized medicine research project. American journal of medical genetics Part A. 2008; 146A(23):3026–33. Epub 2008/11/14. doi: 10.1002/ajmg.a.32559 PMID: 19006210

19. McCarty CA, Mukesh BN, Kitchner TE, Hubbard WC, Wilke RA, Burmester JK, et al. Intraocular pressure response to medication in a clinical setting: the Marshfield Clinic Personalized Medicine Research Project. Journal of glaucoma. 2008; 17(5):372–7. Epub 2008/08/16. doi: 10.1097/IJG.0b013e31815c5f3f PMID: 18703947

20. McCarty CA, Wilke RA, Giampietro PF, Wesbrook SD, Caldwell MD. Marshfield Clinic Personalized Medicine Research Project (PMRP): design, methods and recruitment for a large population-based biobank. Personalized medicine. 2005; 2(1):49–71.

21. Gullapalli RR, Lyons-Weiler M, Petrosko P, Dhir R, Becich MJ, LaFramboise WA. Clinical integration of next-generation sequencing technology. Clinics in laboratory medicine. 2012; 32(4):585–99. Epub 2012/10/20. doi: 10.1016/j.cll.2012.07.005 PMID: 23078661

22. Carter TC, He MM. Challenges of Identifying Clinically Actionable Genetic Variants for Precision Medicine. Journal of healthcare engineering. 2016; 2016. Epub 2016/05/20.

23. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research. 2010; 38(16):e164. Epub 2010/07/06. doi: 10.1093/nar/gkq603 PMID: 20601685

24. Fu W, O'Connor TD, Jun G, Kang HM, Abecasis G, Leal SM, et al. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature. 2013; 493(7431):216–20. Epub 2012/12/04. doi: 10.1038/nature11690 PMID: 23201682

25. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016; 536(7616):285–91. Epub 2016/08/19. doi: 10.1038/nature19057 PMID: 27535533

26. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic acids research. 2016; 44(D1):D862–8. Epub 2015/11/20. doi: 10.1093/nar/gkv1222 PMID: 26582918

27. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. American journal of human genetics. 2005; 76(5):887–93. Epub 2005/03/25. doi: 10.1086/429864 PMID: 15789306

28. Komiya Y, Habas R. Wnt signal transduction pathways. Organogenesis. 2008; 4(2):68–75. Epub 2009/03/13. PMID: 19279717

29. Giardiello FM, Yang VW, Hylind LM, Krush AJ, Petersen GM, Trimbath JD, et al. Primary chemoprevention of familial adenomatous polyposis with sulindac. The New England journal of medicine. 2002; 346 (14):1054–9. Epub 2002/04/05. doi: 10.1056/NEJMoa012015 PMID: 11932472

30. Markowitz SD, Bertagnolli MM. Molecular origins of cancer: Molecular basis of colorectal cancer. The New England journal of medicine. 2009; 361(25):2449–60. Epub 2009/12/19. doi: 10.1056/NEJMra0804588 PMID: 20018966

31. Chen Y, Li J, Yu X, Li S, Zhang X, Mo Z, et al. APC gene hypermethylation and prostate cancer: a systematic review and meta-analysis. European journal of human genetics: EJHG. 2013; 21(9):929–35. Epub 2013/01/10. doi: 10.1038/ejhg.2012.281 PMID: 23299921

32. Aoki K, Taketo MM. Adenomatous polyposis coli (APC): a multi-functional tumor suppressor gene. Journal of cell science. 2007; 120(Pt 19):3327–35. Epub 2007/09/21. doi: 10.1242/jcs.03485 PMID: 17881494

33. James CR, Quinn JE, Mullan PB, Johnston PG, Harkin DP. BRCA1, a potential predictive biomarker in the treatment of breast cancer. The oncologist. 2007; 12(2):142–50. Epub 2007/02/14. doi: 10.1634/theoncologist.12-2-142 PMID: 17296808

34. Wang B, Hurov K, Hofmann K, Elledge SJ. NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control. Genes & development. 2009; 23(6):729–39. Epub 2009/03/06.

35. Andersen TI, Borresen AL, Moller P. A common BRCA1 mutation in Norwegian breast and ovarian cancer families? American journal of human genetics. 1996; 59(2):486–7. Epub 1996/08/01. PMID: 8755943

36. Al-Sukhni W, Rothenmund H, Borgida AE, Zogopoulos G, O'Shea AM, Pollett A, et al. Germline BRCA1 mutations predispose to pancreatic adenocarcinoma. Human genetics. 2008; 124(3):271–8. Epub 2008/09/03. doi: 10.1007/s00439-008-0554-0 PMID: 18762988

37. Friedenson B. The BRCA1/2 pathway prevents hematologic cancers in addition to breast and ovarian cancers. BMC cancer. 2007; 7:152. Epub 2007/08/09. doi: 10.1186/1471-2407-7-152 PMID: 17683622

38. Jensen RB, Carreira A, Kowalczykowski SC. Purified human BRCA2 stimulates RAD51-mediated recombination. Nature. 2010; 467(7316):678–83. Epub 2010/08/24. doi: 10.1038/nature09399 PMID: 20729832

39. Gayther SA, Mangion J, Russell P, Seal S, Barfoot R, Ponder BA, et al. Variation of risks of breast and ovarian cancer associated with different germline mutations of the BRCA2 gene. Nature genetics. 1997; 15(1):103–5. Epub 1997/01/01. doi: 10.1038/ng0197-103 PMID: 8988179

40. Ozcelik H, Schmocker B, Di Nicola N, Shi XH, Langer B, Moore M, et al. Germline BRCA2 6174delT mutations in Ashkenazi Jewish pancreatic cancer patients. Nature genetics. 1997; 16(1):17–8. Epub 1997/05/01. doi: 10.1038/ng0597-17 PMID: 9140390

41. Hu C, Hart SN, Bamlet WR, Moore RM, Nandakumar K, Eckloff BW, et al. Prevalence of Pathogenic Mutations in Cancer Predisposition Genes among Pancreatic Cancer Patients. Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2016; 25(1):207–11. Epub 2015/10/21.

42. Edwards SM, Kote-Jarai Z, Meitz J, Hamoudi R, Hope Q, Osin P, et al. Two percent of men with early-onset prostate cancer harbor germline mutations in the BRCA2 gene. American journal of human genetics. 2003; 72(1):1–12. Epub 2002/12/11. PMID: 12474142

43. Side L, Taylor B, Cayouette M, Conner E, Thompson P, Luce M, et al. Homozygous inactivation of the NF1 gene in bone marrow cells from children with neurofibromatosis type 1 and malignant myeloid disorders. The New England journal of medicine. 1997; 336(24):1713–20. Epub 1997/06/12. doi: 10.1056/NEJM199706123362404 PMID: 9180088

44. Johnson MR, Look AT, DeClue JE, Valentine MB, Lowy DR. Inactivation of the NF1 gene in human melanoma and neuroblastoma cell lines without impaired regulation of GTP.Ras. Proceedings of the National Academy of Sciences of the United States of America. 1993; 90(12):5539–43. Epub 1993/06/15. PMID: 8516298

45. Baralle D, Mattocks C, Kalidas K, Elmslie F, Whittaker J, Lees M, et al. Different mutations in the NF1 gene are associated with Neurofibromatosis-Noonan syndrome (NFNS). American journal of medical genetics Part A. 2003; 119A(1):1–8. Epub 2003/04/23. doi: 10.1002/ajmg.a.20023 PMID: 12707950

46. Tassabehji M, Strachan T, Sharland M, Colley A, Donnai D, Harris R, et al. Tandem duplication within a neurofibromatosis type 1 (NF1) gene exon in a family with features of Watson syndrome and Noonan syndrome. American journal of human genetics. 1993; 53(1):90–5. Epub 1993/07/01. PMID: 8317503

47. Krauthammer M, Kong Y, Bacchiocchi A, Evans P, Pornputtapong N, Wu C, et al. Exome sequencing identifies recurrent mutations in NF1 and RASopathy genes in sun-exposed melanomas. Nature genetics. 2015; 47(9):996–1002. Epub 2015/07/28. doi: 10.1038/ng.3361 PMID: 26214590

48. Surget S, Khoury MP, Bourdon JC. Uncovering the role of p53 splice variants in human malignancy: a clinical perspective. OncoTargets and therapy. 2013; 7:57–68. Epub 2014/01/01. doi: 10.2147/OTT.S53876 PMID: 24379683

49. Ribeiro RC, Sandrini F, Figueiredo B, Zambetti GP, Michalkiewicz E, Lafferty AR, et al. An inherited p53 mutation that contributes in a tissue-specific manner to pediatric adrenal cortical carcinoma. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98(16):9330–5. Epub 2001/08/02. doi: 10.1073/pnas.161479898 PMID: 11481490

50. Colaluca IN, Tosoni D, Nuciforo P, Senic-Matuglia F, Galimberti V, Viale G, et al. NUMB controls p53 tumour suppressor activity. Nature. 2008; 451(7174):76–80. Epub 2008/01/04. doi: 10.1038/nature06412 PMID: 18172499

51. Rutherford J, Chu CE, Duddy PM, Charlton RS, Chumas P, Taylor GR, et al. Investigations on a clinically and functionally unusual and novel germline p53 mutation. British journal of cancer. 2002; 86(10):1592–6. Epub 2002/06/27. doi: 10.1038/sj.bjc.6600269 PMID: 12085209

52. Baker SJ, Fearon ER, Nigro JM, Hamilton SR, Preisinger AC, Jessup JM, et al. Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. Science. 1989; 244(4901):217–21. Epub 1989/04/14. PMID: 2649981

53. Hsu IC, Metcalf RA, Sun T, Welsh JA, Wang NJ, Harris CC. Mutational hotspot in the p53 gene in human hepatocellular carcinomas. Nature. 1991; 350(6317):427–8. Epub 1991/04/04. doi: 10.1038/350427a0 PMID: 1849234

54. Stein T, Crighton D, Boyle JM, Varley JM, White RJ. RNA polymerase III transcription can be derepressed by oncogenes or mutations that compromise p53 function in tumours and Li-Fraumeni syndrome. Oncogene. 2002; 21(19):2961–70. Epub 2002/06/26. doi: 10.1038/sj.onc.1205372 PMID: 12082526

55. Sun Y, Hegamyer G, Cheng YJ, Hildesheim A, Chen JY, Chen IH, et al. An infrequent point mutation of the p53 gene in human nasopharyngeal carcinoma. Proceedings of the National Academy of Sciences of the United States of America. 1992; 89(14):6516–20. Epub 1992/07/15. PMID: 1631151

56. Chen PL, Chen YM, Bookstein R, Lee WH. Genetic mechanisms of tumor suppression by the human p53 gene. Science. 1990; 250(4987):1576–80. Epub 1990/12/14. PMID: 2274789

57. Sonoyama T, Sakai A, Mita Y, Yasuda Y, Kawamoto H, Yagi T, et al. TP53 codon 72 polymorphism is associated with pancreatic cancer risk in males, smokers and drinkers. Molecular medicine reports. 2011; 4(3):489–95. Epub 2011/04/07. doi: 10.3892/mmr.2011.449 PMID: 21468597

58. Stacey SN, Sulem P, Jonasdottir A, Masson G, Gudmundsson J, Gudbjartsson DF, et al. A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. Nature genetics. 2011; 43(11):1098–103. Epub 2011/09/29. doi: 10.1038/ng.926 PMID: 21946351

59. Kyritsis AP, Bondy ML, Xiao M, Berman EL, Cunningham JE, Lee PS, et al. Germline p53 gene mutations in subsets of glioma patients. Journal of the National Cancer Institute. 1994; 86(5):344–9. Epub 1994/03/02. PMID: 8308926

60. Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. Proceedings of the National Academy of Sciences of the United States of America. 2015; 112(17):5473–8. Epub 2015/04/02. doi: 10.1073/pnas.1418631112 PMID: 25827230

61. Biesecker LG. Long QT Syndrome and Potentially Pathogenic Genetic Variants. Jama. 2016; 315(22):2467. Epub 2016/06/15.