



BRCA Share: A Collection of Clinical BRCA Gene Variants

Christophe Bérout, Stanley I Letovsky, Corey D Braastad, Sandrine M Caputo, Olivia Beaudoux, Yves Jean Bignon, Brigitte Bressac-de Paillerets, Myriam Bronner, Crystal M Buell, Gwenaelle Collod-Beroud, et al.

► To cite this version:

Christophe Bérout, Stanley I Letovsky, Corey D Braastad, Sandrine M Caputo, Olivia Beaudoux, et al.. BRCA Share: A Collection of Clinical BRCA Gene Variants. *Human Mutation*, 2016, Next Generation Sequencing and Human Genetic Disease, 37 (12), pp.1318-1328. 10.1002/humu.23113 . hal-01670197

HAL Id: hal-01670197

<https://hal.science/hal-01670197>

Submitted on 21 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

BRCA Share: A Collection of Clinical BRCA Gene Variants

Christophe Bérout, ^{1,2†} Stanley I. Letovsky, ^{3†} Corey D. Braastad, ^{4†} Sandrine M. Caputo, ^{5†} Olivia Beaudoux, ⁶ Yves Jean Bignon, ⁷ Brigitte Bressac-De Paillerets, ⁸ Myriam Bronner, ⁹ Crystal M. Buell, ⁴ Gwenaëlle Collod-Bérout, ¹ Florence Coulet, ¹⁰ Nicolas Derive, ⁵ Christina Divincenzo, ⁴ Christopher D. Elzinga, ⁴ Céline Garrec, ¹¹ Claude Houdayer, ^{5,12} Izabela Karbassi, ⁴ Sarab Lizard, ¹³ Angela Love, ⁴ Danièle Muller, ¹⁴ Narasimhan Nagan, ³ Camille R. Nery, ¹⁵ Ghadi Rai, ¹ Françoise Revillion, ¹⁶ David Salgado, ¹ Nicolas Sévenet, ¹⁷ Olga Sinilnikova, ^{18†} Hagay Sobol, ¹⁹ Dominique Stoppa-Lyonnet, ^{5,12} Christine Toulas, ²⁰ Edwin Trautman, ³ Dominique Vaur, ²¹ Paul Vilquin, ²² Katelyn S. Weymouth, ³ Alecia Willis, ²³ Laboratory Corporation of America Variant Classification Group, ^{3,23§} Quest Diagnostics Variant Classification Group, ^{4,15||} UNICANCER Genetic Group BRCA Laboratory Network, [#] Marcia Eisenberg, ²³ and Charles M Strom ^{15*}

¹Aix Marseille Univ, INSERM, GMGF, Marseille, France; ²APHM, Hôpital TIMONE Enfants, Laboratoire de Génétique Moléculaire, Marseille, France; ³Laboratory Corporation of America, Westborough, Massachusetts; ⁴Quest Diagnostics, Marlborough, Massachusetts; ⁵Service de Génétique, Département de Biologie des Tumeurs, Institut Curie, Paris, France; ⁶CHU et Institut Jean Godinot, Reims, France; ⁷Centre Jean Perrin, Clermont-Ferrand, France; ⁸Institut Gustave Roussy, Villejuif, France; ⁹CHU de Nancy-Brabois, Vandoeuvre-lès-Nancy, France; ¹⁰Groupe hospitalier Pitié-Salpêtrière, Assistance Publique-Hôpitaux de Paris, Laboratoire d'Oncogénétique et Angiogénétique moléculaire, Université Pierre et Marie Curie, Paris, France; ¹¹CHU, Institut de Biologie, Hôtel Dieu, Nantes, France; ¹²Université Paris Descartes, Paris, France; ¹³CHU de Dijon, Hôpital d'Enfants, Service de Génétique Médicale, Dijon, France; ¹⁴Centre Paul Strauss, Strasbourg, France; ¹⁵Quest Diagnostics, San Juan Capistrano, California; ¹⁶Centre Oscar Lambret, Unité d'Oncologie Moléculaire Humaine, Lille, France; ¹⁷Institut Bergonié, Bordeaux, France; ¹⁸Hospices Civils de Lyon and Centre Léon Bérard, Lyon, France; ¹⁹Institut Paoli-Calmettes, Marseille, France; ²⁰Institut Claudius Régau, Toulouse, France; ²¹Laboratoire de biologie et de génétique du cancer, CLCC François Baclesse, INSERM 1079 Centre Normand de Génomique et de Médecine Personnalisée, Caen, France; ²²Laboratoire de Biologie Cellulaire et Hormonale (CHU Arnaud de Villeneuve), Montpellier, France; ²³Laboratory Corporation of America, Research Triangle Park, North Carolina

For the Next Generation Sequencing special issue

ABSTRACT: As next-generation sequencing increases access to human genetic variation, the challenge of determining clinical significance of variants becomes ever more acute. Germline variants in the *BRCA1* and *BRCA2* genes can confer substantial lifetime risk of breast and ovarian cancer. Assessment of variant pathogenicity is a vital part of clinical genetic testing for these genes. A database

of clinical observations of *BRCA* variants is a critical resource in that process. This article describes *BRCA Share*TM, a database created by a unique international alliance of academic centers and commercial testing laboratories. By integrating the content of the Universal Mutation Database generated by the French Unicancer Genetic Group with the testing results of two large commercial laboratories, Quest Diagnostics and Laboratory Corporation of America (LabCorp), *BRCA Share*TM has assembled one of the largest publicly accessible collections of *BRCA* variants currently available. Although access is available to academic researchers without charge, commercial participants in the project are required to pay a support fee and contribute their data. The fees fund the ongoing curation effort, as well as planned experiments to functionally characterize variants of uncertain significance. *BRCA Share*TM databases can therefore be considered as models of successful data sharing between private companies and the academic world.

KEY WORDS: NGS; *BRCA1*; *BRCA2*; breast cancer; ovarian cancer; genetic databases; variant classification

Additional Supporting Information may be found in the online version of this article.

†These authors contributed equally to this work.

‡This author is deceased.

§Laboratory Corporation of America Variant Classification Group: Ekaterina Bogdanova, Wenjie (Jack) Chen, Sabrina Gardner, Hayk Hovhannisyan, Amanda Natalizio, Wei Song Kelly, and Ildiko Thibodeau.

||Quest Diagnostics Variant Classification Group: Alison Bright, Felicita Dubois, Forum Raval, Glenn Maston, Gwendolyn Bennett, Hannah Hayes, Jennifer Bogenschutz, Jennifer Lapierre, Katelyn Medeiros, Michele McCarthy, Domenic Previte, Ricardo Gutierrez, Andrew J. Farinas, Elaine H. Hiller, Jean R. Tirsch, and David Tsao.

#UNICANCER Genetic Group BRCA Laboratory Network: Françoise Bonnet, Natalie Jones, Virginie Buben, Michel Longy, Sophie Krieger, Laurent Castera, Nancy Uhrhammer, Jean-Philippe Peyrat, Aurélie Dumont Mélanie Léone, Nadia Boutry-Kryza, Pierre-Olivier Harmand, Pascal Pujol, Philippe Jonveaux, Joanna Sokolowska, Capucine Delnante, Virginie Guibert, Stéphane Béziau, Florent Soubrier, Errel Guillerm, Virginie Caux-Moncoutier, Lisa Golmard, Rosette Lidereau, Etienne Rouleau, Chantal Delvincourt, Joseph Abecassis, and Marine Guillaud-Bataille.

*Correspondence to: Charles M. Strom, Quest Diagnostics, 33608 Ortega Highway, San Juan Capistrano, CA 92675. E-mail: Charles.M.Strom@questdiagnostics.com

Introduction

Breast cancer is the most common malignancy in women in most countries with 464,000 new cases and 131,000 deaths each year in

Europe and 246,660 new cases and 40,450 deaths in the US [Ferlay et al., 2013]. Globally, it is estimated that more than 1 million women worldwide are diagnosed yearly with breast cancer and that more than 400,000 will die from the disease. Although most cases are sporadic, familial clustering has been reported and it is estimated that approximately 10% are likely to be hereditary. The known pathogenic mutationsⁱ from predisposing genes, including *BRCA1* (MIM# 113705) and *BRCA2* (MIM# 600185), account for 20%–25% of those familial forms [Pharoah et al., 2002].

As breast cancers represent a significant public health problem, many national programs have been dedicated to cancer prevention and early diagnosis. They use various risk assessment (Gail [Chay et al., 2012], Claus [Fischer et al., 2013] and Tyrer-Cuzick [Boughey et al., 2010]) or probability models such as the BReast Cancer risk PRediction mOdel (BRCAPRO) [Mazzola et al., 2015]. The average cumulative risks in *BRCA1* mutation carriers by the age of 70 are 65% (95% confidence interval 44%–78%) for breast cancer and 39% (18%–54%) for ovarian cancer. The corresponding estimates for *BRCA2* mutation carriers are 45% (31%–56%) and 11% (2.4%–19%) [Antoniou et al., 2003]. Women with a high lifetime risk ($\geq 20\%$) are eligible for genetic counseling and testing mainly for mutations of the *BRCA1* [Hall et al., 1990] and *BRCA2* [Wooster et al., 1994] cancer susceptibility genes.

Professional organizations in North America and Europe have published clinical practice guidelines with respect to BRCA counseling and testing [Balmaña et al., 2011; Gadzicki et al., 2011; Graña et al., 2011; National Institute for Health Care Excellence 2013; Cancer Institute NSW 2015; Holter et al., 2015]. The United States National Comprehensive Cancer Network guidelines recommend genetic counseling and/or testing for women that fulfill one of the following criteria: breast cancer at the age of 50 years or younger; a bilateral breast cancer; triple-negative breast cancer (estrogen receptor negative, progesterone receptor negative, Her2Neu negative); breast cancer at any age with close relatives with breast/ovarian/pancreatic cancer; breast cancer with Ashkenazi Jewish ancestry; male breast cancer; known mutation for a breast cancer susceptibility gene within the family, or women with a history of ovarian cancer. These recommendations have resulted in a rapid uptake and utilization of BRCA testing in clinical practice and a consequent large increase in variant classification burdens among laboratories worldwide.

The process of assessing clinical significance of germline variants has been codified in guidelines [Bahcall 2015] and formalized in probabilistic models for evidence integration [Goldgar et al., 2004; Lovelock et al., 2007]. Evidence may come from family studies, functional studies, algorithmic predictions, patterns of allelic co-occurrences within individuals, and other sources. Clinical significance is typically assessed, per the ACMG/AMP guidelines [Richards et al., 2015], by assigning the variant to one of the following classifications: *Pathogenic*, *Likely Pathogenic*, *VUS* (*Variant of Uncertain Significance*), *Likely Benign*, and *Benign*. Integrating evidence to produce a variant classification currently relies on the judgment of geneticists with diverse areas of expertise who primarily assess the information available from the scientific literature and databases, as well as algorithmic predictions, to arrive at a variant classification. Several studies reporting inter reviewer agreements in classifications among experts with widely differing outcomes ranging from “moderate” (Gross k , 0.52) to “high” (K-alpha, 0.91) have been published. Specifically, although a higher concordance of po-

tentially clinically actionable classifications (*Pathogenic* or *Likely Pathogenic*) derived using the ACMG-AMP schemes with those reported in locus-specific databases (LSDBs) and ClinVar have been reported, the concordance rates for classifications deemed as non-actionable (*VUS*, *Likely Benign*, and *Benign*) were lower [Maxwell et al., 2016]. Amendola et al. (2016) reported a 79% concordance when nine participating laboratories classified a set of variants using their internally developed methods and the ACMG-AMP criteria. However, only a 34% concordance for either classification system across laboratories was observed. After consensus discussions and detailed review of the ACMG/AMP criteria, concordance increased to 71%, showing that a common framework can help resolve differences in classifications.

It can require many hours to classify a variant when literature review is required. The results and conclusions often need to be fully vetted with a critical eye toward translation to the mechanism(s) and presentation(s) of the disease and phenotype under consideration. Understandably, this classification adds significant skilled labor cost to the testing process. Fully automated methods for variant classification are not ready for clinical practice. Median time required to curate available lines of evidence and classify variants have been reported to vary between 54 min (range 5–233 min) [Dewey et al., 2014], to 37 min per variant (range 1–175 min) [Amendola et al., 2015]. Furthermore, variant classifications are not static; classifications may change over time as the underlying scientific knowledge changes, requiring a periodic re-evaluation of the literature and other evidence prior to clinical reporting. CLIA laboratory regulations in the United States require each laboratory to be responsible for its own reports, and to update their classifications to reflect the most recent information, which limits the ability to use a paradigm for static classifications (and hence share the cost of classification) across laboratories.

Variant databases play a critical role in the classification process, by providing summary information about the state of scientific knowledge, with clinical assessments of the variant by other laboratories, with supporting evidence, and with data on other patients carrying the variant. LSDBs [Claustres et al., 2002] have been established for many disease genes, whereas other databases, including OMIM, ClinVar, LOVD, and Clinvar collect variant information across genes. Recently, network federation protocols such as Beacon [Krol 2015] and Matchmaker exchange [Philippakis et al., 2015] have been established to facilitate peer-to-peer exchange of variant data. Some commercial software products, including *GeneInsight Network* and Agilent’s *Cartagenia*, are using a similar approach to enable sharing of variant classifications within their user communities.

In the context of BRCA testing, databases available today include: the Breast Cancer Information Core Database (BIC) [Szabo et al., 2000], LOVD [Fokkema et al., 2011], ClinVar [Landrum et al., 2014], BRCA Exchange [Global Alliance for Genomics and Health 2016], the ARUP databases (www.arup.utah.edu/), and the Universal Mutation Database (UMD)-BRCA1/2 databases [Caputo et al., 2012]. Note that Myriad Genetics, a commercial laboratory, has developed its own database whose quality and content could not be evaluated as it is proprietary and not accessible to the community. These databases have been developed with different goals and have different curation processes, data quality, and content as underlined by Vail et al. [2015]. These authors compared BIC, ClinVar, HGMD (paid version), LOVD, and the UMD databases and concluded that these differences inhibit their wider use in clinical practice.

In order to provide high-quality sustainable *BRCA1* and *BRCA2* databases, the BRCA ShareTM public/private partnership was co-founded by Quest Diagnostics and the French National Institute

ⁱWe use the term *variant* to refer generically to differences from the reference sequence, and *mutation* to refer to pathogenic and likely pathogenic variants.

of Health and Medical Research (Inserm) in April 2015, with Laboratory Corporation of America Holdings as the first commercial participant. The goal of the initiative is to share clinical, genetic, epidemiological, and biological data on *BRCA* variants, particularly VUS, in order to improve the quality of laboratory diagnostics to better predict which individuals are at risk of developing hereditary breast and ovarian cancers, and to accelerate research on *BRCA* gene variants. BRCA Share™ builds on an efficient data curation process [Caputo et al., 2012] that follows international recommendations [Eccles et al., 2015].

The BRCA Share™ database now contains over 6,200 total *BRCA* variants, an increase of nearly 30% compared with the previous UMD-BRCA1/2 databases. Of these variants, 334 are newly identified pathogenic or likely pathogenic, increasing by about 20% the total number of pathogenic or likely pathogenic variants to 1,826.

To our knowledge, BRCA Share™ is the first example of a successful public/private partnership for LSDBs that ensures free access to the database for research purposes while creating a user group of commercial partners that will collectively endorse running and development costs while also supporting functional tests to rapidly and efficiently classify VUS.

With increased adoption of whole exome and whole genome sequencing in clinical practice, it is believed that many patients might benefit from the indirect discovery of variants in clinically actionable genes. In a study of 1,000 individuals (500 European- and 500 African-descent participants randomly selected from the National Heart, Lung, and Blood Institute Exome Sequencing Project), Dorschner et al. [2013] report that ~3.4% of European-descent adults and ~1.2% of African-descent adults can be expected to have actionable highly penetrant pathogenic or likely pathogenic mutations identified by exome sequencing. Among those, three of the 1,000 participants had a *BRCA1/2* pathogenic mutation [Dorschner et al., 2013]. This ratio was confirmed by a subsequent study by Amendola et al. [2015].

With the involvement of the major diagnostic companies in US and reference centers in France, for the first time, all variants will be shared and further annotated by a group of experts [Caputo et al., 2012]. The availability of this database is expected to provide researchers and geneticists rapid access to supporting evidence for classification.

Materials and Methods

The BRCA Share Databases

The BRCA Share™ databases are derived from the UMD-BRCA1 and UMD-BRCA2 databases [Caputo et al., 2012], based on 20 years' experience of clinical *BRCA* testing by UGG, and augmented by the results of clinical *BRCA* testing by Quest and LabCorp. The UMD software [Bérout et al., 2005] has been modified to accommodate new features including a registration process and secure access to the databases under individual login IDs.

In order to integrate data from different sources and ensure an optimal curation process, a common template was used for data submission. It contains information related to the individual and related sample(s), including submitter, de-identified subject and family IDs, subject demographic information, disease status of subject and relatives, availability of cell line, tumor, or other physical materials, and so on. Subject and family identifiers are reassigned to yield globally unique anonymous identifiers which can be mapped back to the submitter's anonymous identifiers for future updates. Information related to variants includes the

next-generation sequencing (NGS) screening type (whole coding sequence; targeted; single site), sequencing platform, DNA, RNA and protein HGVS names for the variant [den Dunnen et al., 2016], ID of transcript used to derive the names, the variant class (missense, nonsense, frameshift, intronic, rearrangement, insertion, deletion, isosemantic a.k.a. synonymous), the submitter pathogenicity assessment, and supporting evidence collected from the literature, from functional assays and from in silico tools. Variant nomenclature is automatically validated by the UMD software and the HGVS genomic DNA variant name (g.) is generated using the GRCh37 reference sequence. Allele frequencies are automatically collected from dbSNP [Sherry et al., 2001], the Exome Aggregation Consortium including TCGA data (<http://exac.broadinstitute.org/>), the 1000 genomes and the 6,500 exomes of the Exome Variant Server (<http://evs.gs.washington.edu/EVS/>). Finally, a link to the UMD-Predictor system [Salgado et al., 2016] is provided to ensure access to the most recent in silico predictions from this system.

The BRCA Share™ databases are accessible at <http://umd.be/BRCA1/> and <http://umd.be/BRCA2/>. They integrate a new Web interface that uses dynamic tabs and forms. It was developed using the JQuery library (<http://jquery.com/>). New graphical displays were also created, using the D3.js library (<http://d3js.org/>), featuring dynamic zoom and signals highlights.

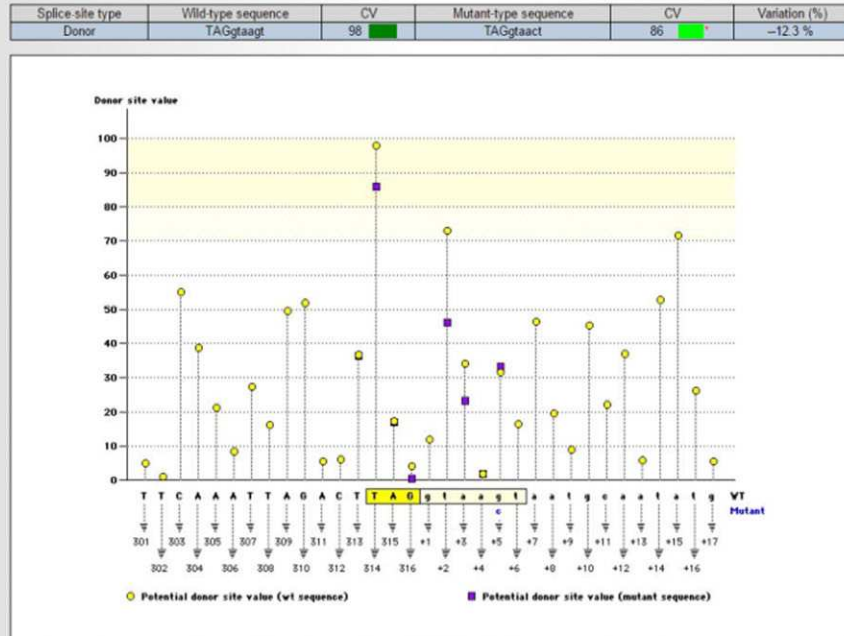
As an example of the depth of annotation provided by these databases, consider the *BRCA2* c.316+5G>C (IVS3+5G>C) variant, listed as pathogenic in BRCA Share™. Figure 1 shows a snippet from BRCA Share™ for this mutation. It provides brief summaries for descriptions in the published literature, in silico splice predictions, frequency, co-occurrence, and even classifications from other public databases. Supporting evidence for pathogenic classification including references to functional studies describing aberrant splicing, frequency, and co-occurrence in patients are included in the variant record. This allows the investigator to review current data and literature for validation of the variant's classification and not solely rely on the classification from other clinical laboratories.

Two new features have been added on the welcome page to highlight changes in classification status: the "Variants reclassification" and "Variants classification" links. The first one gives access to the list of reclassified variants during the last 6 months, whereas the second gives access to variants that have been classified during the last 6 months. Note that during the registration process, the user can select to receive alerts by email when variant reclassification occurs.

Curation

BRCA Share curation is based on processes developed by the Unicancer Genetic Group (UGG) *BRCA* network that ensure clinical grade quality. It relies on a classification working group which meets regularly to provide expert curation. VUS classification is conducted using a combination of available data: frequencies in the general population, in silico predictions using bioinformatics tools at both splice and protein levels, causality or neutrality scores published in the literature [Goldgar et al., 2004; Easton et al., 2007], functional domain [Millot et al., 2012; Guidugli et al., 2013], co-occurrence with causal mutations in the same gene, co-segregation analyses in French families and published results of functional tests or splicing modeling data as described in Caputo et al. [2012]. Evidence for classification are integrated in a likelihood model previously described [Goldgar et al., 2004]. The classifications displayed in the database are limited to those derived by the UGG. Periodically, discordant classifications among participating laboratories are marked with an asterisk to alert the user toward a closer review.

BRCA Share™ (formerly UMD-BRCA2 mutations database)
Mutation c.IVS3+5G>C (c.316+5G>C)



Data for this mutation

(Pathogenic mutations, Unclassified variants, Non-pathogenic variation)

Co-occurrence of BRCA1 and BRCA2 mutations

Sample ID	BRCA2	BRCA1
—121877A001	-	-
—58765A001	-	c.2521C>T (p.Arg841Trp)
—84189A001	-	-
—12765	-	-
—1121060	-	-
—DE45	-	-
—EB65	-	-
—EM03	-	-
—FV26	-	-
—14-47/17-09318/14-1	-	-
—Z4	-	-
—1071	-	-
—1285	-	-
—1416	-	-
—424	-	-
—02-2892/14-7	-	-

Evidences

Analysis	Result	Date	Origin	PMID/dbSNP
Ex vivo splicing assay	Absence of wt product and skipping exon 3 (full exclusion). Interpretation: pathogenic	22/09/09	Bonnet et al., JMG, 2008	18424508
MLPA, QMPSF, CGH, RT-PCR, Allelic expression, Competitive quantitative RT-PCR, p	The results are similar to those reported for a mutation in intron 3, c.316+5 G>C (which was analyzed in a mini-gene system) and the Portuguese Alu insertion, both of which result in total splicing out of exon 3 on the mutated allele. Consequently, these data, taken together, support the hypothesis of a pathogenic effect when there is exclusive synthesis of the delta3-transcript from one allele.	22/09/11	Muller D. et al., BMC 2011	21939546
COsegregation studies	Deleterious (COVAR)	19/03/15	GGC	
Population frequency in ExAC (exac03)	NA	19/04/16	Automatic annotation by UMD	
dbSNP	NA	19/04/16	Automatic annotation by UMD	
Population frequency in 1000 genomes project (1000g2014oct_all)	NA	19/04/16	Automatic annotation by UMD	
Population frequency in ESP (esp6500si_all)	NA	19/04/16	Automatic annotation by UMD	

Curators' annotation

Biological significance	Validated by	Date	Comment
5 - Causal	GGC	19/03/15	dbSNP rs81002840, LOVD // ClinVar: RCV000044131.2 (Not provided - Familial cancer of breast).

Figure 1. Classification information for BRCA2 c.316+5G>C (IVS3+5G>C) is present in BRCA Share™ database with annotations of class 5.

Table 1. Current Counts of Records (Variant Occurrences) and Unique Variants in BRCA Share™ by Pathogenicity

	BRCA1				BRCA2				Combined			
	Records		Unique variants		Records		Unique variants		Records		Unique Variants	
1. Neutral (Benign)	70,388	88.2%	114	4.5%	45,221	78.3%	123	3.3%	115,609	90.2%	237	6.4%
2. Likely Neutral (Likely Benign)	944	1.2%	142	5.6%	2,487	4.3%	218	5.9%	2,487	1.9%	360	9.7%
3. VUS	2,828	3.5%	1,423	56.0%	5,742	9.9%	2,408	64.9%	5,742	4.5%	3,831	103.2%
4. Likely causal (Likely Pathogenic)	39	0.0%	21	0.8%	46	0.1%	34	0.9%	46	0.0%	55	1.5%
5. Causal (Pathogenic)	5,626	7.0%	843	33.1%	4,276	7.4%	928	25.0%	4,276	3.3%	1,771	47.7%
Total	79,825		2,543		57,772		3,711		128,160		6,254	

Table 2. Current Counts of Records and Unique Variants in BRCA Share™ Broken Down by Variant Type

	BRCA1				BRCA2				Combined			
	Records		Unique mutations		Records		Unique mutations		Records		Unique mutations	
	#	%	#	%	#	%	#	%	#	%	#	%
Total	79,825		2,543		57,772		3,711		137,597		6,254	
Large rearrangements (>1 exon)	458	0.6%	79	3.1%	59	0.1%	32	0.9%	517	0.4%	111	1.8%
Deletions	372	0.5%	58	2.3%	37	0.1%	24	0.6%	409	0.3%	82	1.3%
Insertions	86	0.1%	21	0.8%	22	0.0%	8	0.2%	108	0.1%	29	0.5%
Small deletions and insertions	3,145	3.9%	531	20.9%	3,115	5.4%	702	18.9%	6,260	4.5%	1,233	19.7%
Small deletions	2,234	2.8%	389	15.3%	2,538	4.4%	530	14.3%	4,772	3.5%	919	14.7%
Small insertions	911	1.1%	142	5.6%	577	1.0%	172	4.6%	1,488	1.1%	314	5.0%
Point mutations	73,420	92.0%	1,264	49.7%	38,395	66.5%	2,195	59.1%	111,815	81.3%	3,459	55.3%
Missense	72,227	90.5%	1,068	42.0%	37,185	64.4%	1,964	52.9%	109,412	79.5%	3,032	48.5%
Nonsense	1,193	1.5%	196	7.7%	1,210	2.1%	231	6.2%	2,403	1.7%	427	6.8%
Intronic mutations	2,593	3.2%	628	24.7%	16,103	27.9%	743	20.0%	18,696	13.6%	1,371	21.9%
Splice sites (<10 bp from exon)	809	1.0%	167	6.6%	827	1.4%	172	4.6%	1,636	1.2%	339	5.4%
Mid-intronic mutations	1,784	2.2%	461	18.1%	15,276	26.4%	571	15.4%	17,060	12.4%	1,032	16.5%
Indels	209	0.3%	41	1.6%	100	0.2%	39	1.1%	309	0.2%	80	1.3%

Results

Usage and Content Statistics

The BRCA Share databases have been accessible online since July 2015. In this first year, the number of registrants has grown regularly to reach 1,119 users and the sites have been queried on average 17,000 times/month.

As a result of the establishment of the BRCA Share™ effort and the addition of clinical variants from Quest and LabCorp, the number of *BRCA1* and *BRCA2* variants has grown from UMD's total of 4,838 at the time BRCA Share was launched, to its current total of 6,254 unique variants, an increase of 29.3%.

The current count of records and variants are shown in Tables 1 and 2, and broken down by pathogenicity assessment and variant type, respectively. The distribution by variation type and pathogenicity assessment is available in Supp. Table S1.

Overlap and Classification Agreement with other Collections

The Venn diagrams in Figure 2(A and B) show the counts of *BRCA1* and *BRCA2*, variants, respectively, shared with ARUP (www.arup.utah.edu/), ClinVar [Landrum et al.,2014], the ENIGMA collection [Spurdle et al., 2012], and LOVD [Fokkema et al., 2011]. The BRCA Share™ collection currently is the second largest, behind ClinVar; both contain a large number of variants not present in the other. Note that ARUP only reports class 5 variants; its overlap with other systems would be more favorable if limited to these. Discordant classifications (2 or more classes difference) within ClinVar were treated as VUS for the purposes of

comparison. Classification comparisons between the two largest sites, BRCA Share™ and ClinVar, are shown in Figure 3. Seventy-four percent of the variants are classified identically between these two sites. None of the discordances are of the clinically actionable type where a variant goes from likely pathogenic/pathogenic to likely benign/benign. Total VUS discrepancies are shown in the margins. The largest discordance categories are for variants classified as class 2 (Likely Benign) in ClinVar versus class 3 (VUS) in BRCA Share™ and class 5 (Pathogenic) in BRCA Share™ versus class 3 (VUS) in ClinVar. In addition, the comparison of variants classified as class 4/5 in BRCAShare™, ARUP, and ClinVar reveals that BRCAShare™ and ARUP are more in agreement with each other (499/546; 91.4%) than ARUP and ClinVar (400/1,062; 37.7%), possibly reflecting small differences in classification criteria.

Pairwise comparisons among BRCA Share™, ClinVar, and ARUP show that BRCA Share™ and ClinVar agree on 72% of classifications, BRCA Share™ and ARUP on 81%, and ARUP and ClinVar on 60% of shared variants. Twenty four percent of variants classified as VUS by BRCA Share™ are classified as some other category by ClinVar, whereas 19% of variants classified as VUS by ClinVar are classified otherwise by BRCA Share™.

Variant Frequencies

The BRCA Share™ database affords an opportunity to characterize the spectrum of BRCA variation in some detail, to gain understanding the relationships between frequency, type, and classification. Although population frequency is not directly available from the BRCA Share™ collection, it can be estimated from the database. This collection incorporates a number of biases of sampling and reporting, including emphasis on affected subjects and

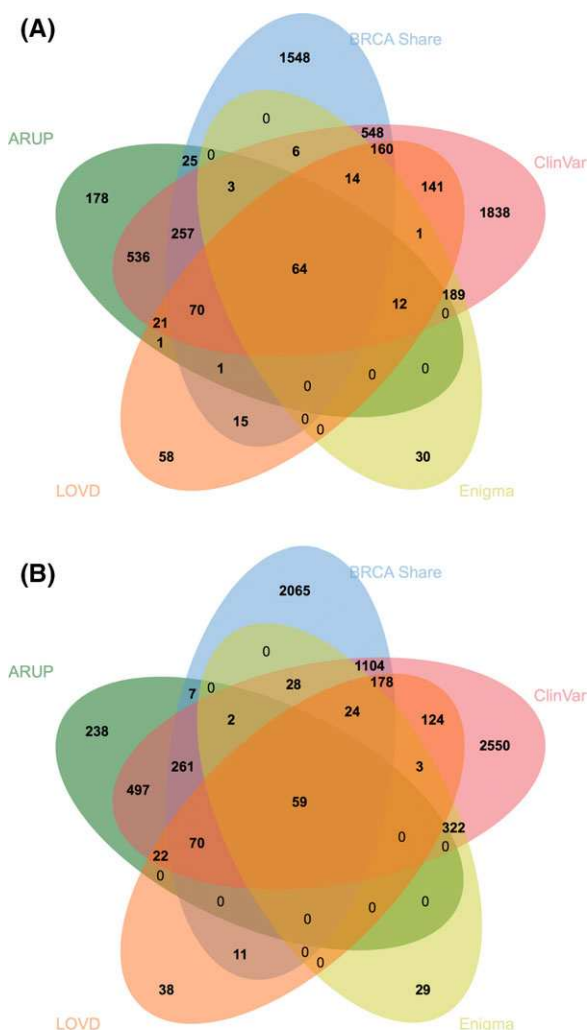


Figure 2. Venn diagram showing sharing of (A) BRCA1 and (B) BRCA2 variants between collections.

their relatives, leading to an emphasis of deleterious alleles. We estimated population frequencies for BRCA ShareTM variants using a log-log regression of frequencies from the Exome Aggregation Consortium (ExAC) dataset against BRCA ShareTM occurrence counts. The ExAC frequencies are based on a collection of 60,706 unrelated individuals sequenced as part of various disease-specific and population genetic studies for which an aggregate frequency is provided [<http://exac.broadinstitute.org/>]. As ExAC is not expected to be enriched for obligate carriers from families with a strong history of breast and ovarian cancer, the contribution of pathogenicity biases within this dataset are minimized. As shown in Figure 4(A), BRCA ShareTM occurrence counts correlate well with aggregated ExAC frequencies overall, especially for common variants. The distribution of pathogenic variants is skewed higher in BRCA ShareTM compared with ExAC, presumably reflecting the collection biases. A regression on benign variants only provides better agreement over most of the dynamic range than based on all variants. This regression was used to estimate frequencies for all variants, including those not shared with ExAC, in Figure 4(B–D). Figure 4B shows a breakdown of classifications in log10 estimated frequency bins, showing the increase in VUS and pathogenic rates among rare variants. Note that sample size effects cause the left most bin to be a truncation of unobserved rarer frequency bins to the left. Figure 4(C) shows the fractional representation of variant types in each log frequency bin, whereas Figure 4(D) shows the same for VUS only.

Co-occurrences

Co-occurrence of a VUS with a known disease-causing mutation, either from the same or the other BRCA gene, can provide evidence to classify the variant, especially if it is reported in multiple patients [Goldgar et al., 2008; Cherbal et al., 2012; Santos et al., 2014]. Co-occurrence data are used in two ways: in a Posterior Probability calculation [Goldgar et al., 2004], as it is one of its elements, and as a standalone criterion for classification. Indeed, if a given variant co-occurs with at least two different pathogenic variants in the same gene or is demonstrated co-occurring in trans with a pathogenic variant of the same gene, it is classified as Neutral. This criterion

	BRCA Share				
ClinVar	1	2	3	4	5
1	151	61	47		
2	19	92	243		
3	35	74	1203	4	174
4			10		39
5			81	11	608

Figure 3. Classification comparisons for BRCA1 and 2 variants between ClinVar and BRCA ShareTM. Counts on the diagonal (yellow) represent variants with concordant classifications between the two databases, whereas those in orange represent variants that are VUS in one database but not in the other.

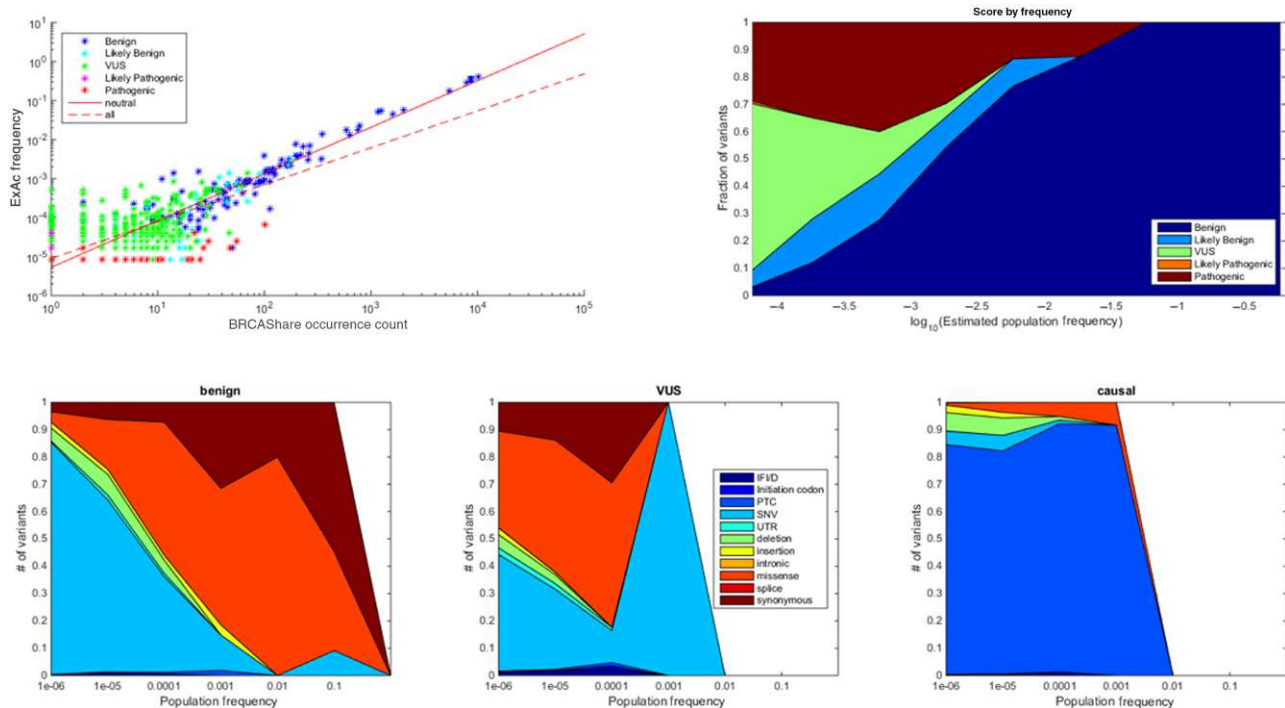


Figure 4. **A:** ExAC aggregate frequency versus BRCA ShareTM occurrence count, colored by pathogenicity, with log–log regression lines based on benign variants only (solid line) and all variants (dashed line). The asymmetric distribution of pathogenic variants reflects the pathogenic bias of the BRCA Share collection compared with the ExAC. The solid regression line was used to estimate population frequencies for BRCA Share variants in **B–D**; collection biases imply that frequencies of pathogenic variants are likely overestimated in BRCA Share. **B:** Classifications of unique variants in BRCA Share binned by log₁₀ estimated frequency, showing the predominance of VUS and pathogenic among rare variants. **C–E:** proportion of variants of each type broken down pathogenicity group (likely grouped with definite).

requires that the patient has been examined by a clinical geneticist to exclude Fanconi-like for *BRCA2*.

Since the BRCA ShareTM databases preserve patient-centric information, they allow co-occurrences to be readily queried, as shown in Figure 5. During curation, such evidence can provide arguments against pathogenicity, as illustrated in Table 3. For each *BRCA1* VUS from a specific sample, the table displays co-occurring pathogenic or likely pathogenic *BRCA1* mutations. We can distinguish cases in which the same pair of variants co-occur repeatedly, as illustrated for *BRCA2* c.324T>C (p.Asn108Asn), which is specifically associated with the c.2612C>A (p.Ser871X) pathogenic mutation. This might suggest a common origin of the patients or the presence of both variants on the same allele. By contrast, the *BRCA2* c.2350A>G (p.Met784Val) variant was reported in two patients from different laboratories. Each of them also harbors a different pathogenic *BRCA2* mutation: c.6952C>T (p.Arg2318X) and c.5771_5774delTTCA (p.Ile1924ArgfsX38). This indicates that those mutations are probably in trans and that these patients are unrelated. Additionally, this is a strong evidence to consider the *BRCA2* c.2350A>G (p.Met784Val) variant as non-pathogenic. Thus, the co-occurrence data facilitate weighting of evidence based upon the number of samples displaying a co-occurrence with another pathogenic variant.

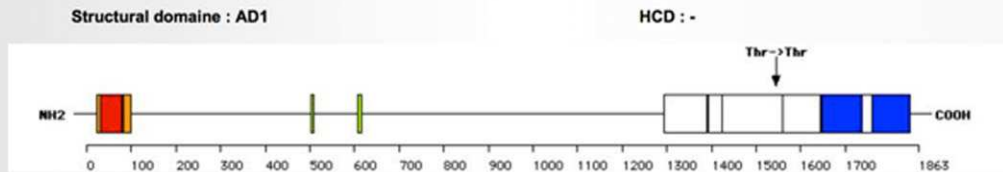
Classification Discordances

The initial combined dataset contained 687 variants that were shared between two or more of the three participating laboratories, 67% ($N = 457$) of which were in *BRCA1* and 33% ($N = 230$) were in *BRCA2*. Classifications were identical across the

laboratories for 57% of the shared variants, whereas 69% were concordant when Likely Benign is grouped with Benign and Likely Pathogenic with Pathogenic. The remaining 31% included variants that crossed over from a VUS to a non-VUS category between laboratories. No drastic differences in classifications (i.e., Benign/Likely Benign vs. Pathogenic/Likely Pathogenic) with a potential to adversely impact patient outcomes were identified. Concordance rates were higher in *BRCA1* than *BRCA2*, for example, exact concordances were 76% and 53%, respectively. We emphasize that these are per variant discordance rates and not per patient; since the discordances involve rare variants, per patient discordance rates would be far lower.

A process to resolve these discordances is underway, involving periodic review of the classification evidence by all the contributing laboratories. Initial review of 148 *BRCA1* discordances revealed that most were attributable to changes in the available evidence since the original classification; after review the number was significantly reduced to 37, or 8.1% of the *BRCA1* shared variants. Of these 37 discordances, 17 were attributable to differences in the approach toward classification of rare synonymous variants in the absence of other supportive evidence, with some groups classifying these as VUS and others as Likely Benign. The remaining 20 variants (4.4%) are all missense (18) or intronic (2), and are still under review, as are the *BRCA2* variants. BRCA Share displays a primary classification for each variant, though submitter classifications are accessible. The primary classification displayed is from the French Consortium (UGG) if available or else the closest to VUS of the submitted variants is displayed. Discordant classifications among participating laboratories are marked with an asterisk.

BRCA Share™ (formerly UMD-BRCA1 mutations database) Mutation p.Thr1548Thr (c.4644G>A)



[Alignment for the surrounding region](#) (Show/Hide)

Data for this mutation

Bioinformatics predictions of the mutation's impact

Conservation (0-1)	SIFT (0-1)	Grantham (0-215)	Max cross species difference (0-215)	UMD-predictor (0-100)
0.83	0.85 (non pathogenous)	0	81	22 (Polymorphism)

(Pathogenic mutations, Unclassified variants, Non-pathogenic variation)

Co occurrence of BRCA1 and BRCA2 mutations

Sample ID	BRCA1	BRCA2
---DB15	-	-
---1526	c.5260G>T (p.Glu1754X)	-
---668	-	-
---996	-	-
---86EE54FQ	c.342_343delTC (p.Pro115X)	-
---1-5479	-	-
---33-1-2014	-	-
--PrzKtUsJQEOjF	-	-
--ZZVjQiBOXYJHD	-	-

Evidences

Analysis	Result	Date	Origin	PMID/dbSNP
Co-occurrence with deleterious mutation	Polymorphism	7/02/11	Judkins T. et al., Cancer Res, 2005	16267036
Population frequency in ESP (esp6500sl_all)	0.000154	21/06/16	Automatic annotation by UMD	
Population frequency in 1000 genomes project (1000g2014oct_all)	NA	21/06/16	Automatic annotation by UMD	
dbSNP	rs28897692	21/06/16	Automatic annotation by UMD	rs28897692
Population frequency in ExAC (exac03)	4.12e-05	21/06/16	Automatic annotation by UMD	

Curators' annotation

Biological significance	Validated by	Date	Comment
3 - UV	GGC	17/03/15	dbSNP rs28897692 C/T. ESP: EA: T=0.02% - AA: T=0.00%.

Figure 5. Example of co-occurrence information provided for patients harboring the BRCA1 c.4644G>A mutation. This mutation has been reported in nine patients, two of them also harboring a pathogenic mutation in the BRCA1 gene: c. 5260G>T (p.Glu1754X) for sample --1526 and c.342_343delTC (p.Pro115X) for patient---86EE54FQ. Also note the availability of in silico predictions above the co-occurrence table.

Discussion

Since its inception a little over a year ago, the BRCA Share™ initiative has significantly augmented the UMD dataset collected by the UGG, by leveraging the clinical sample flow of two large laboratories. The combination of US and European clinical samples has likely increased the ethnic diversity of the BRCA Share™ sample, while still leaving large portions of human diversity undersampled.

The fact that BRCA Share™ and ClinVar, the two largest collection sites, each have a substantial number of variants not present in the other, indicates that no existing collection has yet to uncover all the variants in the human population.

Classification agreement within the participating BRCA Share™ groups was generally higher than previous reports, elsewhere, with most discordances being attributable either to incorporation of newer data or *likely* versus *VUS* differences.

The BRCA Share™ funding strategy, in which academic researchers have free access, while commercial labs pay support fees, may provide a sustainable model in a time when public funding for biological databases is under stress [Check Hayden 2016; Kaiser 2016]. Public collections such as ClinVar and GA4GH provide an alternative model with fewer restrictions on access, but they rely on government funding. ClinVar has historically functioned as a submission archive, without requiring resolution of conflicting

Table 3. Selected Examples of Co-occurrence of BRCA1 Pathogenic Mutations in Samples with a VUS BRCA1 Mutation

HGVS c.	HGVS p.	Sample ID	Class	Pathogenic BRCA1 mutation
c.28A>G	p.Thr10Ala	04-3625	3	c.2701delC (p.Ala902LeufsX2)
c.324T>C	p.Asn108Asn	02-2012	3	c.2612C>A (p.Ser871X)
c.324T>C	p.Asn108Asn	31-3504	3	c.2612C>A (p.Ser871X)
c.324T>C	p.Asn108Asn	01-65884A001	3	c.2612C>A (p.Ser871X)
c.324T>C	p.Asn108Asn	14-829	3	c.2612C>A (p.Ser871X)
c.2320A>G	p.Thr774Ala	15-676	3	c.5576_5579delTTAA (p.Ile1859LysfsX3)
c.2350A>G	p.Met784Val	01-82103A001	3	c.6952C>T (p.Arg2318X)
c.2350A>G	p.Met784Val	02-20417	3	c.5771_5774delTTCA (p.Ile1924ArgfsX38)
c.2416G>C	p.Asp806His	FYxqoGbWhaFkaYn	3	c.IVS6-2A>G (c.476-2A>G)
c.2751A>G	p.Val917Val	20-14YW68IE	3	c.8904delC (p.Val2969CysfsX7)
c.2780T>C	p.Met927Thr	07-A633	3	c.1310_1313delAAGA (p.Lys437IlefsX22)
c.2837A>G	p.Asp946Gly	33-008FCF1006/33-4-292	3	c.2701delC (p.Ala902LeufsX2)
c.2919G>A	p.Ser973Ser	12-FK16	3	c.1184G>A (p.Trp395X)
c.3152T>C	p.Leu1051Ser	01-62951A001	3	c.6656C>G (p.Ser2219X)
c.3170A>G	p.Lys1057Arg	05-026081	3	c.IVS15+1G>A (c.7617+1G>A)
c.3226G>A	p.Val1076Ile	33-1-5524	3	c.3159delA (p.Asp1054IlefsX6)
c.3304A>T	p.Asn1102Tyr	pfayNycYdTDDJm	3	c.1929delG (p.Arg645GlufsX15)
c.3304A>T	p.Asn1102Tyr	VrnAyEGaCFDesIG	3	c.1929delG (p.Arg645GlufsX15)
c.3445A>G	p.Met1149Val	11-2009-012	3	c.2092delC (p.Leu698TyrfsX32)
c.3539A>G	p.Lys1180Arg	02-11705	3	c.8070_8071dup (p.Ser2691PhefsX4)

*First column: cDNA HGVS nomenclature of the mutation; second column: protein HGVS nomenclature of the mutation; third column: unique sample ID; fourth column: mutation class (3 = VUS); fifth column: HGVS c. and p. nomenclature of BRCA1 pathogenic mutations co-occurring in each sample.

pathogenicity classifications between different submitters, instead showing all submitted classifications for each variant. Submitters are encouraged to provide classification evidence but are not required to. Consequently, the quality of data curation can vary significantly from one variant to the next. ClinVar has implemented a star rating system to provide an indication of annotation quality. The NIH-funded ClinGen curation effort, in collaboration with ClinVar [https://www.clinicalgenome.org/data-sharing/clinvar/] is encouraging the formation of expert panels to resolve discordances with the intention to create over time a higher quality subset of ClinVar content that will be more readily usable in the clinic.

GA4GH [Global Alliance for Genomics and Health 2016] has recently announced the creation of BRCAExchange [http://brcaexchange.org/], a consolidated database that integrates all publicly available datasets on BRCA gene variants. The quality assurance and funding models for BRCAExchange are still under development. BRCA Share, by contrast, builds on the curation model developed by the French National Working Group on BRCA VUS Classification, and funds curation efforts from membership dues. As with ClinVar, complete sharing of BRCA ShareTM data with BRCA Exchange would undermine the BRCA ShareTM model, but sharing of variant content, either directly or via the distributed Beacon [Krol 2015] mechanism, is under discussion.

A portion of the BRCA ShareTM funding is earmarked for functional studies to reduce the fraction of BRCA classifications in the variants of uncertain significance category. BRCA functional assays include assessments of E3-ligase activity, BARD1 binding and homology-directed repair [Starita et al., 2015], splicing [Théry et al., 2011], gene expression [Findlay et al., 2014], nuclear localization and P53 phosphorylation in response to DNA damage [Loke et al., 2015], growth restoration in BRCA1-deficient mouse embryonic stem cells [Bouwman et al., 2013], and others. Although there is not yet a consensus as to which of these will best correlate with in vivo phenotypes, the rapid recent progress provides hope that such a consensus may soon be forthcoming.

Options for which VUSes to characterize and when include:

- Per patient, in which a functional study is performed as part of or as a follow up to the BRCA test of a particular patient specimen

with the intent of improving the report for that patient [e.g., Loke et al. 2015]. This model is more appropriate for clinical labs than for a collection.

- Prioritized list, in which variants in a database are prioritized for functional studies according to some cost/benefit criteria.
- Exhaustive, in which all possible variants are generated via high-throughput mutagenesis and characterized in a high-throughput assay [Bouwman et al. 2013; Guidugli et al. 2013; Findlay et al. 2014; Starita et al. 2015; Sun et al. 2016].

An additional dimension for the design of functional studies is whether they require live patient cells or extracted DNA, or whether an in silico description of the variant will suffice. The high-throughput methods tend to require only in silico descriptions. Prioritized lists are compatible with both approaches but must be coupled to a cell, tissue or DNA bank, such as kConfab (http://www.kconfab.org/). No such collection currently exists for BRCA ShareTM.

We are piloting the prioritized list model using BRCA ShareTM data. One obvious prioritization criterion is frequency; given a fixed capacity for doing functional studies, choosing more frequent variants will impact more patients. Since most VUS are rare, the impact of a functional study for a single variant on the total per patient VUS rate will necessarily be minimal; many VUS must be classified to have an impact. Moreover, under current ACMG Guidelines [Richards et al., 2015] functional studies can provide “strong” but not “very strong” evidence, which is insufficient to reclassify a variant as likely pathogenic or benign without 1–2 additional pieces of moderate strength evidence, or >2 pieces of supporting evidence. Probabilistic approaches to evidence integration may provide stronger weighting to functional studies based on demonstrated predictive value, but there is not yet a consensus on how to do this [Grandval et al., 2013].

The BRCA ShareTM team has developed a prioritized list available from the “Statistics” section of the Website using the “Prioritization of VUS for validation” feature. It combines frequency (number of cases), available evidence, and co-occurrence data to produce a prioritization score, a high value means that additional evidence is available and that the addition of a functional study is likely to bring about reclassification by ACMG guidelines. In the future this may

be used to prioritize allocation of BRCA Share™ funding for functional studies. Funding of high-throughput studies or construction of reagents to support them is also under consideration.

Such efforts will ultimately provide high-quality databases for actionable genes for which patients might directly benefit from adequate variant annotation during high-throughput sequencing analysis especially in the context of WES and WGS. We believe that such databases will still be needed in the future as they provide accurate annotations as well as the list of available evidence to help the end-user interpret difficult variants. They will exist in parallel to general databases such as ExAC dedicated to variant frequency. BRCA1/2 BRCA Share™ databases could today be considered as models to demonstrate that data sharing is now a reality between private companies and the academic world; that resulting data are shared with research teams as well as commercial partners engaged in the long-term sustainability of such initiatives; that strong efforts are dedicated to reclassify VUS; and finally that such systems might facilitate NGS handling and secondary findings interpretation.

Finally, although BRCA Share™ is currently limited to BRCA1 and 2, there are other clinically important genes that could benefit from a similar alliance between researchers and commercial labs. A useful next step would be to include other important cancer risk genes into a Cancer Risk Share.

Disclosure statement: The authors declare no conflict of interest.

References

- Amendola LM, Dorschner MO, Robertson PD, Salama JS, Hart R, Shirts BH, Murray ML, Tokita MJ, Gallego CJ, Kim DS, Bennett JT, Crosslin DR, et al. 2015. Actionable exomic incidental findings in 6503 participants: challenges of variant classification. *Genome Res* 25:305–315.
- Amendola LM, Jarvik GP, Leo MC, McLaughlin HM, Akkari Y, Amaral MD, Berg JS, Biswas S, Bowling KM, Conlin LK, Cooper GM, Dorschner MO, et al. 2016. Performance of ACMG-AMP variant-interpretation guidelines among nine laboratories in the clinical sequencing exploratory research consortium. *Am J Hum Genet* 98:1067–1076.
- Antoniou A, Pharoah PDP, Narod S, Risch HA, Eyfjord JE, Hopper JL, Loman N, Olsson H, Johannsson O, Borg A, Pasini B, Radice P, et al. 2003. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 72:1117–1130.
- Bahcall OG. 2015. Genetic testing. ACMG guides on the interpretation of sequence variants. *Nat Rev Genet* 16:256–257.
- Balmaña J, Díez O, Rubio IT, Cardoso F, ESMO Guidelines Working Group. 2011.
- Bérout C, Hamroun D, Collod-Beroud G, Boileau C, Soussi T, Claustres M. 2005. UMD (Universal Mutation Database): 2005 update. *Hum Mutat* 26:184–191.
- Boughey JC, Hartmann LC, Anderson SS, Degnim AC, Vierkant RA, Reynolds CA, Frost MH, Pankratz VS. 2010. Evaluation of the Tyrer-Cuzick (International Breast Cancer Intervention Study) model for breast cancer risk prediction in women with atypical hyperplasia. *J Clin Oncol* 28:3591–3596.
- Bouwman P, van der Gulden H, van der Heijden I, Drost R, Klijn CN, Prasetyanti P, Pieterse M, Wientjens E, Seibler J, Hogervorst FBL, Jonkers J. 2013. A high-throughput functional complementation assay for classification of BRCA1 missense variants. 3:1142–1155.
- Cancer Institute NSW. 2015. EviQ Cancer Treatments; www.eviq.org.au.
- Caputo S, Benboudjema L, Sinilnikova O, Rouleau E, Bérout C, Lidereau R, French BRCA GGC Consortium. 2012. Description and analysis of genetic variants in French hereditary breast and ovarian cancer families recorded in the UMD-BRCA1/BRCA2 databases. *Nucleic Acids Res* 40:D992–D1002.
- Chay WY, Ong WS, Tan PH, Jie Leo NQ, Ho GH, Wong CS, Chia KS, Chow KY, Tan M, Ang P. 2012. Validation of the Gail model for predicting individual breast cancer risk in a prospective nationwide study of 28,104 Singapore women. *Breast Cancer Res* 14:R19.
- Check Hayden E. 2016. Funding for model-organism databases in trouble. *Nature*.
- Cherbal F, Salhi N, Bakour R, Adane S, Boualga K, Maillat P. 2012. BRCA1 and BRCA2 unclassified variants and missense polymorphisms in Algerian breast/ovarian cancer families. *Dis Markers* 32:343–353.
- Claustres M, Horaitis O, Vanevski M, Cotton RGH. 2002. Time for a unified system of mutation description and reporting: a review of locus-specific mutation databases. *Genome Res* 12:680–688.
- den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J, Roux A-F, Smith T, Antonarakis SE, Taschner PEM. 2016. HGVS recommendations for the description of sequence variants: 2016 Update. *Hum Mutat* 37:564–569.
- Dewey FE, Grove ME, Pan C, Goldstein BA, Bernstein JA, Chaib H, Merker JD, Goldfeder RL, Enns GM, David SP, Pakdaman N, Ormond KE, et al. 2014. Clinical interpretation and implications of whole-genome sequencing. *JAMA* 311:1035–1044.
- Dorschner MO, Amendola LM, Turner EH, Robertson PD, Shirts BH, Gallego CJ, Bennett RL, Jones KL, Tokita MJ, Bennett JT, Kim JH, Rosenthal EA, et al. 2013. Actionable, pathogenic incidental findings in 1,000 participants' exomes. *Am J Hum Genet* 93:631–640.
- Easton DF, Deffenbaugh AM, Pruss D, Frye C, Wenstrup RJ, Allen-Brady K, Tavtigian SV, Monteiro ANA, Iversen ES, Couch FJ, Goldgar DE. 2007. A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes. *Am J Hum Genet* 81:873–883.
- Eccles DM, Mitchell G, Monteiro ANA, Schmutzler R, Couch FJ, Spurdle AB, Gómez-García EB, ENIGMA Clinical Working Group. 2015. BRCA1 and BRCA2 genetic testing-pitfalls and recommendations for managing variants of uncertain clinical significance. *Ann Oncol* 26:2057–2065.
- Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JWW, Comber H, Forman D, Bray F. 2013. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* 49:1374–1403.
- Findlay GM, Boyle EA, Hause RJ, Klein JC, Shendure J. 2014. Saturation editing of genomic regions by multiplex homology-directed repair. *Nature* 513:120–123.
- Fischer C, Kuchenbäcker K, Engel C, Zachariae S, Rhiem K, Meindl A, Rahner N, Dikow N, Plendl H, Debatin I, Grimm T, Gadzicki D, et al. 2013. Evaluating the performance of the breast cancer genetic risk models BOADICEA, IBIS, BR-CAPRO and Claus for predicting BRCA1/2 mutation carrier probabilities: a study based on 7352 families from the German Hereditary Breast and Ovarian Cancer Consortium. *J Med Genet* 50:360–367.
- Fokkema IFAC, Taschner PEM, Schaafsma GCP, Celli J, Laros JFJ, Dunnen den JT. 2011. LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat* 32:557–563.
- Gadzicki D, Evans DG, Harris H, Julian-Reynier C, Nippert I, Schmidtke J, Tiben A, van Asperen CJ, Schlegelberger B. 2011. Genetic testing for familial/hereditary breast cancer-comparison of guidelines and recommendations from the UK, France, the Netherlands and Germany. *J Community Genet* 2:53–69.
- Global Alliance for Genomics and Health. 2016. GENOMICS. A federated ecosystem for sharing genomic, clinical data. *Science* 352:1278–1280.
- Goldgar DE, Easton DF, Byrnes GB, Spurdle AB, Iversen ES, Greenblatt MS, IARC Unclassified Genetic Variants Working Group. 2008. Genetic evidence and integration of various data sources for classifying uncertain variants into a single model. *Hum Mutat* 29:1265–1272.
- Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro ANA, Tavtigian SV, Couch FJ, Breast Cancer Information Core (BIC) Steering Committee. 2004. Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. *Am J Hum Genet* 75:535–544.
- Grandval P, Fabre AJ, Olschwang S. 2013. Design of a core classification process for DNA mismatch repair variations of a priori unknown functional significance. *Hum Mutat* 34:920–922.
- Graña B, Lastra E, Lloret G, Brunet J, Isla D. 2011. SEOM clinical guidelines for hereditary cancer. *Clinical and Translational Oncology* 13:580–586.
- Guidugli L, Carreira A, Caputo SM, Ehlen A, Galli A, Monteiro ANA, Neuhausen SL, Hansen TVO, Couch FJ, Vreeswijk MPG, on behalf of the ENIGMA consortium. 2013. Functional assays for analysis of variants of uncertain significance in BRCA2. *Hum Mutat* 35:151–164.
- Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, King MC. 1990. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 250:1684–1689.
- Holter S, Borgida A, Dodd A, Grant R, Semotiuk K, Hedley D, Dhani N, Narod S, Akbari M, Moore M, Gallinger S. 2015. Germline BRCA mutations in a large clinic-based cohort of patients with pancreatic adenocarcinoma. *J Clin Oncol* 33:3124–3129.
- Kaiser J. 2016. Biomedical resources. Funding for key data resources in jeopardy. *Science* 351:14.
- Krol A. 2015. Beacon project cracks the door for genomic data sharing - Bio-IT world.
- Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, Maglott DR. 2014. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* 42:D980–D985.
- Loke J, Pearlman A, Upadhyay K, Tesfa L, Shao Y, Ostrer H. 2015. Functional variant analyses (FVAs) predict pathogenicity in the BRCA1 DNA double-strand break repair pathway. *Hum Mol Genet* 24:3030–3037.
- Lovelock PK, Spurdle AB, Mok MTS, Farrugia DJ, Lakhani SR, Healey S, Arnold S, Buchanan D, kConFab Investigators, Couch FJ, Henderson BR, Goldgar DE, et al.

2007. Identification of BRCA1 missense substitutions that confer partial functional activity: potential moderate risk variants? *Breast Cancer Res* 9:R82.
- Maxwell KN, Hart SN, Vijai J, Schrader KA, Slavin TP, Thomas T, Wubbenhorst B, Ravichandran V, Moore RM, Hu C, Guidugli L, Wenz B, et al. 2016. Evaluation of ACMG-guideline-based variant classification of cancer susceptibility and non-cancer-associated genes in families affected by breast cancer. *Am J Hum Genet* 98:801–817.
- Mazzola E, Blackford A, Parmigiani G, Biswas S. 2015. Recent enhancements to the genetic risk prediction model BRCAPRO. *Cancer Inform* 14:147–157.
- Millot GA, Carvalho MA, Caputo SM, Vreeswijk MPG, Brown MA, Webb M, Rouleau E, Neuhausen SL, Hansen TVO, Galli A, Brandão RD, Blok MJ, et al. 2012. A guide for functional analysis of BRCA1 variants of uncertain significance. *Hum Mutat* 33:1526–1537.
- National Institute for Health Care, Excellence. 2013. Familial breast cancer classification and care of people at risk of familial breast cancer and management of breast cancer and related risks in people with a family history of breast cancer. nice.org.uk.
- Pharoah PDP, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BAJ. 2002. Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet* 31:33–36.
- Philippakis AA, Azzariti DR, Beltran S, Brookes AJ, Brownstein CA, Brudno M, Brunner HG, Buske OJ, Carey K, Doll C, Dumitriu S, Dyke SOM, et al. 2015. The Matchmaker Exchange: a platform for rare disease gene discovery. *Hum Mutat* 36:915–921.
- Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL, et al. 2015. 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. *Genet Med* 17:405–423.
- Salgado D, Desvignes J-P, Raï G, Blanchard A, Miltgen M, Pinard A, Lévy N, Collod-Beroud G, Bérout C. 2016. UMD-Predictor: a high throughput sequencing compliant system for pathogenicity prediction of any human cDNA substitution. *Hum Mutat* 37:439–446.
- Santos C, Peixoto A, Rocha P, Pinto P, Bizarro S, Pinheiro M, Pinto C, Henrique R, Teixeira MR. 2014. Pathogenicity evaluation of BRCA1 and BRCA2 unclassified variants identified in Portuguese breast/ovarian cancer families. *J Mol Diagn* 16:324–334.
- Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. 2001. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 29:308–311.
- Spurdle AB, Healey S, Devereau A, Hogervorst FBL, Monteiro ANA, Nathanson KL, Radice P, Stoppa-Lyonnet D, Tavtigian S, Wappenschmidt B, Couch FJ, Goldgar DE, et al. 2012. ENIGMA—evidence-based network for the interpretation of germline mutant alleles: an international initiative to evaluate risk and clinical significance associated with sequence variation in BRCA1 and BRCA2 genes. *Hum Mutat* 33:2–7.
- Starita LM, Young DL, Islam M, Kitzman JO, Gullingsrud J, Hause RJ, Fowler DM, Parvin JD, Shendure J, Fields S. 2015. Massively parallel functional analysis of BRCA1 RING domain variants. *Genetics* 200:413–422.
- Sun S, Yang F, Tan G, Costanzo M, Oughtred R, Hirschman J, Theesfeld CL, Bansal P, Sahni N, Yi S, Yu A, Tyagi T, et al. 2016. An extended set of yeast-based functional assays accurately identifies human disease mutations. *Genome Res* 26:670–680.
- Szabo C, Masiello A, Ryan JF, Brody LC. 2000. The breast cancer information core: database design, structure, and scope. *Hum. Mutat* 16:123–131.
- Théry JC, Krieger S, Gaildrat P, Révillion F, Buisine M-P, Killian A, Duponchel C, Rousselin A, Vaur D, Peyrat J-P, Berthet P, Frebourg T, et al. 2011. Contribution of bioinformatics predictions and functional splicing assays to the interpretation of unclassified variants of the BRCA genes. *Eur J Hum Genet* 19:1052–1058.
- Vail PJ, Morris B, van Kan A, Burdett BC, Moyes K, Theisen A, Kerr ID, Wenstrup RJ, Egginton JM. 2015. Comparison of locus-specific databases for BRCA1 and BRCA2 variants reveals disparity in variant classification within and among databases. *J Community Genet* 6:351–359.
- Wooster R, Neuhausen SL, Mangion J, Quirk Y, Ford D, Collins N, Nguyen K, Seal S, Tran T, Averill D. 1994. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science* 265:2088–2090.