Published in final edited form as:

Nature. 2019 July ; 571(7766): 576–579. doi:10.1038/s41586-019-1382-1.

Tumor lineage shapes BRCA-mediated phenotypes

Philip Jonsson^{1,2,3}, Chaitanya Bandlamudi¹, Michael L. Cheng⁴, Preethi Srinivasan⁵, Shweta S. Chavan¹, Noah D. Friedman^{2,3}, Ezra Y. Rosen⁴, Allison L. Richards¹, Nancy Bouvier¹, S. Duygu Selcuklu¹, Craig M. Bielski^{1,2,3}, Wassim Abida⁴, Diana Mandelker⁵, Ozge Birsoy⁵, Liying Zhang⁵, Ahmet Zehir⁵, Mark T.A. Donoghue¹, José Baselga^{4,†}, Kenneth Offit⁴, Howard I. Scher⁴, Eileen M. O'Reilly⁴, Zsafia K. Stadler⁴, Nikolaus Schultz^{1,3}, Nicholas D. Socci¹, Agnes Viale¹, Marc Ladanyi⁵, Mark E. Robson⁴, David M. Hyman^{4,6}, Michael F. Berger^{1,5,6,*}, David B. Solit^{1,2,4,6,*}, Barry S. Taylor^{1,2,3,6,*}

¹Marie-Josée and Henry R. Kravis Center for Molecular Oncology, New York NY

²Human Oncology and Pathogenesis Program, New York NY

³Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York NY

⁴Department of Medicine, Memorial Sloan Kettering Cancer Center, New York NY

⁵Department of Pathology, Memorial Sloan Kettering Cancer Center, New York NY

⁶Weill Cornell Medical College, New York, NY

Abstract

BRCA1 and *BRCA2* mutations predispose to select cancers^{1–3}, and disease-specific screening and preventative strategies have reduced cancer mortality in affected patients^{4,5}. These classical tumors suppressor genes have tumorigenic effects associated with somatic biallelic inactivation, though haploinsufficiency may also promote tumor formation and progression^{6,7}. Moreover, *BRCA1/2*-

Reprints and permissions information is available at www.nature.com/reprints.

*Correspondence and requests for materials should be addressed to: M.F.B. (bergerm1@mskcc.org), D.B.S. (solitd@mskcc.org), or B.S.T. (taylorb@mskcc.org).

†Present address: AstraZeneca, Maryland, USA

Author Contributions

P.J., M.F.B., D.B.S., and B.S.T. conceived the study. C.B. and P.S. generated germline variant calling and pathogenicity assessments. P.J., C.B., P.S., S.S.C., N.D.F., A.L.R., C.M.B., A.Z., M.T.A.D., N.S., M.F.B., D.B.S., and B.S.T. designed and performed data analysis. D.M., O.B., L.Z., Z.K.S., K.O., and M.E.R. aided germline variant pathogenicity assessment. N.B., S.D.S., N.D.S., and A.V. assisted with exome re-sequencing. M.L.C., E.Y.R., N.B., S.D.S., W.A., J.B., K.O., H.I.S., E.M.O., Z.K.S., M.L., M.E.R., D.M.H., M.F.B., and D.B.S. assisted with prospective genomic and clinical data collection, sample annotation, and consent infrastructure. P.J. and B.S.T. wrote the manuscript with input from all authors.

Competing Interests

M.L.C. reports receiving travel/accommodation funding from Allergan, Sanofi-Aventis, Daiichi Sankyo. W.A. reports receiving honoraria from Caret, advisory board activities for Clovis Oncology, Janssen, and MORE Health, travel/accommodation expenses from Clovis Oncology and GlaxoSmithKline, and research funding from AstraZeneca, Zenith Epigenetics, Clovis Oncology, and GlaxoSmithKline. E.M.O. reports receiving consulting fees from BioLineRx, Targovax, Halozyme, Celgene, Cytomx, and Bayer and research funding support from Genentech, Roche, BMS, Halozyme, Celgene, MabVax Therapeutics, and ActaBiologica. D.M.H. reports receiving research funding from AstraZeneca, Puma Biotechnology, Loxo Oncology and personal fees from Atara Biotherapeutics, Chugai Pharma, Boehringer Ingelheim, AstraZeneca, Pfizer, Bayer, Debiopharm Group, and Genentech. M.F.B. reports receiving research funding from Illumina and advisory board activities for Roche. D.B.S. reports advisory board activities for Loxo Oncology, Pfizer, Illumina, Lilly Oncology, Vivideon, and Intezyme. All stated activities were outside of the work described herein. No other disclosures were noted.

mutant tumors often harbor homologous recombination deficiency^{8–13}, and consequently exhibit increased therapeutic sensitivity to platinum-containing therapy and poly [adenosine diphosphate (ADP)-ribose] polymerase (PARP) inhibitors^{14,15}. However, the phenotypic and therapeutic relevance of *BRCA1* and *BRCA2* mutations remains poorly defined in most cancer types. Here we show that in the 2.7% and 1.8% of advanced cancer patients with *BRCA1/2* germline pathogenic or somatic loss-of-function alterations, respectively, selective pressure for biallelic inactivation, zygosity-dependent phenotype penetrance, and PARP inhibitor sensitivity was only observed in tumor types associated with increased heritable cancer risk in *BRCA1/2* carriers (BRCA-associated cancer types). Conversely, in non-BRCA associated cancer types, most carriers of these *BRCA1/2* mutation types had evidence for tumor pathogenesis independent of mutant *BRCA1/2*. Overall, mutant BRCA is an indispensable founding event for some tumors, whereas in a significant proportion of others, it appears to be biologically neutral, a difference predominantly conditioned by tumor lineage, with implications for disease pathogenesis, screening, clinical trial design, and therapeutic decision making.

As mutant *BRCA1/2* emerges as an important therapeutic target, the current clinical development and continued off-label use of PARP inhibitors broadly assumes that germline and somatic *BRCA1/2* mutations are tumor-agnostic biomarkers, having similar biologic and therapeutic significance wherever they are identified. Moreover, as oncology shifts from limited germline screening in high-risk populations to broader panel or unbiased sequencing of matched tumor and normal DNA in unselected cancer patients, studies are revealing a higher than anticipated prevalence of pathogenic germline mutations in cancer predisposition genes, including *BRCA1* and *BRCA2*^{16,17}. Consequently, we sought to understand which common and rare cancers are phenotypically and therapeutically dependent on somatic and germline BRCA alterations.

We analyzed the germline blood and matched tumor tissue of 17,152 cancer patients diagnosed with one of 55 cancer types in whom prospective clinical sequencing of up to 468 cancer-associated genes was performed to guide treatment decisions for advanced and metastatic disease (Extended Data Fig. 1A, Supplementary Tables 1–2). We defined somatic loss-of-function (LoF) alterations in *BRCA1* and *BRCA2* and identified germline pathogenic and likely pathogenic variants (hereafter pathogenic *BRCA1/2*) utilizing a clinically validated variant discovery pipeline and a custom pathogenicity classifier trained on expert curation by medical geneticists (see Methods)^{18,19}. Hypermutated tumors were considered separately as a biologically distinct class (Extended Data Fig. 1B–D).

Overall, 2.7% of patients (n=462) harbored a germline pathogenic *BRCA1/2* allele (Fig. 1A), a higher than expected incidence that reflected, in part, the demography of the cohort (at least 18% of Ashkenazi Jewish ancestry; see Methods, Extended Data Fig. 1E–G and Supplementary Table 3). Of the remaining 16,690 germline wildtype (WT) patients, 919 had a somatic mutation in *BRCA1* or *BRCA2*, of which most were missense mutations of uncertain significance (59%). In total, 307 patients (1.8%) had a somatic presumed LoF mutation in *BRCA1/2* in a non-hypermutated cancer (Fig. 1A and Supplementary Table 4). Multi-exonic or whole-gene homozygous deletions represented 24% of all somatic

BRCA1/2 LoF alterations (n=74) and more often affected *BRCA2* (p-value=1.5e-9, Extended Data Fig. 1H).

Germline pathogenic or somatic LoF *BRCA1/2*-mutant patients (4.9% in total) had one of 38 cancer types of which 53% were breast, ovary, prostate, or pancreas cancers, types known to be associated with *BRCA1/2* germline carrier status (hereafter referred to as BRCA-associated cancer types, see Methods)²⁰. Strikingly, these were the only four cancer types that upon ancestry-adjusted association testing were significantly enriched among *BRCA1/2* germline carriers, so all other cancer types were considered non-BRCA-associated for subsequent analyses (Fig. 1B). Finally, *BRCA1/2* germline carriers were younger at first cancer diagnosis and had a higher proportion of multiple independent cancer diagnoses than did germline WT cancer patients (Extended Data Fig. 1I–J)²¹.

Both germline pathogenic and somatic LoF mutations were modestly more common in *BRCA2* than in *BRCA1* pan-cancer (p-values<10⁻⁷, Extended Data Fig. 1K), though considerable variability existed among individual cancer types, suggesting lineage-based dependencies perhaps related to the distinct roles of these two genes in tumor initiation and progression (Fig. 1C)^{3,22}. The germline and somatic origin of *BRCA1/2* alterations were approximately equal pan-cancer, though key cancer types favored one over the other. As expected, among all BRCA-associated cancers, pathogenic germline mutations were more common than somatic LoF mutations (5.8 versus 2.7%, respectively, p-value=9e-14, Fig. 1D), except in prostate cancer where they occurred with approximately equal frequency. By contrast, in the non-BRCA associated cancer uterine sarcoma, somatic *BRCA2* homozygous deletions accounted for nearly all *BRCA1/2* alterations.

To assess the dependence on BRCA dysfunction, we determined the selective pressure for WT BRCA loss in tumors of germline or somatic carriers by integrating purity- and ploidy-corrected allele-specific copy number with high-precision mutant allele fractions (see Methods, Extended Data Fig. 2A–C). To model neutral evolution, we determined the specificity of WT BRCA loss by comparing it to a cancer type-specific background distribution of LOH spanning either benign germline variants or somatic VUSs in *BRCA1/2* predicted to have no effect on fitness. As expected, LOH affected the WT and benign germline/somatic VUS mutant alleles with equal frequency (Extended Data Fig. 2D). By comparison, 86% of zygosity changes in germline pathogenic *BRCA1/2* carriers targeted loss of the remaining WT BRCA allele (p-value=4.4e-36), consistent with positive selective pressure for biallelic inactivation in these tumors.

Pan-cancer, 61% of *BRCA1/2* germline carriers had somatic LOH affecting the WT allele, which was significantly enriched over the rate of background LOH in tumors with benign variants (20%, p-value<10⁻¹⁰⁰). While LOH affecting somatic LoF *BRCA1/2* mutations was also enriched over background (28% versus 11%, p-value=3.8e-7), it was more than 2-fold less prevalent than for germline pathogenic variants (p-value=4.2e-15) and lower than the rate of LOH affecting oncogenic mutations in *TP53* (p-value=9e-17; Fig. 2A). This pattern was consistent across BRCA gene and specimen type (primary versus metastatic; Extended Data Fig. 2E–H). In hypermutated tumors with somatic LoF *BRCA1/2* mutations, LOH occurred in only 14% of cases. Enrichment of LOH in pathogenic germline carriers was

significantly higher in BRCA-associated cancer types than in cancer types with no strong prior association such as lung, bladder, melanoma, and colorectal cancers in which it was largely absent (Fig. 2B), a pattern evident for somatic mutations as well (75 versus 39%, p -value=7.5e-22; Fig. 2C). Gene- and origin-specific differences were also evident in breast cancer (Extended Data Fig. 2I). Prostate was the only cancer type in which biallelic loss in germline and somatic LoF carriers was approximately equal (~70%, Fig. 2C).

Given the different rates of LOH across tumor types, we postulated that not all tumors in germline carriers were dependent on BRCA dysfunction and that somatic chromosomal losses common in cancer genomes may lead to the loss of the pathogenic germline *BRCA1/2* allele if it were dispensable. Indeed, we found that 8% of *BRCA1/2* germline carriers lost their pathogenic germline allele somatically. While least common in BRCA-associated cancer types, this occurred significantly more frequently in lung cancers (5 versus 20%, respectively, p -value=6.4e-11, Fig. 2D), apparently driven in some cases by selective pressure for biallelic inactivation of a proximal tumor suppressor gene somatically mutated *in trans* (Extended Data Fig. 3). In sum, these data suggest that selection for loss of WT BRCA was tumor lineage-dependent, while varying both by the affected gene and the origin of the first BRCA hit, and that the pathogenesis of some cancers arising in *BRCA1/2* carriers is likely independent of mutant BRCA.

To explore how *BRCA1/2* mutations dictated phenotypic changes in affected cancers, we performed whole-exome sequencing of tumor/normal pairs from 815 patients (293 and 522 *BRCA1/2*-mutant and WT cases, respectively; Extended Data Fig. 4A–C, Supplementary Table 5). We inferred two orthogonal signatures of HRD, one each from somatic mutations and large-scale transitions in contiguous CNAs^{11,23}, that we combined into a single composite measure of HRD (see Methods, Extended Data Fig. 4D–E and Supplementary Table 6). *BRCA1/2*-mutant tumors of BRCA-associated cancer types had a greater degree of HRD than did mutant tumors of non-BRCA-associated cancers, and far more than HR-intact tumors (see Methods, Fig. 3A). Hypermutated tumors with somatic *BRCA1/2* mutations had no more HRD than WT cancers, which together with their low rate of biallelic inactivation, is consistent with these mutations being likely passenger events.

We next investigated how tumor lineage and *BRCA1/2* zygosity interact to determine the degree of HRD. Pan-cancer, heterozygous *BRCA1/2*-mutant tumors had only modestly greater HRD than WT cancers, which increased upon biallelic inactivation, independent of germline or somatic origin (Fig. 3B). BRCA-associated cancers had a greater degree of HRD at all levels of zygosity and was mutant dose-dependent, increasing with each successive BRCA hit. In contrast, only the small subset of non-BRCA-associated cancers with biallelic inactivation had any evidence of HRD (Fig. 3C and Extended Data Fig. 4F–H). Among germline *BRCA1/2* carriers, this lineage- and zygosity-dependent effect on the HRD phenotype was also evident within the same patient diagnosed with multiple independent primary tumors (Extended Data Fig. 5). Collective, these data suggest that different tumor lineages vary in their phenotypic susceptibility to a *BRCA1/2* defect, which plays a central role in mediating tumor development in some patients, whereas in others tumorigenesis is apparently BRCA-independent, both within and across patients.

Given the near complete lack of HRD in non-BRCA associated cancers with heterozygous *BRCA1/2* mutations, we hypothesized that those arising somatically were the consequence, rather than cause, of tumorigenesis in these patients. In 50% of evaluable cases independent of cancer type (n=21 of 42), which had a higher mutational burden overall, the nucleotide context of the somatic *BRCA1/2* mutation was consistent with the mutation being introduced by a dominant non-HRD mutational process (Fig. 3D and Extended Data Fig. 6–7). While this analysis cannot exclude BRCA haploinsufficiency mediating non-HRD driven tumorigenesis⁷, it implies that most somatic heterozygous *BRCA1/2* mutations in non-BRCA-associated cancers are neutral passenger mutations, which like those in hypermutated tumors, may be a consequence rather than the cause of tumorigenesis.

These results imply that a reliance on germline only testing to select patients of any cancer type for treatment with agents targeting DNA damage repair, such as PARP inhibitors, would likely result in the inclusion of a significant number of patients with non-BRCA driven tumors. To test this hypothesis, we curated treatment outcomes for patients treated with PARP inhibitors and/or immune checkpoint blockade (Supplementary Table 2). As expected, *BRCA1/2*-mutant patients with BRCA-associated cancer types derived greater clinical benefit from PARP inhibitor therapy than did patients that lacked these alterations (n=110 and 73, respectively; HR 0.58, 95% CI 0.46–0.73, log-rank p-value=3.7e-6; Fig. 4A, left). However, this was not true in patients with non-BRCA-associated cancer types (n=14 and 20, respectively; HR 1.02, 95% CI 0.6–1.7, p-value=0.98, Fig. 4A, right). Nevertheless, the longest clinical benefit from PARP inhibition among these tumors was a uterine sarcoma with a homozygous *BRCA2* deletion, which arise in 6.5% of all such tumors (Fig. 1D, Extended Data Fig. 1H), indicating that they may represent a previously unrecognized BRCA-dependent cancer type.

Among BRCA-associated cancer types, patients with somatic LoF *BRCA1/2* mutations had an improved response to PARP inhibition similar to germline carriers (n=26 and 84, respectively; log-rank p-value=2e-5, Fig. 4B) and the clinical benefit to PARP inhibition was also similar in heterozygous and biallelic patients (n=30 and 80, respectively; HR=0.44 and 0.47, p-values=6.7e-4 and 3.6e-05, respectively; Fig. 4C and Extended Data Fig. 8). Biallelic inactivation as determined by DNA analysis alone may, therefore, not be required for PARP inhibitor sensitivity in BRCA-associated cancers and even the modest HRD levels in the heterozygotes may be sufficient to confer sensitivity. While underpowered for formal analysis, none of the five patients with *BRCA1/2* heterozygous mutations in non-BRCA-associated cancer types in whom HRD was largely absent remained on PARP inhibitor therapy for more than 4 months. Collectively, these data suggest that lineage rather than the mutational origin, zygosity, or the extent of HRD may be the primary determinant of response in *BRCA1/2*-mutant patients.

Given the rate of *BRCA1/2* mutations increased with increasing mutational burden and these tumors generally lacked the phenotypic evidence of BRCA dependence, we postulated that despite suggested links^{24,25}, *BRCA1/2* alterations would not further discriminate patients likely to benefit from immune checkpoint blockade (ICB) beyond the influence of increased tumor mutational burden. Indeed, we found no association pan-cancer between LoF *BRCA1/2* alterations and event-free survival from the start of ICB therapy in 2,246 treated

patients (HR 0.99, 95% CI 0.78–1.3, p-value=0.9; Fig. 4D), even after adjusting for tumor mutational burden and affected cancer type or when considering the individual BRCA genes, mutation origin, or zygosity (Fig. 4E).

Our results indicate that mutant *BRCA1* and *BRCA2* have pleiotropic effects that are tumor lineage-dependent and that the majority of *BRCA1/2* alterations in non-BRCA associated cancers may be incidental findings unrelated to tumor pathogenesis and unlikely to be therapeutically actionable. This is consistent with studies in mice as well as clinical experience in germline *BRCA1/2* carriers, which demonstrate that these alterations strongly promote oncogenesis in some but not all tumor lineages. Further mechanistic study is required to determine the mechanisms that mediate differences in tolerance to HR defects across various tumor lineages. Despite the recent emergence of tissue-agnostic biomarkers of cancer therapy response such as NTRK fusions²⁶ and microsatellite instability²⁷, our data suggest that mutant *BRCA1/2* is unlikely to be of similar therapeutic relevance in all cancer types in which it is found. While the identification of germline carriers will continue to be important for broader cancer risk reduction, we caution that the integration of both germline and somatic *BRCA1/2* mutational status with broader somatic tumor profiles will ultimately be necessary to identify the subset of non-BRCA-associated cancers with true phenotypic dependence on mutant BRCA.

Methods

Study cohort and prospective sequencing

The study cohort was comprised of 18,392 tumor samples from 17,152 patients. All patients underwent prospective sequencing as part of their clinical care (February 2014 to July 2017). This study was approved by the Memorial Sloan Kettering Cancer Center Institutional Review Board and all patients provided written informed consent for tumor sequencing and review of medical records for detailed demographic, pathology, and treatment information. Genomic sequencing was performed on tumor DNA extracted from formalin-fixed paraffin-embedded (FFPE) tissue and normal DNA was sequenced in all patients. Patient samples were sequenced in a CLIA-certified environment with one of three versions of the FDA-authorized MSK-IMPACT targeted sequencing panel using methods and analysis as described previously^{19,28}. As the patients studied here were molecularly profiled to guide treatment decisions for advanced and metastatic disease, 42% of samples sequenced were obtained from metastatic tumors, while the remainder were primary tumor specimens predominantly acquired from patients with active metastatic disease.

Clinical and treatment data

Basic patient demographic data as well as treatment histories were retrieved from electronic health records in accordance with established IRB-approved processes. The cancer diagnosis history of each patient was matched against the cancer type of the prospectively sequenced tumor specimen to determine individual diagnosis date and corresponding age at diagnosis, as well as overall survival from disease onset. Diagnosis of pre-malignant disease was excluded from age of onset analyses. Treatment data were curated for the patients receiving one or more lines of either PARP inhibitor or immune checkpoint blockade therapy via free-

text search of medical and pharmacy records followed by expert review in a subset of cases, as follows. Overall, 217 patients received at least one line of PARP inhibitor treatment at the time of the clinical data freeze. For each of these, the duration of treatment was manually curated and for patients who had received more than one therapeutic regimen containing a PARP inhibitor, only the data from the first line of therapy was included in analyses. In total, 199 of these patients had at least one sequenced specimen prior to the start of therapy. Patients treated with one or more lines of immune checkpoint blockade therapy included those receiving either anti-PD-1/PD-L1 agents (n=1,771), anti-CTLA-4 agents (n=119), or their combination (n=356). All clinical data were frozen in March 2018 and anonymized where necessary in a manner identical to molecular data and integrated post-anonymization for comprehensive analysis (see below).

Germline variant discovery and pathogenicity assessment

At the time of clinical data freeze, 3,358 patients in the cohort had consented to identified analyses of germline variants via an IRB protocol (#12-245, part C;). For the remainder of patients, their genomic and clinical data were anonymized prior to analysis with a deterministic one-way hash function. In these patients, germline variant calling was performed using the clinically validated pipeline utilized for the identified analysis described above¹⁸. Those germline variants that were suspected to have resulted from clonal hematopoiesis (CH) or circulating cell-free DNA from the tumor were excluded as described in Srinivasan P, et al. (*manuscript in preparation*). Briefly, excluded variants were those with insufficient aligned read coverage (<100- or 50-fold in the blood normal or tumor, respectively) unless spanned by a somatic homozygous deletion, or those occurring in a CH-associated gene²⁹ with a variant allele frequency (VAF) in the blood normal <0.35 or 0.25 for SNVs and indels, respectively, or a somatic VAF <0.25. Germline variants were retained for further analysis if their observed VAF in the tumor specimen was within the 95% confidence interval of the expected VAF for the variant inferred using the tumor sample purity and total and minor copy number estimated from FACETS analysis (see below) except in cases for which FACETS was indeterminant, after which we used VAF thresholds of 0.4 and 0.3 in the normal and tumor specimens respectively. Finally, known oncogenic variants of low VAF in normal blood specimens and with a ratio of tumor-to-normal sample VAFs greater than three were presumed to have originated in tumor-derived circulating cell-free DNA (cfDNA) and excluded. The remaining germline variants were annotated with Variant Effect Predictor (VEP, v88) using vcf2maf (<http://github.com/mskcc/vcf2maf>, v1.6.10) and subsequently filtered to include only rare variants by excluding common variants with a minor allele frequency above 2% using population frequencies from gnomAD³⁰ as well as ClinVar and myvariant.info (accessed as of September 2017).

Variant classification as pathogenic or likely pathogenic (P/LP) was determined using a random forest classifier trained on expert curation by medical geneticists. Briefly, the classifier used a feature matrix comprised of multiple independently and partly overlapping feature categories including variant and gene type (oncogene versus tumor suppressor gene), sequence context (distance to splice site and C-terminal end), known pathogenicity, population frequency, *in silico* functional prediction, experimental validation annotation, and structured literature curation. Classifier training was performed on a set of 6,120 unique

variants for which annotated pathogenicity exists (473 pathogenic) across 88 genes from 6,009 patients using the ACMG guidelines for clinical interpretation. Overall, we used 450 trees and accounted for class imbalance using a ratio of weights of 1:4.5 between the benign and pathogenic classes. Ten-fold cross-validation achieved 94% [$\pm 6\%$ confidence interval (CI)] and 89% ($\pm 7\%$ CI) average precision and recall, respectively (Srinivasan P, et al. *manuscript in preparation*). Of a total of 60,697 unique rare variants, 3,158 were predicted as pathogenic across all genes in the MSK-IMPACT panel. Four additional *BRCA1* and *BRCA2* germline variants initially classified as benign were reclassified pathogenic based on sufficient level of evidence from the ENIGMA consortium³¹. Subsequent reclassification of a subset of variants presumed benign was performed, including C-terminal truncating variants beyond enzymatic domains. In total, 456 pathogenic or likely pathogenic *BRCA1/2* germline variants were identified (208 unique SNVs and indels) as were 6,764 rare variants considered either benign or having unknown significance. Notably, the pathogenicity of only two variants was discordant between this classification of *BRCA1/2* germline variant pathogenicity and a recent high-throughput functional validation study³². These variants (*BRCA1* C24Y and E85K) were classified as LoF by experimental methods but not predicted pathogenic by our classification. Of note, neither tumor harboring these two variants had evidence of HRD typical of known pathogenic germline *BRCA1/2* variants.

Germline *BRCA1* and *BRCA2* copy number variants were detected using a custom algorithm comparing coverage in targeted regions in a normal sample against a reference pool of blood normal samples²⁸. After GC-bias correction, genic and intragenic gains and losses at gene and individual exon level were called at fold-change thresholds 1.2 and -1.5 , respectively. Only contiguous events spanning at least two exons and with a q-value less than or equal to 0.01 were retained (Benjamini and Hochberg, Z-test). Germline deletions not present in the patient tumors (comparing tumor sample coverage against the normal pool) were excluded.

Ancestry estimation and ancestry-adjusted association testing

Ancestry estimation was performed for all 17,152 individuals using principal-component analysis (PCA) of SNPs covered by the MSK-IMPACT assay design. Subpopulations were identified using self-reported race for the patients consented for identified germline testing. Ashkenazi Jewish (ASJ) ancestry was identified within the predicted Caucasian subpopulation as those individuals carrying any of the 60 alleles enriched in the gnomAD ASJ subpopulation compared to the non-Finnish European (NFE) subpopulation. To test for associations between mutant *BRCA1* and *BRCA2* and cancer types, we developed an ancestry-controlled permutation-based framework (Srinivasan P, et al. *manuscript in preparation*). Briefly, for every cancer type of sufficient sample size (50 or more), we tested for the enrichment of *BRCA1* or *BRCA2* germline mutations relative to a background distribution of frequencies generated from all other tumor types while maintaining an underlying population structure in the null distribution consistent with that of the tested cancer type. Significant associations were those with a false discovery-corrected (Benjamini and Hochberg) q-value < 0.15 .

Somatic mutational analyses

Somatic mutations (substitutions and small insertions and deletions), gene-level focal CNAs, and structural rearrangements were detected with a clinically validated pipeline as previously described^{19,28}. Somatic alterations were classified as oncogenic or likely oncogenic using OncoKB³³. Any somatic mutation not otherwise classified known or likely oncogenic were considered variants of uncertain significance (VUS). Loss-of-function (LoF) somatic mutations in *BRCA1* and *BRCA2* included those annotated as oncogenic by OncoKB via literature review and curation or otherwise truncating mutations of any type (nonsense or frameshift indel). All focal *BRCA1* and *BRCA2* homozygous deletions were also considered oncogenic. Tumors were classified as hypermutated and thus analyzed separately if they had LoF mutations due to one of three sources of somatic hypermutation in affected samples: microsatellite instability, DNA polymerase epsilon-mediated ultra-mutation, or alkylating chemotherapy-induced hypermutation. Microsatellite instability (MSI) was determined for all tumor samples using MSIsensor as previously clinically validated^{34,35}. Additional MMR-mediated hypermutated tumors were those for whom univariate k-means clustering established an elevated tumor mutational burden compared to other tumors of the given cancer type of which 50% or greater of all somatic mutations were attributed to MMR/MSI mutational signature as indicated from signature decomposition analysis (see below). POLE ultra-mutated tumors were those with a mutation in the *POLE* exonuclease domain mutation (amino acid residues 86-427, transcript ENST00000320574), accompanied by hypermutation as determined by univariate k-means clustering of tumor mutational burdens by cancer type as described above, with at least 50% of the mutations attributable to the POLE-associated mutational signature. Tumors with alkylating therapy-induced hypermutation were classified similarly and required 50% or more of somatic mutations attributable to the mutational signature associated with exposure to temozolomide³⁶. Hypermutated tumors were considered as a biologically distinct class for analyses of somatic correlates of *BRCA1/2* status unless otherwise noted. The systems for annotation of relevant pathogenic germline (see above) versus somatic presumed LoF mutations in *BRCA1* and *BRCA2* are distinct to reflect the presence of distinct endogenous mutational processes or exogenous mutagens, selective pressures, clonal outgrowths, fitness gains and losses, and co-mutational patterns that accompany the acquisition of driver mutations in somatic cells but not inherited or *de novo* germline variants.

Allele-specific copy number, zygosity, and clonality inference

We determined total, allele-specific, and integer DNA copy number genome-wide as well as tumor purity and ploidy using FACETS (v.0.5.6, <http://github.com/mskcc/facets>)³⁷. Each tumor and matched normal specimen was processed in a two-pass manner, an initial run for purity and ploidy estimation followed by a second run for focal event detection (Srinivasan P, et al. *manuscript in preparation*). A subset of tumor-normal pairs was manually reviewed for fit accuracy. Tumor samples with a concordance less than 60% between observed and expected mutant allele frequencies for heterozygous SNPs in the tumor in regions of total copy number less than or equal to three were excluded from zygosity analyses. For tumors for which FACETS was unable to estimate a tumor purity directly, this value was estimated based on mutant allele fraction of somatic mutations in balanced diploid regions. A total of

15,195 patients had at least one tumor sample that qualified for zygosity analysis based on these criteria.

Zygosity of germline and somatic variants

The tumor-specific zygosity of both germline pathogenic variants and somatic mutations was assessed by integrating the read support for the mutant allele with total coverage and the estimated locus-specific total and allele-specific copy number (determined as described above). We determined if the observed variant allele frequency (VAF) for clonal events in the tumor was consistent with the expected VAF given the tumor purity and local copy number defined for germline variants as $(\Phi \times n + (1 - \Phi) / (\Phi \times N + 2 \times (1 - \Phi)))$ and as $(\Phi \times n) / (\Phi \times N + 2 \times (1 - \Phi))$ for somatic variants, where the tumor purity is Φ whereas N and n are the locus-specific copy number and allele-specific copy number, respectively. For variants with allelic imbalance in favor of the mutant allele, loss of heterozygosity (LOH) was considered present if somatic VAF was within the 95% binomial CI of the expected VAF of the lesser allele having a copy number of zero. LOH favoring the mutant allele was those variants for which the observed VAF exceeded the lower bound of the 95% CI of the expected VAF, and conversely, loss of the mutant (favoring the reference allele) was the reverse. For somatic LoF *BRCA1/2* mutations, zygosity inference was limited to clonal mutations (see below). Finally, the zygosity of a given variant was considered indeterminate if no tumor VAF was estimated, the variant was homozygous in the normal, or the tumor depth at the variant site was less than 50.

While specimens of low tumor cell content will limit the sensitivity of LOH inference based on somatic VAFs, we previously estimated that low tumor cell content affects sensitivity of clinical sequencing with MSK-IMPACT in less than 8.5% of all tumors profiled¹⁹. Moreover, we routinely called LOH in *BRCA1/2*-mutant tumor specimens with tumor cell content <20–30% (Extended Data Fig. 2A–B). Indeed, while tumor cell content will impact the somatic VAF of germline variants, this had no impact on the sensitivity for LOH detection, while only those somatic mutations detected in samples of tumor cell content less than ~20% had lower rates of LOH detection (Extended Data Fig. 2C). These specimens, however, represent only 4.3% of all somatic *BRCA1/2*-mutant cases in the study cohort. There was, therefore, no statistically significant difference in the rate of biallelic inactivation of somatic LoF *BRCA1/2* mutations when limiting the analysis to only those samples of tumor cell content >30% nor is tumor purity a statistically significant predictor of LOH for somatic mutations in such cases. Nevertheless, we cannot formally exclude the possibility that some heterozygous mutant tumors were actually biallelic and false negatives from the zygosity analysis. To determine whether, in the absence of somatic LOH, biallelic inactivation was achieved in *BRCA1/2* mutant tumors via monoallelic promoter methylation, we analyzed data generated by The Cancer Genome Atlas. Here, epigenetic silencing was restricted to *BRCA1*, arose primarily in breast and ovarian cancers, and was mutually exclusive with *BRCA1* mutations (Extended Data Fig. 4F). This is, therefore, unlikely to fully explain the modest HRD phenotype in heterozygous mutant tumors (see below and Fig. 3).

Enrichment of LOH above background

To assess the selective pressure for LOH leading to somatic WT allele loss in carriers of germline pathogenic *BRCA1* or *BRCA2* variants, we compared the rate of such LOH with a background distribution of the same zygosity changes spanning benign variants in the two genes. For the purposes of this analysis, we considered benign variation in *BRCA1* and *BRCA2* to consist of those variants whose gnomAD population allele frequency was less than 5% and have been annotated in ClinVar as “benign” with two or more gold stars indicating that at least two submitters have asserted non-pathogenic status. The zygosity of each such variant was determined in the corresponding tumor specimens of all affected patients using the same method for pathogenic variants described above. To verify benign germline *BRCA1/2* variants were not under selective pressure for somatic biallelic inactivation, we confirmed that the rate of such changes were consistent with the rate of LOH genome-wide in germline WT patients by cancer type and that zygosity changes in tumors of patients with these benign variants selected for loss of the WT allele and loss of the mutant allele with equal frequency, which together is consistent with these benign variants having no effect on fitness (Extended Data Fig. 2D). Having modeled neutral selection with this background distribution, we proceeded to test for the enrichment for biallelic inactivation (loss of WT) for carriers of pathogenic germline *BRCA1* or *BRCA2* variants by comparing their observed distribution of zygosity changes with those of the background using the Fisher’s exact test both pan-cancer and in individual cancer types (Srinivasan P, et al. *manuscript in preparation*). We constructed a similar background distribution of zygosity changes spanning somatic *BRCA1/2* variants of uncertain significance (VUSs, all variants that are not designated as pathogenic here) and tested for the enrichment of biallelic inactivation for somatic LoF *BRCA1/2* mutations in the same manner. Multiple hypothesis correction was performed with the using Bonferroni method.

Estimating the clonality of somatic mutations

For all somatic *BRCA1* and *BRCA2* mutations, we estimated clonality in each affected tumor specimen (fraction of tumor cells harboring the indicated mutation) as described previously^{38,39}. Briefly, we inferred the cancer cell fraction (CCF) for all mutations using the mutant allele fraction, locus-specific read coverage, and an analytical estimate of tumor purity using a binomial distribution and maximum likelihood estimation to generate posterior probabilities. A CCF is generated for the possibility that the mutant allele existed on the major copy number (as used in the zygosity estimate) and for the expected number of copies of the mutant allele (n_E), as determined by $\mu = (\text{VAF}/\Phi) \times (N \times \Phi + 2 \times (1 - \Phi))$, where $n_E = 1$ for $\mu < 1$ and is otherwise rounded to closest integer. For the purpose of zygosity calling and clonality analyses, a somatic mutation was considered subclonal if the upper bound of the 95% CI of its CCF was less than 0.8.

Exome re-sequencing

To increase our sensitivity for broader somatic mutational and DNA copy number alteration signature detection beyond that which is achievable by targeted sequencing of known cancer genes, we performed exome sequencing of 148 germline and 145 somatic *BRCA1/2*-mutated tumors, respectively, as well as 522 *BRCA1/2*-wildtype tumors. While most were

recaptures of existing sequencing libraries from clinical MSK-IMPACT sequencing, a subset were generated from remaining genomic DNA that was first quantified with PicoGreen and quality-controlled by Agilent BioAnalyzer. In total, 15.8–500 ng of DNA were used to prepare libraries using the KAPA Hyper Prep Kit (Kapa Biosystems KK8504) with 8–10 cycles of PCR. For existing libraries that corresponded to clinical MSK-IMPACT sequencing, 74–500 ng of remaining barcoded library were captured by hybridization using either the SureSelectXT Human All Exon V4 (Agilent catalog #5190-4632) or xGen Exome Research Panel v1.0 (IDT) according to manufacturer's protocol. PCR amplification of the post-capture libraries was carried out for 8 cycles. Samples were run on either HiSeq 4000 or HiSeq 2500 in rapid mode in a 100bp or 125bp paired-end run using the HiSeq 3000/4000 SBS Kit or HiSeq Rapid SBS Kit v2 (Illumina) or on NovaSeq 6000 in a 100bp paired-end run using the NovaSeq 6000 SBS v1 Kit and an S2 flow cell (Illumina).

Briefly, demultiplexed FASTQ files were trimmed of adaptors and short reads with TrimGalore (v0.2.5mod) and aligned to the b37 assembly of the human reference genome with BWA mem v0.7.5a. Read groups were annotated and PCR duplicates were marked with Picard Tools v2.9. Indel realignment was performed using the Assembly Based ReAligner (ABRA) v2.12⁴⁰ and base quality recalibration was performed with GATK v3.3-0⁴¹. Somatic mutations (point mutations and small insertions and deletions) were identified in tumor-normal pairs using MuTect v1.1.4⁴² and Vardict v1.5.1⁴³. Variants were then annotated as described above. Somatic mutation filtering was performed as follows. Initially, variants were whitelisted for retention if they were a known recurrent hotspot mutation (<http://www.cancerhotspots.org>)^{44,45} or a known non-truncating oncogenic variant per OncoKB literature curation³³. All non-whitelisted variants were excluded if they correspond to blacklisted sites that include those occurring: 1) in RepeatMasker-annotated low-complexity or simple-repeat regions or blacklisted regions from the ENCODE consortium (hgdownload.cse.ucsc.edu/goldenPath/hg19/database/rmsk.txt.gz and genome.ucsc.edu/cgi-bin/hgFileUi?db=hg19&g=wgEncodeMapability, respectively); 2) more than 10 times in any subpopulation of a non-TCGA subset of ExAC; 3) in more than four samples in a panel of 546 normal blood samples; or 4) with insufficient frequency or read support in the affected sample (VAF < 0.05, supported by fewer than 3 variant reads in the tumor or greater than 3 reads in the matched normal, tumor and matched normal coverage less than 20 and 10 respectively). Any mutation flagged by any of these filters in three or more independent tumor samples were also excluded from all samples. Additional criteria to exclude likely false positive complex indels or variants identified by VarDict were if the product of the VAF and supporting read depth in the tumor is less than six and the mean mapping quality (MQ) < 45, or if the MQ < 55 with a mean number of mismatches per supporting read (NM) greater than one or if the MQ < 60 where the NM > 2. Other excluded VarDict variants included those: 1) with VAF < 0.2 with MQ < 55 and p-value (SSF) > 0.05; 2) with MQ < 60 and SSF > 0.01; 3) spanning repetitive regions of length greater than 10; or 4) were 1bp repeats occurring at a sequence flanked by at least two repeats of the same base. FACETS analysis was performed on the exome data in a two-pass manner identical to that described for MSK-IMPACT sequencing using cval thresholds of 300 and 100 respectively. All DNA copy number profiles were manually reviewed for the presence of hyper-segmentation or other artifacts and re-run with altered parameters where necessary. Purity and ploidy

estimates were obtained and clonality inference completed in a manner identical to that described for MSK-IMPACT data.

Mutational signature inference

Mutation signatures were inferred from single-nucleotide mutations for all WES samples and those MSK-IMPACT sequenced samples with five or more such mutations. The fraction of mutations attributable to each of 30 known mutational signatures⁴⁶ was determined using a basin-hopping algorithm (<https://github.com/mskcc/mutation-signatures>), which assigns a weight to each of the 30 signatures based on the distribution of six types of single-nucleotide substitutions (C to A, G, or T; T to A, C, or G) and their trinucleotide context in a sample. For the purposes of cross-validating the source of somatic hypermutation, signatures 6, 14, 15, 20, 21 and 26 were considered together as mismatch-repair deficiency/MSI-associated, signature 10 was POLE-associated, and signature 11 was alkylating therapy exposure-related. In addition to signature 3 association with HRD, we also calculated three separate surrogate markers of the HRD phenotype¹³ at the level of DNA copy number alterations. Scores for large-scale transitions (LST), number of telomeric allelic imbalances (NtAI), and HRD deficiency (HRD-LOH) were inferred as previously described¹³ using total or allele-specific copy number segmentation data from FACETS analysis of exome and targeted sequencing data (Supplementary Table 6). Due to the correlation of the HRD phenotype measured by multiple independent metrics among patients, genotypes, and zygosity in the study cohort (Extended Data Fig. 4D–E), we calculated a single composite score that combined two individual scores that together represents orthogonal molecular measures of HRD: somatic single-nucleotide mutational signature 3 and LST among DNA copy number changes. We first standardized each of the individual scores into a Z-score (mean-centered and standard deviation-scaled), and then averaged these standardized values per case into an aggregate Z-score. This composite HRD score is, consequently, highly correlated with the values of the individual scores ($\rho=0.89$, $p\text{-value}=1\text{e-}270$).

To establish a control population of presumed HR-intact tumors for comparison, we inferred the mutational signatures of HRD in tumors that excluded any case with a germline or somatic mutation in an expanded set of genes presumed to be HR effector genes: *ATM*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CHEK1*, *CHEK2*, *FANCA*, *FANCC*, *MRE11A*, *NBN*, *PALB2*, *RAD50*, *RAD51*, *RAD51B*, *RAD51C*, *RAD51D*, *RAD52*, *RAD54L*, and all serous ovarian cancers given work indicating that the majority of such tumors harbor evidence of HRD⁴⁷. This does not preclude the possibility of occult HRD in a small number of cases (including those with promoter hypermethylation or those driven by still unknown mechanisms), but nevertheless established a baseline distribution of signature values for HRD from which to draw comparison. To evaluate samples with a predominant mutational signature other than HRD (signature 3), we considered samples with at least 50 somatic SNVs for which <5% of the mutational burden was attributed to HRD and greater than 25% were attributed to another signature of known etiology.

Publicly available methylation and mutational data

To assess the relationship between *BRCA1/2* mutations and epigenetic silencing pan-cancer, we used The Cancer Genome Atlas dataset. Somatic mutations in *BRCA1* and *BRCA2* were

classified as loss-of-function and mutational signatures were inferred in a manner identical to our prospective cohort using the consensus MC3 mutation calls (gdc.cancer.gov/about-data/publications/mc3-2017). Pathogenic and likely pathogenic germline variants and promoter methylation for *BRCA1/2* were acquired from published sources^{48,49}. For ovarian cancers, which were excluded from published resources, we determined *BRCA1/2* promoter methylation using a linear discriminant analysis that leveraged gene expression data and methylation status as a training set⁴⁹. Germ cell tumors were excluded from analysis due to the dynamic nature of DNA methylation changes in primordial germ cells that are not observed in somatic tissues.

Outcome analyses

Clinical endpoints were determined as follows. For PARP inhibitor therapy, outcomes were the percent of patients on treatment relative to treatment duration. For assessing clinical benefit of immune checkpoint blockade therapy, the clinical endpoint was estimated as event-free survival defined as the time to the next line of non-immunotherapy treatment (agent/modality), death, or censoring in cases lost to follow up. All therapeutic outcomes for both PARP inhibitor therapy and checkpoint blockade were assessed for the first line of the indicated therapy only. All outcomes-associated p-values and estimates of hazard ratios were generated from univariate log-rank tests or multivariable Cox proportional hazards models.

Data and code availability

The whole-exome sequencing data as well as germline calls have been deposited in the NCBI dbGaP archive under accession numbers phs001783.v1.p1 and phs001858.v1.p1, respectively. All other genomic and clinical data accompanies the manuscript and is available as Extended Data. Source code for these analyses is available at <https://github.com/taylor-lab/BRCA>.

Extended Data

Extended Data Fig. 1: Study cohort and *BRCA1/2* germline and somatic mutation distribution.

a) The number of tumor and matched normal specimens are shown by cancer type. NSCLC, non-small cell lung cancer; CUP, cancer of unknown primary; GIST, gastrointestinal stromal tumor; SCLC, small-cell lung cancer; GINET, gastrointestinal neuroendocrine tumor; CNS, non-glioma central nervous system tumors; NHL, non-Hodgkin's lymphoma. **b)** Somatic mutational burden (log₂ mutations per Mb) in tumors defined as non-hypermutated or hypermutated (via microsatellite instability, DNA polymerase epsilon mutations, or alkylating therapy-induced; see Methods). **c)** *BRCA1* and *BRCA2* somatic mutation rates in deciles of increasing tumor mutational burden. The highest mutational burden tumors also had the highest rate of *BRCA1/2* mutations. **d)** The percent of tumors in each tumor type harboring either somatic VUS or LoF *BRCA1/2* mutations as a function of the median somatic mutational burden of that cancer type (excluding hypermutated cases). Overall, the rate of somatic LoF *BRCA1/2* mutations by cancer type modestly increased with their increasing mutational burden, while this was much more pronounced for *BRCA1/2* variants of uncertain significance. **e)** Population frequency comparisons are shown between the study cohort and gnomAD for allele frequencies (AF) of *BRCA1/2* germline pathogenic and likely pathogenic (P/LP) alleles and VUS (dark and light blue, respectively). Left, all alleles; center, only P/LP alleles; right, comparison between the study cohort and the germline results from the TCGA cohort. Ashkenzai Jewish (ASJ) founder *BRCA1/2* alleles are shown

(see legend). **f**) As in panel (e) for only the ASJ sub-populations. **g**) As in panels (e-f) but for the non-ASJ white subpopulation. NFE, non-Finnish European. **h**) The prevalence of homozygous deletions in *BRCA1* or *BRCA2* in affected cancer types. Count of affected tumors in parentheses, inset is the fraction of all homozygous deletions of either gene. **i**) Average age of first cancer diagnosis for *BRCA1/2* germline carriers compared to those patients lacking any pathogenic germline alteration (germline WT) in BRCA-associated cancer types and all other cancer types. Error bars are 95% confidence interval (CI), asterisk reflects significance (two-sided Wilcoxon test, p-value<0.01; NS, not significant). **j**) Percent of *BRCA1/2* germline carriers with multiple independent cancer diagnoses compared to germline WT patients. Error bars are 95% CIs, asterisk is p-value=0.02, chi-square test. **k**) The fraction of all germline pathogenic or somatic LoF alterations in *BRCA1* versus *BRCA2* (in non-hypermutated tumors).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Extended Data Fig. 2: *BRCA1/2* zygosity.

a) Diagrammatic representation of the integration of allele-specific copy number with purity and mutant allele frequencies to determine the zygosity of the germline pathogenic allele in the corresponding tumor (and the mechanism of its selection; CN-LOH, copy-neutral LOH).

b) In only a subset of cases of low tumor cell content (<30%) does the LOH inference become increasingly analytically challenging (increasing rate of indeterminate calls).

c) The percent of cases with LOH affecting the germline pathogenic or somatic LoF *BRCA1/2* mutations (as labeled) as a function of tumor purity (see Methods). While somatic mutant allele frequencies are impacted by tumor purity, this does not affect the sensitivity for LOH detection for germline variants and only affects sensitivity for LOH of somatic mutations in tumors of <30% purity.

d) In tumors with benign germline variants in *BRCA1* and *BRCA2*, the ratio of zygosity changes affecting the WT or mutant *BRCA1/2* allele is approximately 0.5, indicating neutral selection. By contrast, the rate of zygosity changes leading to loss of the WT allele in patients with germline pathogenic *BRCA1* or *BRCA2* mutations (>80%) is consistent with selective pressure for biallelic inactivation.

e) Integrating all measurable sources of biallelic inactivation (inset, somatic sequence variants as the source of second hits to WT *BRCA1/2*), the percent of tumors by cancer type harboring a biallelic *BRCA1* or *BRCA2* loss.

f) The rate of biallelic inactivation of *BRCA1* versus *BRCA2* in patients with germline pathogenic or somatic LoF mutations (in hypermutated and non-hypermutated tumors). P-values as indicated, two-sided Fisher's exact test. NS, not significant.

g) The rate of loss of WT *BRCA1* or *BRCA2* (LOH) in patients with germline deleterious *BRCA1* or *BRCA2* mutations compared to rare benign variants in either gene in BRCA-associated cancer types and in those not conventionally associated with BRCA germline carriers.

Asterisks reflects significance (Fisher's exact test).

h) The rate of biallelic inactivation of *BRCA1/2* in patients with germline pathogenic or somatic LoF mutations pan-cancer as a function of primary or metastatic specimen type. At right, the four BRCA-associated cancer

types are shown individually. P-value as indicated, two-sided Fisher's exact test. N.S. is not significant. Error bars are 95% CIs. **i)** The rate of LOH spanning germline or somatic mutant *BRCA1* and *BRCA2* in breast cancers (colored as in Fig. 2B–C) as well as other somatically mutated tumor-suppressor genes. Error bars are binomial CIs.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Extended Data Fig. 3: Somatic loss of the pathogenic germline *BRCA1/2* allele.

a) Schematic representation of the different allelic configurations that would lead to the retention or loss of a germline allele in the presence of a somatically mutated tumor suppressor gene (TSG) responsible for driving biallelic inactivation. **b)** Among tumors with loss of the pathogenic germline allele (in either *BRCA1* or *BRCA2*, as indicated), the pattern of somatic mutations in known TSGs on their respective chromosomes (*TP53* and *NF1* are encoded on chromosome 17 on which *BRCA1* also appears, while *RB1* is encoded on chromosome 13 on which *BRCA2* also appears) arising in the same tumors and *in trans* with, and presumed to drive the loss of, the germline allele. **c)** In a representative EML4-ALK positive lung adenocarcinoma diagnosed in a *BRCA1* E23Vfs*17 carrier, LOH preceding whole-genome doubling spanned chromosome 17 encoding *TP53* R248Q arising *in trans* with the mutant *BRCA1* allele. Dark and light blue represent the major and minor copy number at the indicated loci. **d)** Somatic mutant allele fractions (for case in panel C) are consistent with deletion of the allele harboring the *BRCA1* founder mutation as compared to the observed and expected values for clonal heterozygous somatic mutations (*RAD50*) or biallelic inactivation of mutant *TP53* (tumor purity is Φ). The selective pressure for biallelic *TP53* inactivation driven by the initial R248Q mutation likely precipitated the subsequent heterozygous loss of the WT *TP53* allele, leading to deletion of the *BRCA1* pathogenic mutation and retention of the *BRCA1* WT allele, indicating that mutant *BRCA1* was likely dispensable for its pathogenesis.

Extended Data Fig. 4: HRD phenotype in *BRCA1/2*-mutant cancers characterized by WES.

a) Total number of prospectively sequenced cases by cancer type for which exome re-sequencing was obtained. Abbreviations as in Extended Data Fig. 1; PNS, peripheral nervous system. **b)** The distribution of cancer types among *BRCA1/2*-mutant (germline or somatic) cases with exome re-sequencing data. **c)** The proportion of *BRCA1/2*-mutant cases with exome re-sequencing data by germline or somatic mutational origin. **d)** The somatic single-nucleotide mutational signature 3 of HRD, and **e)** the DNA copy number-based LST metric of HRD as inferred from exome sequencing data are shown as a function affected cancer types (left) and *BRCA1/2* mutation origin and zygosity (right, see legend at bottom of panel **E**) as in main text Fig. 3. Asterisks reflect p-values < 0.01, 10^{-10} , and 10^{-20} respectively, two-sided Student's t-test. The individual metrics are highly correlated with the composite HRD score ($\rho=0.89$, p-value = $1e-270$; see Methods) and consequently the

qualitative results based on lineage, mutational origin and zygosity are similar. **f)** The rate of *BRCA1* promoter methylation in ovarian, breast, and other cancer types (no evidence of *BRCA2* silencing via promoter methylation was apparent). Inset, *BRCA1* germline mutations and promoter methylation leading to *BRCA1* silencing are mutually exclusive in affected cancers, indicating that heterozygous *BRCA1*-mutant tumors typically do not acquire biallelic inactivation via epigenetic silencing of the remaining allele. Epigenetic silencing is therefore unlikely to fully explain the modest HRD phenotype in heterozygous mutant tumors (Fig. 3B). Both germline and somatic mutational data and DNA methylation data was acquired from The Cancer Genome Atlas (see Methods). **g)** The composite measure of HRD in pan HR-wildtype tumors (light gray) and in tumors with either germline or somatic *BRCA1* or *BRCA2* mutations (dark gray) grouped by BRCA-associated cancer types (dark red: breast, ovary, pancreas, prostate) versus other cancer types (red), and tumors with somatic hypermutation (light red, see legend at bottom). Significant differences as indicated, two-sided Student's t-test. NS, not significant. **h)** Same as in panel **G** and main text Fig. 3c, grouped by a combination of lineage, origin, and zygosity (see legend).

Extended Data Fig. 5: Intra-individual BRCA phenotypic divergence.

Exome sequencing of two cancer diagnoses in a founder *BRCA2* S1982Rfs*22 germline carrier revealed two independent and clonally unrelated cancer, one an HR-deficient serous ovarian cancer (left) with loss of WT *BRCA2*, the other a co-incident cholangiocarcinoma with intact WT *BRCA2* (right). The latter had a different pattern of somatic abnormality and lacked any evidence of HRD (top) despite the shared germline pathogenic *BRCA2* allele.

Extended Data Fig. 6: Tumor mutational burden by *BRCA1/2* genotype.

The somatic mutational burden of tumors as a function of cancer type, *BRCA1/2* mutation origin, and somatic *BRCA1/2* zygosity. Statistically significant differences among the indicated comparisons are shown. Two-sided Student's t-test. NS, not significant. Error bars are the 95% confidence intervals.

Extended Data Fig. 7: *BRCA1/2* mutations attributable to other mutational signatures.

The somatic mutations in each of the evaluable cancers in main Fig. 3D in which a *BRCA1* or *BRCA2* somatic heterozygous mutation arose in a motif consistent with an alternative non-HRD mutational signature. The mutation (trinucleotide context, base change, and protein annotation) is indicated in each case as is its cancer type.

Extended Data Fig. 8: PARP inhibitor therapy by *BRCA1/2* mutation origin and zygosity.

A single PARP inhibitor outcome analysis of all four BRCA genotypes as independent classes (*BRCA1/2* mutational origin and zygosity) with BRCA-associated cancer types (as in main text Fig. 4). All four classes of BRCA-mutant patients (see legend, as indicated) achieve significantly greater clinical benefit to PARP inhibitor therapy than do treated patients with WT BRCA tumors [*BRCA1/2*-mutant classes are germline carrier somatic heterozygous (HR=0.39, 0.21–0.72, p-value=0.003); germline carrier, somatic biallelic (HR=0.5, 0.35–0.72, p-value=2e-4); somatic heterozygous LoF (HR=0.5, 0.26–0.95, p-value=0.03); and somatic LoF biallelic (HR=0.34, 0.16–0.72, p-value=0.005)].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank our patients and their families for participating in this study and the members of Taylor lab and Marie-Josée and Henry R. Kravis Center for Molecular Oncology for discussions and support. We thank the Jonathan and Mindy Gray Foundation for their support from project inception. This work was also supported by National Institutes of Health awards P30 CA008748, U54 OD020355 (D.B.S., B.S.T.), R01 CA207244 (D.M.H., B.S.T.), R01 CA204749 (B.S.T.); and the American Cancer Society (RSG-15-067-01-TBG), Cycle for Survival, Sontag Foundation, Prostate Cancer Foundation, Anna Fuller Fund, and the Josie Robertson Foundation (B.S.T.).

References

1. Wooster R et al. Identification of the breast cancer susceptibility gene BRCA2. Nature 378, 789–792 (1995). [PubMed: 8524414]

2. Miki Y et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 266, 66–71 (1994). [PubMed: 7545954]
3. Roy R, Chun J & Powell SN BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nat. Rev. Cancer* 12, 68–78 (2011). [PubMed: 22193408]
4. Kuchenbaecker KB et al. Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. *JAMA* 317, 2402–2416 (2017). [PubMed: 28632866]
5. Paluch-Shimon S et al. Prevention and screening in BRCA mutation carriers and other breast/ovarian hereditary cancer syndromes: ESMO Clinical Practice Guidelines for cancer prevention and screening. *Ann. Oncol* 27, v103–v110 (2016). [PubMed: 27664246]
6. Maxwell KN et al. BRCA locus-specific loss of heterozygosity in germline BRCA1 and BRCA2 carriers. *Nat. Commun* 8, 319 (2017). [PubMed: 28831036]
7. Lord CJ & Ashworth A BRCAness revisited. *Nat. Rev. Cancer* 16, 110–120 (2016). [PubMed: 26775620]
8. Yu VP et al. Gross chromosomal rearrangements and genetic exchange between nonhomologous chromosomes following BRCA2 inactivation. *Genes Dev.* 14, 1400–1406 (2000). [PubMed: 10837032]
9. Moynahan ME, Pierce AJ & Jasin M BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol. Cell* 7, 263–272 (2001). [PubMed: 11239455]
10. Moynahan ME, Chiu JW, Koller BH & Jasin M Brca1 controls homology-directed DNA repair. *Mol. Cell* 4, 511–518 (1999). [PubMed: 10549283]
11. Alexandrov LB et al. Signatures of mutational processes in human cancer. *Nature* 500, 415–421 (2013). [PubMed: 23945592]
12. Davies H et al. HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. *Nat. Med* 23, 517–525 (2017). [PubMed: 28288110]
13. Marquard AM et al. Pan-cancer analysis of genomic scar signatures associated with homologous recombination deficiency suggests novel indications for existing cancer drugs. *Biomark Res* 3, 9 (2015). [PubMed: 26015868]
14. Moore K et al. Maintenance Olaparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. *N. Engl. J. Med* 379, 2495–2505 (2018). [PubMed: 30345884]
15. Robson M et al. Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *N. Engl. J. Med* 377, 523–533 (2017). [PubMed: 28578601]
16. Manickam K et al. Exome Sequencing–Based Screening for *BRCA1/2* Expected Pathogenic Variants Among Adult Biobank Participants. *JAMA Network Open* 1, e182140 (2018). [PubMed: 30646163]
17. Mandelker D et al. Mutation Detection in Patients With Advanced Cancer by Universal Sequencing of Cancer-Related Genes in Tumor and Normal DNA vs Guideline-Based Germline Testing. *JAMA* 318, 825–835 (2017). [PubMed: 28873162]
18. Cheng DT et al. Comprehensive detection of germline variants by MSK-IMPACT, a clinical diagnostic platform for solid tumor molecular oncology and concurrent cancer predisposition testing. *BMC Med. Genomics* 10, 33 (2017). [PubMed: 28526081]
19. Zehir A et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med* 23, 703–713 (2017). [PubMed: 28481359]
20. Levy-Lahad E & Friedman E Cancer risks among BRCA1 and BRCA2 mutation carriers. *Br. J. Cancer* 96, 11–15 (2007). [PubMed: 17213823]
21. Mersch J et al. Cancers associated with BRCA1 and BRCA2 mutations other than breast and ovarian. *Cancer* 121, 269–275 (2015). [PubMed: 25224030]
22. Scully R & Livingston DM In search of the tumour-suppressor functions of BRCA1 and BRCA2. *Nature* 408, 429–432 (2000). [PubMed: 11100717]
23. Timms KM et al. Association of BRCA1/2 defects with genomic scores predictive of DNA damage repair deficiency among breast cancer subtypes. *Breast Cancer Res.* 16, 475 (2014). [PubMed: 25475740]
24. Mouw KW, Goldberg MS, Konstantinopoulos PA & D'Andrea AD DNA damage and repair biomarkers of immunotherapy response. *Cancer Discov.* 7, 675–693 (2017). [PubMed: 28630051]

25. Nolan E et al. Combined immune checkpoint blockade as a therapeutic strategy for BRCA1-mutated breast cancer. *Sci. Transl. Med* 9, (2017).
26. Drilon A et al. Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. *N. Engl. J. Med* 378, 731–739 (2018). [PubMed: 29466156]
27. Le DT et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357, 409–413 (2017). [PubMed: 28596308]
28. Cheng DT et al. Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology. *J Mol Diagn* 17, 251–264 (2015). [PubMed: 25801821]
29. Coombs CC et al. Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. *Cell Stem Cell* 21, 374–382.e4 (2017). [PubMed: 28803919]
30. Karczewski KJ et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *BioRxiv* (2019). doi:10.1101/531210
31. Spurdle AB et al. 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 (2012). [PubMed: 21990146]
32. Findlay GM et al. Accurate classification of BRCA1 variants with saturation genome editing. *Nature* 562, 217–222 (2018). [PubMed: 30209399]
33. Chakravarty D et al. Oncokb: A precision oncology knowledge base. *JCO Precis. Oncol* 2017, (2017).
34. Niu B et al. MSIsensor: microsatellite instability detection using paired tumor-normal sequence data. *Bioinformatics* 30, 1015–1016 (2014). [PubMed: 24371154]
35. Middha S et al. Reliable Pan-Cancer Microsatellite Instability Assessment by Using Targeted Next-Generation Sequencing Data. *JCO Precis. Oncol* 2017, (2017).
36. Johnson BE et al. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science* 343, 189–193 (2014). [PubMed: 24336570]
37. Shen R & Seshan VE FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res.* 44, e131 (2016). [PubMed: 27270079]
38. Bielski CM et al. Genome doubling shapes the evolution and prognosis of advanced cancers. *Nat. Genet* 50, 1189–1195 (2018). [PubMed: 30013179]
39. McGranahan N et al. Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. *Sci. Transl. Med* 7, 283ra54 (2015).
40. Mose LE, Wilkerson MD, Hayes DN, Perou CM & Parker JS ABRA: improved coding indel detection via assembly-based realignment. *Bioinformatics* 30, 2813–2815 (2014). [PubMed: 24907369]
41. DePristo MA et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet* 43, 491–498 (2011). [PubMed: 21478889]
42. Cibulskis K et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol* 31, 213–219 (2013). [PubMed: 23396013]
43. Lai Z et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res.* 44, e108 (2016). [PubMed: 27060149]
44. Chang MT et al. Accelerating discovery of functional mutant alleles in cancer. *Cancer Discov.* 8, 174–183 (2018). [PubMed: 29247016]
45. Chang MT et al. Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. *Nat. Biotechnol* 34, 155–163 (2016). [PubMed: 26619011]
46. Alexandrov LB et al. Clock-like mutational processes in human somatic cells. *Nat. Genet* 47, 1402–1407 (2015). [PubMed: 26551669]
47. Wang YK et al. Genomic consequences of aberrant DNA repair mechanisms stratify ovarian cancer histotypes. *Nat. Genet* 49, 856–865 (2017). [PubMed: 28436987]

48. Huang K-L et al. Pathogenic germline variants in 10,389 adult cancers. *Cell* 173, 355–370.e14 (2018). [PubMed: 29625052]
49. Knijnenburg TA et al. Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. *Cell Rep.* 23, 239–254.e6 (2018). [PubMed: 29617664]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1: The prevalence and origins of *BRCA1/2* mutations.

a) Prevalence and type of germline pathogenic and somatic mutations in *BRCA1* and *BRCA2*. Hypermutated tumors considered separately. **b)** The percent of *BRCA1/2*-mutant patients with a *BRCA*-associated cancer type (breast, ovary, prostate, pancreas) versus all other cancer types. **c)** *BRCA* alteration rates by gene and cancer type. Error bars are binomial confidence intervals (CIs), asterisks are p-value < 0.05 for differences between *BRCA1* and *BRCA2* per tumor type, McNemar's chi-square. Bottom, distribution of *BRCA1/2* alteration types. **d)** Percent of patients by cancer type with germline or somatic *BRCA1/2* alterations (red; p-value < 0.05, enrichment for germline or somatic, McNemar's chi-square).

Fig. 2: Lineage variation in selection for *BRCA1/2* biallelic inactivation.

a) The rate of LOH in *BRCA1/2* pathogenic germline carriers [blue; diamond, biallelic inactivation via any mechanism (66%)], somatic LoF *BRCA1/2* mutations, or *TP53* oncogenic mutations (red). In gray, the background rate of LOH spanning benign germline variants or somatic passenger mutations (see Methods, asterisks are p-values < 0.01, 10^{-10} , or 10^{-100} , respectively). NS, not significant. In all panels, error bars are binomial CIs. **b)** The rate of LOH in germline carriers compared to benign variants in BRCA-associated and select non-BRCA-associated cancer types. Asterisks reflects significance (two-sided Fisher's exact test). **c)** Cancer type-specific rates of biallelic inactivation by mutation origin (bottom) and mechanism thereof (top). Dotted line, background rate of LOH pan-cancer. **d)** Percent of *BRCA1/2* germline carriers that lose the pathogenic allele somatically in the indicated cancer types (p-value=0.003, two-sided Fisher's exact test).

Fig. 3: BRCA phenotypes are tumor lineage and zygosity-dependent.

a) The composite measure of HRD in pan-HR wildtype tumors versus those with *BRCA1/2* germline or somatic mutations stratified by cancer type (BRCA-associated, non-BRCA-associated, and hypermutated tumors; see legend at right). Asterisks are p-values < 0.01, 10^{-10} , and 10^{-20} respectively, two-sided Student's t-test. NS, not significant. **b)** As in panel **a**, grouped by mutation origin and zygosity (see legend). All individual comparisons with pan-HR wildtype tumors unless indicated with individual p-values. **c)** Same as in panels **a-b**, but grouped by a combination of lineage, origin, and zygosity (see legend). **d)** Tumors with somatic *BRCA1/2* mutations with a dominant non-HRD mutational signature indicating an alternative mechanism of pathogenesis (0–5% HRD, which does not exceed the background rate in pan-HR wildtype tumors, Supplementary Table 6). Bottom, the likelihood the *BRCA1/2* mutation was induced by the indicated signature (position and trinucleotide context indicative of the signature motif). Only samples with a dominant signature of known etiology are shown.

Fig. 4: Context-specific therapeutic sensitivity of *BRCA1/2*-mutant tumors.

a) Left, clinical benefit to PARP inhibitor therapy in patients with BRCA-associated cancer types with and without *BRCA1/2* mutations (germline or somatic) (HR 0.58, 95% CI 0.46–0.73, log-rank p-value=3.7e-6). Right, patients with all other cancer types (HR 1.02, 95% CI 0.6–1.7, log-rank p-value=0.98). **b)** In BRCA-associated cancer types, clinical benefit to PARP inhibition in patients with somatic LoF versus germline pathogenic *BRCA1/2* alterations. **c)** As in panel (b) but comparing clinical benefit in heterozygous versus biallelic *BRCA1/2*-mutant BRCA-associated cancers. **d-e)** Event-free survival from the start of the first line of immune checkpoint blockade therapy in patients pan-cancer with or without *BRCA1/2* germline or somatic mutations (HR=0.99, p-value=0.9 for non-hypermutated tumors). Multivariable model includes tumor mutational burden (>75th percentile) and affected cancer type.