Published in final edited form as:

Hum Mutat. 2015 February ; 36(2): 260–269. doi:10.1002/humu.22736.

Genome Annotation by Shotgun Inactivation of a Native Gene in Hemizygous Cells: Application to *BRCA2* with Implication of Hypomorphic Variants

Soma Ghosh, Anil K. Bhunia, Bogdan C. Paun, Samuel F. Gilbert, Urmil Dhru, Kalpesh Patel, and Scott E. Kern*

Department of Oncology, The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, Baltimore, Maryland 21287

Abstract

The greatest interpretive challenge of modern medicine may be to functionally annotate the vast variation of human genomes. Demonstrating a proposed approach, we created a library of *BRCA2* exon 27 shotgun-mutant plasmids including solitary and multiplex mutations to generate human knockin clones using homologous recombination. This 55-mutation, 13-clone syngeneic variance library (SyVaL) comprised severely affected clones having early-stop nonsense mutations, functionally hypomorphic clones having multiple missense mutations emphasizing the potential to identify and assess hypomorphic mutations in novel proteomic and epidemiologic studies, and neutral clones having multiple missense mutations. Efficient coverage of nonessential amino acids was provided by mutation multiplexing. Severe mutations were distinguished from hypomorphic or neutral changes by chemosensitivity assays (hypersensitivity to mitomycin C and acetaldehyde), by analysis of RAD51 focus formation, and by mitotic multipolarity. A multiplex unbiased approach of generating all-human SyVaLs in medically important genes, with random mutations in native genes, would provide databases of variants that could be functionally annotated without concerns arising from exogenous cDNA constructs or interspecies interactions, as a basis for subsequent proteomic domain mapping or clinical calibration if desired. Such gene-irrelevant approaches could be scaled up for multiple genes of clinical interest, providing distributable cellular libraries linked to public-shared functional databases.

Keywords

BRCA2; shotgun library; hypomorphs; homologous recombination

Introduction

Clinical gene sequencing is now common but clinical uses of sequence results is hindered by the limited functional annotation for the innumerable variants identified, even in well-

*Correspondence to: Scott E. Kern, Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, CRBI, Room No. #451, 1650 Orleans Street, Baltimore, Maryland 21287. sk@jhmi.edu.

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

Disclosure statement: The authors have no conflict of interest to declare.

studied genes. As one example, *BRCA2* (MIM #600185) sequencing to evaluate an individual's inherited risks for breast, ovarian, and pancreatic cancer often reveals innumerable variants of unknown significance (VUS), but the absence of a simple and facile functional assay for risk assessment of 2,000 unique VUS [Couch et al., 2014] (Breast Cancer Information Core, www.research.nhgri.nih.gov/bic/) has posed significant problems in counseling and in optimizing preventative care for thousands of carriers of these variants.

The functional annotation of human gene variation is the greatest interpretive challenge in modern genetic medicine. Functional annotation based on clinical observation studies is particularly difficult for rare variants and minority populations. Similar problems afflict functional annotation in hundreds of other clinically relevant genes relating to diverse diseases.

In the above example, several functional assays have been proposed to annotate the variants in *BRCA1* (MIM #113705) and *BRCA2*. Analysis of cosegregation of the VUS with families having cancer is a powerful approach based on the cancer phenotype [Goldgar et al., 2004]. Functional studies that measure the role of mutations on the known activity of a protein, in assays based on the functions of the protein, have been used with some success. In vitro assays to evaluate effects of *BRCA2* VUS on *BRCA2* function by transient transfection using plasmids based on the known activities of *BRCA2* like cellular survival and viability, homologous recombination repair, and genome instability have been used [Wu et al., 2005]. The results of these assays, however, can be discrepant and therefore not helpful for clinical decision making and patient counseling [Wu et al., 2005].

Complementation tests to evaluate functions of VUS in *BRCA1*- and *BRCA2*-deficient cells have been conducted [Carvalho et al., 2007]. A murine embryonic stem cell-based complementation functional assay to evaluate mutations in *BRCA2* was established [Kuznetsov et al., 2008], this method appearing robust in evaluating functions of *BRCA2* variants. Human *BRCA2* protein in mouse models, however, could in theory create nonphysiologic interactions.

We explore a gene-irrelevant general strategy of generating human SyVaL, Syngeneic Variance Library [Hucl et al., 2008], originally reported using a site-directed mutational approach in *BRCA2*. Here, we extend the SyVaL technology to shotgun mutagenic libraries and multiplex mutations to efficiently analyze multiple variants simultaneously. Intriguing properties of lethal genes, such as tetraploid rescue and polyploidy, were uncovered. The modularity of this technique would enable its application to study naturally occurring variants in any disease-associated gene. The results obtained by such approaches could provide useful information for genetic counseling and for personalized therapy for a gene of clinical interest.

As a demonstration of principle in a given gene, we generated random mutations in *BRCA2* exon 27 using a PCR misincorporation shotgun approach to create a human SyVaL and then functionally annotated the solitary or multiplex variants as functionally damaged, neutral, or hypomorphic. We chose exon 27 of *BRCA2* because it is functionally diverse. The extreme C-terminal is nonessential, but a proximal region is necessary for a fully functional

interaction between BRCA2 and Rad51 due to the presence of a Rad51 interaction domain [Donoho et al., 2003; Esashi et al., 2005; Davies and Pellegrini, 2007]. Two nuclear-localization signals (NLS) exist in exon 27 [Spain et al., 1999; Yano et al., 2000]. Exon 27 of BRCA2 is nearly 200 amino acids in length as compared with the large size of BRCA2 protein (3,418 amino acids), and therefore offered a technically feasible size to test and validate novel approaches.

Materials and Methods

Cell Lines and Cell Culture

DLD1 cells were obtained from the American Type Culture Collection and cultured in RPMI supplemented with 10% FBS, and antibiotics at 37°C, 5% CO₂. CAPAN1 cells were cultured in Iscove's modified Dulbecco's medium, supplemented with 10% FBS and antibiotics. *BRCA2*-null (B2-A10A3) and *BRCA2*-hemizygous (B2-A6) cell lines were taken from previous work [Hucl et al., 2008] and cultured under conditions similar to DLD1 cells.

Generation of NXA16 Mutant Library

The construct used by Hucl et al. (2008) to target *BRCA2* (Gen-Bank Accession: #U43746.1) exon 27 was slightly modified such that the NotI site at the end of the right homology arm (RHA) was replaced by XhoI (Fig. 1). The modified construct named NXA16 was linearized with SpeI and the RHA fragment containing *BRCA2* exon 27 was amplified with primers B2-I-F and B2-I-R (Supp. Table S1) using Phusion polymerase (New England Biolabs, Ipswich, MA) to obtain a high-quality template for Mn⁺⁺-mediated mutation-prone PCR (MPP) [Xu et al., 1999].

The amplified product was used subsequently for MPP using published protocols (extension time was changed to 8 min; annealing temperature was 60°C). Ten products were collected over the course of progressive mutagenesis and serially labeled 1 through 10. In brief, 4 µl aliquots were harvested from a 100-µl reaction mix after every four cycles. The reaction was replenished with fresh reagents and continued for another four cycles, and so on. The products were analyzed on agarose gels, gel-purified, and inserted back into linearized NXA16 such that an enriched library of *BRCA2* exon 27 mutant plasmids having solitary and multiplex mutations in the RHA was generated. Following sequencing of mutation-carrying plasmids, 72 plasmids from the library were selected (Supp. Table S2) to form a targeting pool of plasmids so as to yield a *BRCA2* exon 27 syngeneic variance library (SyVaL) following homologous recombination in *BRCA2*-hemizygous cells (B2-A6).

Composition of the Targeting Pool of Plasmids

By mass, 20% of total pooled transfectable DNA was taken from each of the following categories of plasmids: a nonsense codon as the first (the 5'-most) mutation (10 clones); a low number of missense mutations preceding a nonsense mutation (20 clones); a low number of missense mutations only (20 clones); and an intermediate number of missense mutations (20 clones). The remaining 20% comprised the positive (an early stop mutation clone) and negative (a silent or late truncating mutation clone) controls, 10% each in the total pooled DNA. Supp. Table S2 has a list of these plasmids and their sequence variations.

Generation of a SyVaL by Homologous Recombination in Hemizygous Cells

An enriched library of *BRCA2* exon 27 mutant plasmids, described above, was used as the targeting pool of transfectable DNA. This pool was cotransfected with pRC and pHelper plasmids into HEK 293 cells using Lipofectamine (Invitrogen, Grand Island, NY) [Kohli et al., 2004]. Two days later, virus was harvested from HEK 293 cells and used for infection and gene targeting of exon 27 in *BRCA2*-hemizygous cells (B2-A6), which had one disrupted allele from a prior targeted deletion of exon 11 [Hucl et al., 2008]. Infected cells were split into 96-well plates. Three weeks later, hygromycin-resistant clones having homologous integration of the targeting construct were identified by PCR using a primer inside the selectable marker (vector-specific primer) and a primer outside of the homology arms (SF/SR; Fig. 1). The allelic phase of the constructs was determined by PCR using primers PF/PR (Fig. 1), followed by sequencing. All protein variants are preceded by the symbol “p.”

Plasmid Purification, RNA and Genomic DNA Isolation, PCR, RT-PCR, and Sequencing

Plasmid DNA was purified (Qiagen, Valencia, CA). Genomic DNA (crude) was released from cancer cells using either the Lyse-N-Go PCR Reagent (Pierce Biotechnology, Rockford, IL) or by purification (Qiagen). Primers were obtained from IDT (Integrated DNA Technologies, Coralville, IA). For most PCR reactions, the touch-down method was used. RNA was isolated (RNeasy Mini kit; Qiagen). Complementary DNA was made from RNA (Superscript III; Invitrogen). PCR products were resolved on agarose gels using lithium boric acid electrophoresis (Faster Better Media LLC, Hunt Valley, MD) [Brody et al., 2004]. Automated sequencing was performed by MacroGen USA (Rockville, MD).

Cell Cycle Analysis

Cells (10^6) were treated with mitomycin C (MMC; Sigma, St. Louis, MO) for 48 hr, washed once in PBS, fixed in 70% ethanol at -20°C for at least 30 min, washed again with PBS, and incubated in 0.2 ml of PI solution (0.1% Triton X-100, $1\times$ PBS, 200 $\mu\text{g/ml}$ RNase A [Sigma], 20 $\mu\text{g/ml}$ propidium iodide [Sigma]) at 37°C for 1 hr. The cells were diluted with 1 ml PBS for flow cytometry (FACScalibur; BD Biosciences, San Jose, CA). CellQuest (BD Biosciences) software was used to interpret cell cycle profiles.

Cell Proliferation Assay

One-thousand cells (for wild type) and 2,000 cells (for experimental cultures and A10A3) were plated per well of a 96-well plate. The cells were allowed to adhere and after 24 hr were exposed to either MMC or acetaldehyde (Sigma). After 6 days, cells were washed, lysed in 40 μl 0.03% SDS, and 0.5% Picogreen (Molecular Probes, Grand Island, NY) was added. Fluorescence was measured and the relative cell numbers were calculated, defining the untreated samples as “1.” Three independent replication of the experiment were performed per drug, with each graphed data point reflecting the average readings from six wells in a given representative experiment.

Immunofluorescence

Cells were grown on eight-chamber slides, washed with PBS, fixed with either -20°C methanol (for multipolarity assay using mouse anti- α -tubulin clone DMA1 antibody; Santa Cruz Biotechnology, Santa Cruz, CA) or with 4% paraformaldehyde (for RAD51 assay using mouse anti-RAD51 antibody; Santa Cruz Biotechnology) for 10 min, incubated with 0.1% Triton X-100 in PBS for 10 min, and blocked with 1% goat serum, 0.05% NP40 in PBS for 10 min. Samples were incubated with primary antibody for 2 hr and with fluorophore-tagged secondary antibodies (Molecular Probes) for 1 hr in blocking solution, and mounted with ProLong Gold antifade containing DAPI (Life Technologies, Grand Island, NY).

Statistical Analyses

The mean (X), standard deviation (SD), and standard error of the mean (SEM) were computed for replicated experiments. The 90% confidence intervals (CIs) of the fold difference (X, pharmacogenetic window) [Gallmeier et al., 2006; Hucl et al., 2007] between parental and null cells, for three experimental replicates of the drug-sensitivity assays, were calculated using the formula $[x \pm 2.92 (SD/ \sqrt{3})]$.

Results

Generation of SyVaL by Homologous Recombination

Exon 27 of *BRCA2* was targeted in hemizygous cells B2-A6 [Hucl et al., 2008] using the enriched library of 72 *BRCA2* exon 27 mutant plasmids (Supp. Table S2) having solitary and multiplex mutations in the RHA. We screened 1,500 hygromycin-resistant clones after four separate rounds of infection in B2-A6 cells. Of the 1,500 clones, we identified 34 that had integration in the *BRCA2* locus, of which 21 had integrated into the previously disrupted allele; 13 integrated into the wild-type allele, giving rise to 11 unique knockin clones, as two of the clones were identified more than once in separate infections. Homologous integration of these 34 clones and the desired sequence variations were confirmed using PCR with primers SF and SR (primer positions in Fig. 1) followed by sequencing. The desired sequence variations in the transcripts were confirmed by RT-PCR using primers RF and RR followed by sequencing. Allelic phases of the constructs were determined by PCR using primers PF and PR (primer position in Fig. 1) followed by sequencing to confirm whether the vector had integrated in the previously disrupted allele (containing intron 25 polymorphisms G and T; Fig. 1) or in the wild-type allele (having intron 25 polymorphisms C and C; Fig. 1) (list of primers, provided in Supp. Table S1). *BRCA2* intron 25 polymorphisms G/C and T/C were established in our previous work [Hucl et al., 2008].

Of the 13 confirmed knockin clones, two had early stop (nonsense) mutations (B2-587 and B2-616, [p.L3230X]) (Fig. 2; Table 1). The targeting construct generating this mutation, B27-365 (Supp. Table S2), was the positive control plasmid in the targeting pool (10% of all plasmids). Most of the clones were multiple-missense clones (Fig. 2). Two multiple-missense clones (B2-40 p.[M3248L, N3264T, F3273S, and K3315X] and B2-246 p.[S3231L, C3233G, K3266N, and L3274X]) had nonsense mutations (p.K3315X in clone B2-40, and p.L3274X in clone B2-246). Homologous integration of 34 clones of 1,500 total

integrants reflected an efficiency of site-specific integration of 2.3%. Obtaining 11 unique knockin clones out of a library depth of 72 clones (15% representation) indicated diversity generally desired from library-based techniques. A greater integration efficiency could be expected by employing TALEN or CRISPR techniques not used for the current SyVaL construction ([Hockemeyer et al., 2011; Qiu et al., 2013; Feng et al., 2014], and S.G., S.K., unpublished results of separate experiments). The ratio, of integration among clones into the already-disrupted allele to that of the wild-type allele, was 1.6:1; most of the knockin clones, unlike knockout or null clones, had wild-type function and could yield viable cells upon integration into either of the two alleles.

Purification of Clones Having Persistent Wild-Type Alleles

Vector-specific primers (SF and SR; PF and PR, shown in Fig. 1) were used to amplify the knockin clones, followed by sequencing to confirm the desired sequence variants and the allelic phase of the clones, respectively. In instances when the same region of the knockin clone was amplified using human-specific primers HF and HR (Supp. Table S1), and when sequencing failed to confirm the expected strong mutation peaks, a mixed population (wild type and mutant) of cells could be identified for purification using a second round of limiting dilution to obtain pure mutant clones. The purified clones were confirmed using human- and vector-specific primers (HF/HR and PF/PR, respectively; Supp. Table S1) followed by sequencing. RT-PCR was done to confirm the sequence variants in the transcripts. We performed flow cytometry to determine the ploidy and cell cycle profiles of the SyVaL clones. While nearly all clones had a normal diploid cell cycle profile, clone B2–246, a multiple-missense clone with a stop p.(S3231L, C3233G, K3266N, L3274X), was tetraploid (Fig. 3A). Transcript analysis of this clone revealed wild-type sequence, indicating the dominance of wild-type alleles over the mutated allele, which was difficult to evaluate by transcript sequencing alone. This observation revealed spontaneous tetraploidy in culture, a phenomenon we noticed while developing *PALB2*-deficient cells by homologous recombination, wherein two of the four *PALB2*-deficient cells revealed tetraploidy upon culturing (Fig. 3B) ([Ghosh et al., 2014] and unpublished results). Unlike cells with chemically induced tetraploidy or polyploidy [Kosik et al., 2009], it was interesting to note that the tetraploidy of these new cells was not transient but was stable over successive passages.

Functional Validation of SyVaL Clones

Human FANC-null cells are known to have hypersensitivity to interstrand cross-linkers, such as MMC. We evaluated the function of the SyVaL clones by estimating cell survival after a 6-day exposure to graded doses of MMC [van der Heijden et al., 2004, 2005; Gallmeier et al., 2006; Hucl et al., 2008]. Based on their responses, the clones were classified as damaged, hypomorphic, or neutral in chemosensitivity assays (MMC and acetaldehyde) [Hucl et al., 2008]. Clones B2–587 and B2–616 (p.L3230X) had 10–17x hypersensitivity when exposed to MMC (Fig. 4A) and were classified as functionally damaged clones. DLD1 parental and *BRCA2*^{−/−} (B2-A10A3) cell lines were used as controls [Hucl et al., 2008]. In three replicate MMC treatment experiments, the 90% CIs of the pharmacogenetic window between the parental and the knockin clones were found to be 10.43–18.22 for B2–587 and 7.73–12.63 for B2–616. SyVaL clones B2–27, B2–72

(genotypically similar clones p.[V3240A, D3260E, K3267R, Q3299R, K3326Q, N3348S, Q3350L, S3354F, and K3383N], B2–40 p.[M3248L, N3264T, F3273S, and K3315X], and B2–173 p.[M3248K, C3253R, D3260E, S3275C, K3314E]) had intermediate responses when exposed to MMC (Fig. 4B), with pharmacogenetic windows ranging 2–5x (90% CI for B2–72 was 3.67–5.67 and for B2–27 was 1.86–2.54), and were classified as functionally hypomorphic clones. The hypomorphic clones had altered amino acids in the putative NLS sequence (Fig. 2). The remaining missense clones (B2–241, B2–245, B2–259, B2–222, B2–352, B2–169; amino acid alterations of clones in Table 1) and the tetraploid clone B2–246 (Fig. 4C) revealed no material differences in IC_{50} values between parental and knockin clones, and they were classified as neutral clones.

The null and hypomorphic knockin clones were also exposed to the ethanol metabolite, acetaldehyde, a chemical that may have had a significant role in the evolution of the FANC genes *BRCA2* and *PALB2* [Ghosh et al., 2014]. We observed a 4–8x pharmacogenetic window using the functionally null clones B2–587 and B2–616 p.L3230X (Fig. 5A) but nearly no differential response to acetaldehyde using the functionally hypomorphic clones B2–27 and B2–72 (p.[V3240A, D3260E, K3267R, Q3299R, K3326Q, N3348S, Q3350L, S3354F, and K3383N]) (Fig. 5B). The 90% CI for clones B2–587 and B2–616 (p.L3230X) were 7–8x and 3.3–6.7x, respectively.

Analysis of Multipolar Mitosis in SyVaL Clones

BRCA2 has a role in the formation of the midbody and in the completion of mitosis [Daniels et al., 2004; Mondal et al., 2012; Takaoka et al., 2014]. However, a study using a *BRCA2*-null cell line (B2-A10A3), engineered in our laboratory [Hucl et al., 2008], was reported to give contrasting results [Lekomtsev et al., 2010]. In order to reexamine this subject using our previous *BRCA2*-null (B2-A10A3) and present SyVaL clones, we labeled the cells with mouse anti- α -tubulin antibody and assayed for the presence of bipolar or multipolar mitosis. Knockin clones, functionally null for *BRCA2* (B2–587 and B2–616, [p.L3230X]), had numerous (3%–8%) multipolar cells, as did the engineered *BRCA2*-null (B2-A10A3) and the naturally derived *BRCA2*-null cell line CAPAN1 (Table 2; representative α -tubulin-labeled images are shown in Supp. Fig. S1). The knockin hypomorphic clones (B2–27 and B2–72) (p.[V3240A, D3260E, K3267R, Q3299R, K3326Q, N3348S, Q3350L, S3354F, and K3383N]) had 1%–3% multipolarity. The DLD1 parental cell line and the *BRCA2* hemizygous cell line (B2-A6) had no detectable multipolarity, consistent with having wild-type *BRCA2* function. Representative α -tubulin-labeling results are shown in Supp. Figure S1.

RAD51 Focus Formation in SyVaL Clones

Cells were labeled with anti-RAD51 antibody with or without MMC treatment and assayed for RAD51 focus formation. Control cells generating a positive assay result (DLD1 parental cells and B2-A6 hemizygous cells) and negative results (B2-A10A3 and CAPAN1) were included. RAD51 focus formation was significantly reduced and interpreted as negative assay results in B2–587 (p.L3230X), and B2–72 (p.[V3240A, D3260E, K3267R, Q3299R, K3326Q, N3348S, Q3350L, S3354F, and K3383N]) (functionally chemosensitivity damaged and chemosensitivity hypomorphic clones, respectively) (Table 3).

Discussion

Novel and Modular Approach of This Study

This study is the first unbiased gene-irrelevant general technical approach wherein we used a shotgun method for systematically generating random solitary and multiplex mutations in an exon of a clinically relevant gene (using a PCR misincorporation approach) to create a human SyVaL, followed by functional annotation of the variants. The variants generated by such a library technology form a distributable resource for research. We present a limited application of our technology to *BRCA2* as a demonstration of principle. Generating multiplex mutations, and annotating them functionally, enables efficient coverage of coding regions in which missense changes might be tolerated without any functional deficits. Ideally, such readouts or conclusions might be informative for patient counseling if applied to various genes.

One could also expect to now apply and improve the efficiency of the shotgun technology by (1) synthesizing a comprehensive diversity of variant clones using automated variant gene synthesis and cell culture robotics and (2) drive the efficiency of generating integrant clones by employing TALEN or CRISPR systems.

Desirable features of such a general gene-irrelevant technique include the ability to infer functional status even in essential genes, to detect and quantify hypomorphic dysfunction of a mutation, to provide a distributable resource for flexible and wide use, to link clones to expandable databases for functional and clinical annotation, to avoid nonphysiologic gene expression or interspecies interactions, to test native gene loci including splice sites, and to gain a numerical efficiency from multiplexing and from a diversity of mutational contexts. The proposed method meets these criteria.

Demonstration of Principle: Applying Shotgun SyVaL Technology to a Given Gene Region, *BRCA2* Exon 27

To demonstrate the principles desired, as given in the preceding paragraph, we explored variants introduced into *BRCA2* exon 27. Using this terminal region, we demonstrated lethal mutation, hypomorphic mutation, simplex and multiplex mutation, and annotations linking new functional SyVaL data with the prior domain/structural literature. The goal was to examine an adequate clone set to establish the suitability of shotgun SyVaL technology for these stated principles, without specific intention to contribute to the knowledge set of any given gene.

RAD51-interacting domain—*BRCA2* protein is presumed to function in homologous recombination through interactions with RAD51 (MIM #179617). Both exons 11 and 27 of *BRCA2* code for domains interacting with Rad51: exon 11 encodes eight BRC motifs [Carreira and Kowalczykowski, 2011], whereas exon 27 encodes a single, distinct interaction domain. Mice with a homozygous deletion of only the RAD51-interacting region encoded by exon 27 have shorter lifespan than that of control littermates, attributed to early onset of cancer and sepsis [Donoho et al., 2003]. Cells having homozygous deletion of the Rad51-interacting domain in exon 27 suffer reduced viability, hypersensitivity to DNA

interstrand cross-linkers such as MMC, and gross chromosomal instability. Therefore, recombinational repair is impaired by truncations affecting the Rad51 domain of exon 27 but not completely ablated. Our early-stop (nonsense) damaged mutants (B2–587 and B2–616 [p.L3230X]) had hypersensitivities to MMC and acetaldehyde but the pharmacogenetic windows were not as large as in *BRCA2*-knockout null clone B2-A10A3. In addition, the clones developed some resistance over successive passages, plausibly due to secondary functional compensation relying on the intact BRC repeats in exon 11.

Nuclear-localization signal—*BRCA2* is a nuclear protein, but subcellular localization of truncated *BRCA2* was shown to be cytoplasmic [Spain et al., 1999]. NLSs are located in the final 156 residues of the *BRCA2* protein (NLS1, 3263–69; NLS2, 3381–85) [Spain et al., 1999; Yano et al., 2000]. The endogenous-truncated *BRCA2* mutant in CAPAN1 cells (6174delT) was also found to be cytoplasmic [Spain et al., 1999], explaining how many *BRCA2* mutants might be nonfunctional due to their impaired ability to translocate to the nucleus.

SyVaL clones B2–27 and B2–72 (p.[V3240A, D3260E, K3267R, Q3299R, K3326Q, N3348S, Q3350L, S3354F, and K3383N]) have amino acid alterations in both NLS1 (p.K3267R; see Fig. 2) and NLS2 (p.K3383N; see Fig. 2) and are functionally classified as hypomorphic clones by our functional assays. Another clone classified as hypomorphic is B2–40 (p.[M3248L, N3264T, F3273S, K3315X]). This clone has an altered amino acid sequence in NLS1 (p.N3264T; see Fig. 2) and a stop thereafter (p.K3315X; see Fig. 2), plausibly rendering NLS2 nonfunctional in this clone. On the contrary, clone B2–259 (p.[C3222R, A3235T, R3237W, K3267R, A3270T, L3271W, Q3299R, E3330G, and S3336N]) shares the same variation as clone B2–27 and B2–72 (p.[V3240A, D3260E, K3267R, Q3299R, K3326Q, N3348S, Q3350L, S3354F, and K3383N]) at NLS1 (p.K3267R) but has an intact NLS2. Presumably, an intact NLS2 could be sufficient to translocate the *BRCA2* protein to the nucleus, thereby retaining the wild-type function in this clone, which has an altered NLS1 sequence. This study highlights the role of hypomorphic mutants of a medically relevant gene. It might be very useful to clinically identify such hypomorphs in the general population and study the resultant and presumably more subtle phenotypic alterations. Without such a pre-existing functional map provided by in vitro annotation, such studies might never be adequately informative.

Tetraploid Rescue

Our study reveals how spontaneous tetraploidy might convey to clones a survival advantage when essential genes become inactivated. We observed this phenomenon multiple times. Chemically induced tetraploidy or polyploidy aids malignant transformation [Kosik et al., 2009]. Human tumors with mutations in genes such as *BRCA2* often have gross chromosomal instability and karyotype variability [Grigorova et al., 2004] and exhibit alterations in chromosome number, have structurally aberrant chromosomes, and have impaired completion of cytokinesis [Daniels et al., 2004]. Our data on multipolar mitosis in *BRCA2*-deficient or *BRCA2*-null cells are consistent with these observations. Multipolar mitosis causes mis-segregation of chromosomes and gives rise to polysomy, therefore potentially contributing to increased cancer susceptibility in *BRCA2* mutation carriers. The

Author Manuscript

detrimental effects of spontaneous tetraploidy, as in our cultures, however, might be offset by a selective advantage that may temporarily perpetuate the normal phenotype of a gene exploring “lethal” mutations. Such phenomena could give individual clones a survival advantage and therefore help in clonal progression. Owing to a load of various lethal gene mutations in typical normal populations, we suggested that the random loss of heterozygosity could be a propellant for cells to explore tetraploid genotypes, an intermediate phase often suggested to generate aneuploidy in early tumorigenesis [Ganem et al., 2007; Davoli and de Lange, 2011; Ghosh et al., 2014].

Clonal Variation

Author Manuscript

Clonal variation is expected in any study of organisms. Libraries of variants offer the advantage of quantifying the magnitude of clonal variation, which will be essential to interpreting assay results and offering guidance for clinical decision making. Among other causes, clonal variation could arise due to unknown and fortuitous compensatory mechanisms that affect the chemical hypersensitivities of the clones. A twofold to threefold variation ((as seen in hypomorphic clones B2–27 and 72 (p.[V3240A, D3260E, K3267R, Q3299R, K3326Q, N3348S, Q3350L, S3354F, and K3383N]) and in null clones B2–587 and 616 (p.L3230X) was suggested to often represent trivial variance within the typical behavior profiles of drug-sensitive assays [Hucl et al, 2007]. In any shotgun approach, fortuitous complementation of one mutation by another or other forms of cellular compensation are possible. The compensatory mechanisms might affect the functional readout of the variants under study. The advantage of a SyVaL is the evaluation of such clonal variance by using parallel and related variant clones. By such an approach, even fortuitous complementation or interactions between individual mutations and each other could be recognized due to the testing of mutations in multiple sequence contexts.

Mapping of Optimal Distances for Homologous Recombination

Author Manuscript

Use of recombinant adeno-associated virus-mediated (rAAV) gene targeting was a major advance in human gene targeting, as rAAV-based vectors were found to be 25-fold more efficient than comparable plasmid vectors [Russell and Hirata, 1998; Topaloglu et al., 2005]. Typically, for generation of knockin alleles by homologous recombination, small genetic alterations can be engineered in the homology arms of a vector by site-directed mutagenesis [Rago et al., 2007]. During the integration process, the engineered alteration(s) can become uncoupled from the selectable marker gene, and therefore it is advisable to design the altered sequence in close proximity to the marker [Rago et al., 2007]. Here, we introduced random mutations along exon 27, from amino acid 3,217 until amino acid position 3,391 near the end of the protein at (i.e., ranging 174 amino acids or 522 bp). Of the 13 SyVaL clones obtained, nine had break points of integration within 119 amino acids or ~350 bp (Fig. 2), though the alterations in the remaining clones were yet farther away. The range of alterations introduced by homologous recombination was determined by sequencing each clone. We were therefore able to map optimal distances (~to 350 bp) for homologous recombination in AAV-based viral systems based on the break points of integration of various mutant clones.

This multiplex unbiased approach of generating SyVaL in a medically important gene such as BRCA2 with random mutations was technically demonstrated. The strategy could be scaled up or improved by incorporating advances like custom gene synthesis, use of CRISPRs or TALENs for gene snipping to aiding gene targeting, and automation to customize and characterize an indefinite number of knockin clones. The modular nature of this approach permits the study of variants in any disease-associated gene. Creation of functionally annotated databases for clinical interpretation and distributable syngeneic human cell libraries encompassing the full range of human variation in clinically relevant genes would be possible by (1) improving the efficiency of recovering the clinically observed variants, (2) statistically distinguishing the magnitudes of clonal variation in cell lines that happen to bear the same mutation-specific substitution, from the variation attributed to different mutations, and (3) application of functional assays that are calibrated against epidemiologically defined rates of susceptibility to the human diseases of interest.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

Contract grant sponsors: NIH grants CA62924 and CA123483; The Everett and Marjorie Kovler Professorship in Pancreas Cancer Research.

References

- Brody JR, Calhoun ES, Gallmeier E, Creavalle TD, Kern SE. Ultra-fast high-resolution agarose electrophoresis of DNA and RNA using low-molarity conductive media. *Biotechniques*. 2004; 37:598–600. 602. [PubMed: 15517972]
- Carreira A, Kowalczykowski SC. Two classes of BRC repeats in BRCA2 promote RAD51 nucleoprotein filament function by distinct mechanisms. *Proc Natl Acad Sci USA*. 2011; 108:10448–10453. [PubMed: 21670257]
- Carvalho MA, Couch FJ, Monteiro AN. Functional assays for BRCA1 and BRCA2. *Int J Biochem Cell Biol*. 2007; 39:298–310. [PubMed: 16978908]
- Couch FJ, Nathanson KL, Offit K. Two decades after BRCA: setting paradigms in personalized cancer care and prevention. *Science*. 2014; 343:1466–1470. [PubMed: 24675953]
- Daniels MJ, Wang Y, Lee M, Venkitaraman AR. Abnormal cytokinesis in cells deficient in the breast cancer susceptibility protein BRCA2. *Science*. 2004; 306:876–879. [PubMed: 15375219]
- Davies OR, Pellegrini L. Interaction with the BRCA2 C terminus protects RAD51-DNA filaments from disassembly by BRC repeats. *Nat Struct Mol Biol*. 2007; 14:475–483. [PubMed: 17515903]
- Davoli T, de Lange T. The causes and consequences of polyploidy in normal development and cancer. *Annu Rev Cell Dev Biol*. 2011; 27:585–610. [PubMed: 21801013]
- Donoho G, Brenneman MA, Cui TX, Donoviel D, Vogel H, Goodwin EH, Chen DJ, Hasty P. Deletion of Brca2 exon 27 causes hypersensitivity to DNA crosslinks, chromosomal instability, and reduced life span in mice. *Genes Chromosomes Cancer*. 2003; 36:317–331. [PubMed: 12619154]
- Esashi F, Christ N, Gannon J, Liu YL, Hunt T, Jasin M, West SC. CDK-dependent phosphorylation of BRCA2 as a regulatory mechanism for recombinational repair. *Nature*. 2005; 434(7033):598–604. [PubMed: 15800615]
- Feng Y, Zhang S, Huang X. A robust TALENs system for highly efficient mammalian genome editing. *Sci Rep*. 2014; 4:3632. [PubMed: 24407151]
- Gallmeier E, Calhoun ES, Rago C, Brody JR, Cunningham SC, Hucl T, Gorospe M, Kohli M, Lengauer C, Kern SE. Targeted disruption of FANCC and FANCG in human cancer provides a

- preclinical model for specific therapeutic options. *Gastroenterology*. 2006; 130:2145–2154. [PubMed: 16762635]
- Ganem NJ, Storchova Z, Pellman D. Tetraploidy, aneuploidy and cancer. *Curr Opin Genet Dev*. 2007; 17:157–162. [PubMed: 17324569]
- Ghosh S, Sur S, Yerram SR, Rago C, Bhunia AK, Hossain MZ, Paun BC, Ren YR, Iacobuzio-Donahue CA, Azad NA, Kern SE. Hypersensitivities for acetaldehyde and other agents among cancer cells null for clinically relevant Fanconi anemia genes. *Am J Pathol*. 2014; 184:260–270. [PubMed: 24200853]
- Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro AN, Tavtigian SV, Couch FJ. Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. *Am J Hum Genet*. 2004; 75:535–544. [PubMed: 15290653]
- Grigороva M, Staines JM, Ozdag H, Caldas C, Edwards PA. Possible causes of chromosome instability: comparison of chromosomal abnormalities in cancer cell lines with mutations in BRCA1, BRCA2, CHK2 and BUB1. *Cytogenet Genome Res*. 2004; 104:333–340. [PubMed: 15162061]
- Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol*. 2011; 29:731–734. [PubMed: 21738127]
- Hucl T, Gallmeier E, Kern SE. Distinguishing rational from irrational applications of pharmacogenetic synergies from the bench to clinical trials. *Cell Cycle*. 2007; 6:1336–1341. [PubMed: 17568192]
- Hucl T, Rago C, Gallmeier E, Brody JR, Gorospe M, Kern SE. A syngeneic variance library for functional annotation of human variation: application to BRCA2. *Cancer Res*. 2008; 68:5023–5030. [PubMed: 18593900]
- Kohli M, Rago C, Lengauer C, Kinzler KW, Vogelstein B. Facile methods for generating human somatic cell gene knockouts using recombinant adeno-associated viruses. *Nucleic Acids Res*. 2004; 32:e3. [PubMed: 14704360]
- Kosik A, Bekier ME, Katusin JD, Kaur H, Zhou X, Diakonova M, Chadee DN, Taylor WR. Investigating the role of Aurora kinases in RAS signaling. *J Cell Biochem*. 2009; 106:33–41. [PubMed: 19009561]
- Kuznetsov SG, Liu P, Sharan SK. Mouse embryonic stem cell-based functional assay to evaluate mutations in BRCA2. *Nat Med*. 2008; 14:875–881. [PubMed: 18607349]
- Lekomtsev S, Guizetti J, Pozniakovsky A, Gerlich DW, Petronczki M. Evidence that the tumor-suppressor protein BRCA2 does not regulate cytokinesis in human cells. *J Cell Sci*. 2010; 123:1395–1400. Pt 9. [PubMed: 20356927]
- Mondal G, Rowley M, Guidugli L, Wu J, Pankratz VS, Couch FJ. BRCA2 localization to the midbody by filamin A regulates cep55 signaling and completion of cytokinesis. *Dev Cell*. 2012; 23:137–152. [PubMed: 22771033]
- Qiu Z, Liu M, Chen Z, Shao Y, Pan H, Wei G, Yu C, Zhang L, Li X, Wang P, Fan HY, Du B, et al. High-efficiency and heritable gene targeting in mouse by transcription activator-like effector nucleases. *Nucleic Acids Res*. 2013; 41:e120. [PubMed: 23630316]
- Rago C, Vogelstein B, Bunz F. Genetic knockouts and knockins in human somatic cells. *Nat Protoc*. 2007; 2:2734–2746. [PubMed: 18007609]
- Russell DW, Hirata RK. Human gene targeting by viral vectors. *Nat Genet*. 1998; 18:325–330. [PubMed: 9537413]
- Spain BH, Larson CJ, Shihabuddin LS, Gage FH, Verma IM. Truncated BRCA2 is cytoplasmic: implications for cancer-linked mutations. *Proc Natl Acad Sci USA*. 1999; 96:13920–13925. [PubMed: 10570174]
- Takaoka M, Saito H, Takenaka K, Miki Y, Nakanishi A. BRCA2 phosphorylated by PLK1 moves to the midbody to regulate cytokinesis mediated by nonmuscle myosin IIC. *Cancer Res*. 2014; 74:1518–1528. [PubMed: 24448238]
- Topaloglu O, Hurley PJ, Yildirim O, Civin CI, Bunz F. Improved methods for the generation of human gene knockout and knockin cell lines. *Nucleic Acids Res*. 2005; 33:e158. [PubMed: 16214806]
- van der Heijden MS, Brody JR, Dezentje DA, Gallmeier E, Cunningham SC, Swartz MJ, DeMarzo AM, Offerhaus GJ, Isacoff WH, Hruban RH, Kern SE. In vivo therapeutic responses contingent on

- Fanconi anemia/BRCA2 status of the tumor. Clin Cancer Res. 2005; 11:7508–7515. [PubMed: 16243825]
- van der Heijden MS, Brody JR, Gallmeier E, Cunningham SC, Dezentje DA, Shen D, Hruban RH, Kern SE. Functional defects in the fanconi anemia pathway in pancreatic cancer cells. Am J Pathol. 2004; 165:651–657. [PubMed: 15277238]
- Wu K, Hinson SR, Ohashi A, Farrugia D, Wendt P, Tavtigian SV, Deffenbaugh A, Goldgar D, Couch FJ. Functional evaluation and cancer risk assessment of BRCA2 unclassified variants. Cancer Res. 2005; 65:417–426. [PubMed: 15695382]
- Xu H, Petersen EI, Petersen SB, el-Gewely MR. Random mutagenesis libraries: optimization and simplification by PCR. Biotechniques. 1999; 27:1102–1104. 1106, 1108. [PubMed: 10631483]
- Yano K, Morotomi K, Saito H, Kato M, Matsuo F, Miki Y. Nuclear localization signals of the BRCA2 protein. Biochem Biophys Res Commun. 2000; 270:171–175. [PubMed: 10733923]

Figure 1.

Structural schematic of knockin clones created upon *BRCA2* exon 27 targeting. *BRCA2* hemizygous (B2-A6) cells were infected with virus carrying the shotgun-mutated constructs, each harboring simplex or multiplex mutation in the RHA. Integration of the construct into the desired locus is determined by screening primers SF and SR. The polymorphisms G/C and T/C in intron 25 are indicated. Integration in either the already disrupted (G-T) allele or the wild-type (C-C) allele is determined by PCR using phase-determination primers PF and PR. The numbers in the boxes indicate exons; LHA and RHA, left and RHAs; asterisks, various randomly generated mutations in the RHA.

Figure 2.
Physical map of SyVaL clones following *BRCA2* exon 27 targeting by a shotgun method. The amino acid changes and their positions in exon 27 of the Brca2 protein are indicated. The NLSs in exon 27 are shown.

Figure 3.

Tetraploidy in knockout and knockin cells. **A:** Cell cycle profiles of diploid *BRCA2*-hemizygous cell line and of *BRCA2* multiplex missense tetraploid clone, B2-246. **B:** Cell cycle profiles of parental and *PALB2*^{-/-}1 ([Ghosh et al., 2014] and previously unpublished data) diploid cell lines are compared with tetraploid clones *PALB2*^{-/-}3 and *PALB2*^{-/-}4 clones.

Figure 4.

Functional evaluation of SyVaL clones using drug sensitivity to MMC. Cell populations, following treatment with MMC at the indicated concentrations, are compared with untreated cells. **A:** Pharmacogenetic windows (X) comparing parental cells (DLD1; negative control), *BRCA2* knockout (*BRCA2*^{-/-} B2-A10A3; positive control), and *BRCA2* early stop knockin clones (B2-587 and B2-616) are given. **B:** Multiplex missense clones found to exhibit a hypomorphic chemosensitivity phenotype (B2-27, B2-72, B2-40, and B2-173) are compared with parental cells and *BRCA2* knockout cells (*BRCA2*^{-/-}). **C:** Multiplex missense clones having a wild-type phenotype (B2-169, B2-222, B2-241, B2-245, B2-246, B2-259, and B2-352) are compared with parental cells and *BRCA2* knockout cells (*BRCA2*^{-/-} B2-A10A3). The pharmacogenetic windows are averaged from three independent experiments, whereas data points are displayed from a representative experiment. Error bars, SEM of six replicate wells in each single experiment.

Figure 5.

Functional evaluation of SyVaL clones using sensitivity to acetaldehyde. Cell populations following treatment with acetaldehyde at the indicated concentrations are compared with untreated cells. **A:** Parental cells (DLD1), and *BRCA2* knockout cells (*BRCA2*^{-/-} B2-A10A3), are compared with exon 27 early-stop codon clones (B2-587 and B2-616). **B:** Multiplex missense clones having a hypomorphic chemosensitivity phenotype (B2-27 and B2-72) are compared with parental cells and *BRCA2* knockout (*BRCA2*^{-/-} B2-A10A3) cells. The pharmacogenetic windows (X) are averaged from three independent experiments, whereas data points are displayed from a representative experiment. Error bars, SEM of six replicate wells in each single experiment.

Table 1

Amino Acid Alterations of SyVaL Clones

Clone (library plasmid ID ^a)	Amino acid alterations
B2-27 and B2-72 (B27-347)	V3240A, D3260E, K3267R, Q3299R, K3326Q, N3348S, Q3350L, S3354F, and K3383N
B2-40 ^b	M3248L, N3264T, F3273S, and K3315X
B2-169 (B27-246)	Q3227P
B2-173 (B27-326)	M3248K, C3253R, D3260E, S3275C, and K3314E
B2-222 (B27-292)	S3219P, K3236E, S3241T, D3260E, C3265G, L3277Q, and D3378E
B2-241 (B27-303)	G3255R and C3304R
B2-245 (B27-373)	M3217T, I3259Y, E3309V, F3328S, L3333P, T3357A, D3378N, L3380P, C3386G, and L3390M
B2-246 (B27-317)	S3231L, C3233G, K3266N, and L3274X
B2-259 (B27-340)	C3222R, A3235T, R3237W, K3267R, A3270T, L3271W, Q3299R, E3330G, and S3336N
B2-352 (B27-300)	N3221D, K3263M, P3282H, R3302T, C3304R, and T3306S
B2-587 and B2-616 (B27-365)	L3230X

^a Each human cell clone carried a single plasmid-derived sequence replacing wild-type exon 27 *BRCA2* native sequence. See Supp. Table S1.

^b The integrated plasmid sequence in clone B2-40 presumably was a minor contaminant of another plasmid among those plasmids combined to form the library, and it did not correspond to an intentionally selected sequence.

Table 2

Analysis of Multipolar Mitosis

Name	Genotype	Hypersensitivity to MMC	Multipolar mitosis (%)
DLD1	BRCA2 ^{+/+}	No	0
B2-A6	BRCA2 ^{+/-}	No	0
CAPAN1	BRCA2 ^{fs/-}	Yes (6–8X)	8%
B2-A10A3	BRCA2 ^{-/-}	Yes (16–20X)	4%
B2-587	BRCA2 ^{mut/-}	Yes (15–20X)	3%
B2-616	BRCA2 ^{mut/-}	Yes (10–12X)	4%
B2-72	BRCA2 ^{mut/-}	Yes (4–5X)	3%
B2-27	BRCA2 ^{mut/-}	Yes (2–3X)	1%
B2-246	BRCA2 ^{+/-}	No	0
PAL B 2 ^{-/-} 1	PALB2 ^{-/-}	Yes (16X)	4.5%

+, wild-type exons;

–, long deletions; fs, frameshift; mut, engineered codon change.

Table 3

Analysis of RAD51 Focus Formation

Clone	Genotype	Assay ^a		
		MMC	Acetaldehyde	Rad51 ^b
DLD1-parental	BRCA2 ^{+/+}	Neutral	Neutral	Positive (11 ± 1.3/65 ± 0.9)
B2-A10A3	BRCA2 ^{-/-}	Deleterious	Deleterious	Negative (7 ± 1.9/11 ± 1.4)
CAPAN1	BRCA2 ^{fs/-}	Deleterious	Deleterious	Negative (10 ± 2.3/24 ± 1.4)
B2-587	BRCA2 ^{mut/-}	Deleterious	Deleterious	Negative (13 ± 1.4/18 ± 1.8)
B2-72	BRCA2 ^{mut/-}	Hypomorphic	Hypomorphic	Negative (8 ± 0.9/11 ± 1.0)

^aInterpretations according to *Materials and Methods* section.

^bNumbers within parentheses represent average percent of MMC-untreated/-treated cells having more than five RAD51 foci ± SEM.