

# Tumour lineage shapes BRCA-mediated phenotypes

Philip Jonsson<sup>1,2,3</sup>, Chaitanya Bandlamudi<sup>1</sup>, Michael L. Cheng<sup>4,7</sup>, Preethi Srinivasan<sup>5</sup>, Shweta S. Chavan<sup>1</sup>, Noah D. Friedman<sup>2,3</sup>, Ezra Y. Rosen<sup>4</sup>, Allison L. Richards<sup>1</sup>, Nancy Bouvier<sup>1</sup>, S. Duygu Selcuklu<sup>1</sup>, Craig M. Bielski<sup>1,2,3</sup>, Wassim Abida<sup>4</sup>, Diana Mandelker<sup>5</sup>, Ozge Birsoy<sup>5</sup>, Liying Zhang<sup>5</sup>, Ahmet Zehir<sup>5</sup>, Mark T. A. Donoghue<sup>1</sup>, José Baselga<sup>4,8</sup>, Kenneth Offit<sup>4</sup>, Howard I. Scher<sup>4</sup>, Eileen M. O'Reilly<sup>4</sup>, Zsofia K. Stadler<sup>4</sup>, Nikolaus Schultz<sup>1,3</sup>, Nicholas D. Soccil<sup>1</sup>, Agnes Viale<sup>1</sup>, Marc Ladanyi<sup>2,5</sup>, Mark E. Robson<sup>4</sup>, David M. Hyman<sup>4,6</sup>, Michael F. Berger<sup>1,5,6\*</sup>, David B. Solit<sup>1,2,4,6\*</sup> & Barry S. Taylor<sup>1,2,3,6\*</sup>

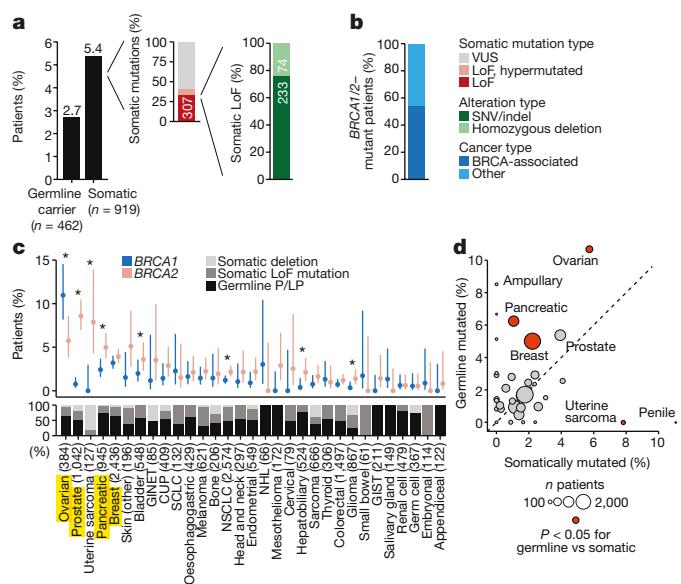
**Mutations in *BRCA1* and *BRCA2* predispose individuals to certain cancers<sup>1–3</sup>, and disease-specific screening and preventative strategies have reduced cancer mortality in affected patients<sup>4,5</sup>. These classical tumour-suppressor genes have tumorigenic effects associated with somatic biallelic inactivation, although haploinsufficiency may also promote the formation and progression of tumours<sup>6,7</sup>. Moreover, *BRCA1/2*-mutant tumours are often deficient in the repair of double-stranded DNA breaks by homologous recombination<sup>8–13</sup>, and consequently exhibit increased therapeutic sensitivity to platinum-containing therapy and inhibitors of poly-(ADP-ribose)-polymerase (PARP)<sup>14,15</sup>. However, the phenotypic and therapeutic relevance of mutations in *BRCA1* or *BRCA2* remains poorly defined in most cancer types. Here we show that in the 2.7% and 1.8% of patients with advanced-stage cancer and germline pathogenic or somatic loss-of-function alterations in *BRCA1/2*, respectively, selective pressure for biallelic inactivation, zygosity-dependent phenotype penetrance, and sensitivity to PARP inhibition were observed only in tumour types associated with increased heritable cancer risk in *BRCA1/2* carriers (BRCA-associated cancer types). Conversely, among patients with non-BRCA-associated cancer types, most carriers of these *BRCA1/2* mutation types had evidence for tumour pathogenesis that was independent of mutant *BRCA1/2*. Overall, mutant BRCA is an indispensable founding event for some tumours, but in a considerable proportion of other cancers, it appears to be biologically neutral—a difference predominantly conditioned by tumour lineage—with implications for disease pathogenesis, screening, design of clinical trials and therapeutic decision-making.**

Mutations in *BRCA1* and *BRCA2* are emerging as important therapeutic targets, but the current clinical development and continued off-label use of PARP inhibitors broadly assumes that germline and somatic mutations in *BRCA1* or *BRCA2* are tumour-agnostic biomarkers, with similar biological and therapeutic importance wherever they are identified. Moreover, as oncology shifts from limited germline screening in high-risk populations to broader panel or unbiased sequencing of matched tumour and normal DNA in unselected patients with cancer, studies are revealing a higher than anticipated prevalence of pathogenic germline mutations in cancer predisposition genes, including *BRCA1* and *BRCA2*<sup>16,17</sup>. Consequently, we sought to understand which common and rare cancers are phenotypically and therapeutically dependent on somatic and germline BRCA alterations.

We analysed the germline blood and matched tumour tissue of 17,152 patients with cancer diagnosed with 1 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 the *BRCA1* and *BRCA2* genes, and identified germline pathogenic and probable

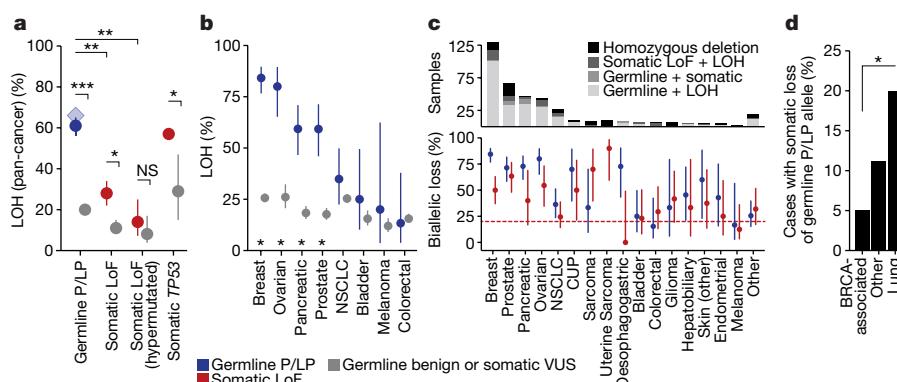
pathogenic variants (hereafter pathogenic *BRCA1/2*) using a clinically validated variant discovery pipeline and a custom pathogenicity classifier trained on expert curation by medical geneticists<sup>18,19</sup> (see Methods). Hypermutated tumours were considered separately as a biologically distinct class (Extended Data Fig. 1b–d).

Overall, 2.7% of patients ( $n = 462$ ) contained 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, Supplementary Table 3). Of the remaining 16,690 germline wild-type patients, 919 had a somatic mutation in *BRCA1* or *BRCA2*, of which



**Fig. 1 | The prevalence and origins of *BRCA1/2* mutations.** **a**, Prevalence and type of germline pathogenic and somatic mutations in *BRCA1* and *BRCA2* among 17,152 patients. Hypermutated tumours were considered separately. indel, insertion or deletion; SNV, single-nucleotide variant. **b**, The percentage of *BRCA1/2*-mutant patients with a BRCA-associated cancer type (breast, ovary, prostate or pancreas) compared with all other cancer types. **c**, The percentage of patients with BRCA alterations by gene and cancer type. P/LP denotes pathogenic and likely pathogenic. Error bars are binomial confidence intervals, number of patients per cancer type in parentheses. \* $P < 0.05$  for differences between *BRCA1* and *BRCA2* per tumour type, McNemar's  $\chi^2$  test. Bottom, distribution of types of *BRCA1/2* alterations. CUP, cancer of unknown primary; GINET, gastrointestinal neuroendocrine tumour; NHL, non-Hodgkin lymphoma; NSCLC, non-small-cell lung cancer. **d**, Percentage of patients by cancer type with germline or somatic *BRCA1/2* alterations.  $P < 0.05$ , enrichment for germline or somatic (red), McNemar's  $\chi^2$  test.

<sup>1</sup>Marie-Josée and Henry R. Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>2</sup>Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>3</sup>Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>4</sup>Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>5</sup>Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA. <sup>6</sup>Weill Cornell Medical College, New York, NY, USA. <sup>7</sup>Present address: Dana-Farber Cancer Institute, Boston, MA, USA. <sup>8</sup>Present address: AstraZeneca, Waltham, MA, USA. \*e-mail: bergerm1@mskcc.org; solitd@mskcc.org; taylorb@mskcc.org



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

**a**, The rate of LOH in *BRCA1/2* pathogenic germline carriers (blue,  $n = 430$  evaluable; diamond denotes biallelic inactivation via any mechanism (66%)), or somatic LoF *BRCA1/2* mutations ( $n = 202$ ), or *TP53* oncogenic mutations (red,  $n = 6,455$ ). The background rate of LOH spanning benign germline variants or somatic passenger mutations is shown in grey (see Methods).  $*P < 0.01$ ,  $**P < 1 \times 10^{-10}$ ,  $***P < 1 \times 10^{-100}$ , two-sided Fisher's exact test. NS, not significant. In all panels, error bars are binomial confidence intervals. **b**, The rate of LOH in germline

carriers compared with benign variants in *BRCA*-associated and select non-*BRCA*-associated cancer types.  $*P < 0.01$ , two-sided Fisher's exact test. P/LP variants,  $n = 120, 40, 59, 54, 43, 16, 5$  and 15, from left to right, respectively. **c**, Cancer-type-specific rates of biallelic inactivation by mutational origin (bottom) and mechanism thereof (top). Dotted line denotes background rate of LOH pan-cancer. **d**, Percentage of *BRCA1/2* germline carriers ( $n = 273, 45$  and 133, from left to right, respectively) that lose the pathogenic allele somatically in the indicated cancer types.  $*P = 0.003$ , two-sided Fisher's exact test.

most were missense mutations of uncertain importance (59%). In total, 307 patients (1.8%) had a somatic presumed LoF mutation in *BRCA1/2* in a non-hypermutated cancer (Fig. 1a, 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* ( $P = 1.5 \times 10^{-9}$ ; Extended Data Fig. 1h).

Patients with germline pathogenic or somatic LoF mutations in *BRCA1/2* (4.9% in total) had one of 38 cancer types of which 53% were breast, ovary, prostate or pancreatic cancers—types known to be associated with *BRCA1/2* germline carrier status<sup>20</sup> (hereafter referred to as *BRCA*-associated cancer types, see Methods). Notably, these were the only four cancer types that after 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 wild-type patients with cancer<sup>21</sup> (Extended Data Fig. 1i, j).

Both germline pathogenic and somatic LoF mutations were modestly more common in *BRCA2* than in *BRCA1* pan-cancer ( $P < 10^{-7}$ ; Extended Data Fig. 1k), although considerable variability existed among individual cancer types, suggesting lineage-based dependencies perhaps related to the distinct roles of these two genes in tumour initiation and progression<sup>3,22</sup> (Fig. 1c). The germline and somatic origins of *BRCA1/2* alterations were approximately equal across all cancer types, although key cancer types favoured 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,  $P = 9 \times 10^{-14}$ , Fig. 1d), except in prostate cancer, in which they occurred with approximately equal frequency. By contrast, in the non-*BRCA*-associated cancer uterine sarcoma, somatic *BRCA2* homozygous deletions accounted for nearly all alterations in *BRCA1/2*.

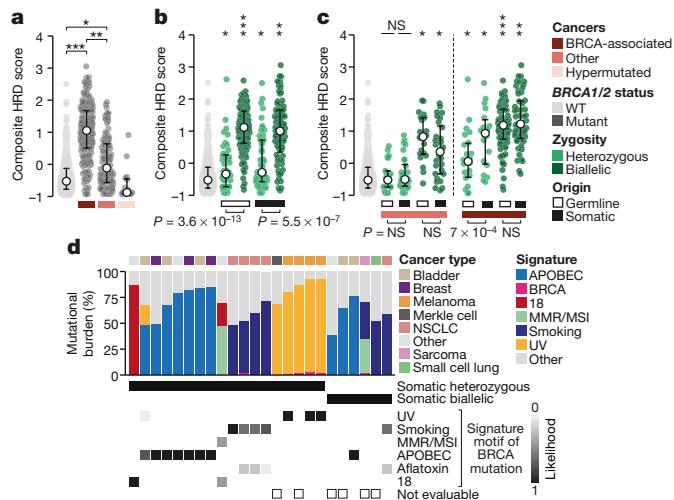
To assess the dependence on *BRCA* dysfunction, we determined the selective pressure for wild-type *BRCA* loss in tumours of germline or somatic carriers by integrating purity- and ploidy-corrected allele-specific copy number with high-precision mutant allele fractions (see Methods, Extended Data Fig. 2a–c). To model neutral evolution, we determined the specificity of wild-type *BRCA* loss by comparing it with a cancer-type-specific background distribution of loss of heterozygosity (LOH) spanning either benign germline variants or somatic variants of unknown significance (VUS) in *BRCA1/2* predicted to have no effect on fitness. As expected, LOH affected the wild-type and benign germline or 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 wild-type *BRCA* allele ( $P = 4.4 \times 10^{-36}$ ), consistent with positive selective pressure for biallelic inactivation in these tumours.

Across all cancer types, 61% of *BRCA1/2* germline carriers had somatic LOH affecting the wild-type allele, which was significantly enriched over the background rate of LOH in tumours with benign variants (20%,  $P < 10^{-100}$ ). Although LOH affecting somatic LoF mutations in *BRCA1/2* was also enriched over background (28% versus 11%,  $P = 3.8 \times 10^{-7}$ ), it was more than twofold less prevalent than for germline pathogenic variants ( $P = 4.2 \times 10^{-15}$ ) and lower than the rate of LOH affecting oncogenic mutations in *TP53* ( $P = 9 \times 10^{-17}$ ; Fig. 2a). This pattern was consistent across *BRCA* gene and specimen type (primary versus metastatic; Extended Data Fig. 2e–h). In hypermutated tumours 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 previous association such as lung, bladder, melanoma and colorectal cancers in which it was largely absent (Fig. 2b), a pattern also evident for somatic mutations (75% versus 39%,  $P = 7.5 \times 10^{-22}$ ; 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 (approximately 70%, Fig. 2c).

Given the different rates of LOH across tumour types, we postulated that not all tumours 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. Although this was least common in *BRCA*-associated cancer types, it occurred significantly more frequently in lung cancers (5% versus 20%, respectively,  $P = 6.4 \times 10^{-11}$ ; Fig. 2d), apparently driven in some cases by selective pressure for biallelic inactivation of a proximal tumour suppressor gene somatically mutated *in trans* (Extended Data Fig. 3). In summary, these data suggest that selection for loss of wild-type *BRCA* was dependent on tumour lineage, 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 to be independent of mutant *BRCA*.

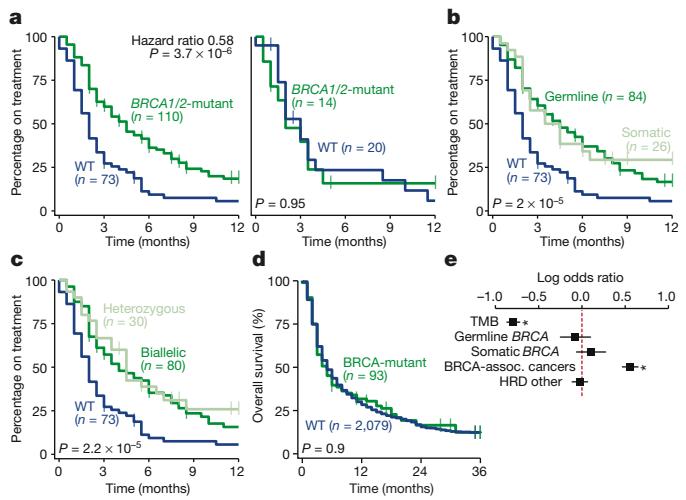
To explore how *BRCA1/2* mutations dictated phenotypic changes in affected cancers, we performed whole-exome sequencing of tumour-normal pairs from 815 patients (293 and 522 *BRCA1/2*-mutant and



**Fig. 3 | BRCA phenotypes are tumour lineage and zygosity-dependent.** **a**, The composite measure of HRD in pan-homologous-recombination wild-type (WT,  $n = 461$ ) tumours versus those with *BRCA1/2* germline or somatic mutations stratified by cancer type (BRCA-associated, non-BRCA-associated, and hypermutated tumours,  $n = 176$ , 97 and 15, respectively). \* $P < 0.01$ , \*\* $P < 1 \times 10^{-10}$ , \*\*\* $P < 1 \times 10^{-20}$ , two-sided Student's *t*-test. Circles and horizontal lines denote median and lower and upper quartiles, respectively. **b**, As in **a**, grouped by mutational origin and zygosity. All individual comparisons with pan-homologous-recombination wild-type tumours unless indicated by individual *P* values. From left to right,  $n = 461$ , 43, 104, 46 and 79, respectively. **c**, As in **a** and **b**, but grouped by a combination of lineage, origin and zygosity. From left to right,  $n = 461$ , 21, 28, 19, 28, 22, 18, 85 and 51, respectively. **d**, Tumours 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-homologous-recombination wild-type tumours, 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 aetiology are shown. MMR, mismatch repair; MSI, microsatellite instability.

wild-type cases, respectively; Extended Data Fig. 4a–c, Supplementary Table 5). We inferred two orthogonal signatures of homologous recombination deficiency (HRD), one each from somatic mutations and large-scale transitions in contiguous copy number alterations<sup>11,23</sup>, that we combined into a single composite measure of HRD (see Methods, Extended Data Fig. 4d, e, Supplementary Table 6). *BRCA1/2*-mutant tumours of BRCA-associated cancer types had a greater degree of HRD than did mutant tumours of non-BRCA-associated cancers, and far more than homologous-recombination-intact tumours (see Methods, Fig. 3a). Hypermutated tumours with somatic *BRCA1/2* mutations had no more HRD than wild-type cancers, which together with their low rate of biallelic inactivation, is consistent with these mutations being probable passenger events.

We next investigated how tumour lineage and *BRCA1/2* zygosity interact to determine the degree of HRD. Pan-cancer, heterozygous *BRCA1/2*-mutant tumours had only modestly greater HRD than wild-type cancers, which increased after 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 were mutant dose-dependent, increasing with each successive *BRCA* hit. By contrast, only the small subset of non-BRCA-associated cancers with biallelic inactivation had any evidence of HRD (Fig. 3c, 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 tumours (Extended Data Fig. 5). Collectively, these data suggest that different tumour lineages vary in their phenotypic susceptibility to a *BRCA1/2* defect, which has a central role in mediating tumour development in



**Fig. 4 | Context-specific therapeutic sensitivity of *BRCA1/2*-mutant tumours.** **a**, Left, clinical benefit to PARP inhibitor therapy in patients with BRCA-associated cancer types with and without *BRCA1/2* mutations (germline or somatic) (hazard ratio 0.58, 95% confidence interval 0.46–0.73, log-rank  $P = 3.7 \times 10^{-6}$ ). Right, patients with all other cancer types (hazard ratio 1.02, 95% confidence interval 0.6–1.7, log-rank  $P = 0.98$ ). **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 **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 with pan-cancer with or without *BRCA1/2* germline or somatic mutations (hazard ratio 0.99,  $P = 0.9$  for non-hypermutated tumours). Multivariable model includes tumour mutational burden (above the 75th percentile) and affected cancer type, lines are 95% confidence intervals.

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 ( $n = 21$  out 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, Extended Data Figs. 6, 7). Although this analysis cannot exclude BRCA haploinsufficiency mediating non-HRD driven tumorigenesis<sup>7</sup>, it indicates that most somatic heterozygous *BRCA1/2* mutations in non-BRCA-associated cancers are neutral passenger mutations, which similarly to those in hypermutated tumours, may be a consequence rather than the cause of tumorigenesis.

These results indicate 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 probably result in the inclusion of a considerable number of patients with non-BRCA-driven tumours. 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; hazard ratio 0.58, 95% confidence interval 0.46–0.73, log-rank  $P = 3.7 \times 10^{-6}$ ; Fig. 4a, left). However, this was not true in patients with non-BRCA-associated cancer types ( $n = 14$  and 20, respectively; hazard ratio 1.02, 95% confidence interval 0.6–1.7,  $P = 0.98$ ; Fig. 4a, right). Nevertheless, the longest clinical benefit from PARP inhibition among these tumours was a uterine sarcoma with a homozygous *BRCA2* deletion, which arises in 6.5% of all such tumours (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  $P = 2 \times 10^{-5}$ ; Fig. 4b) and the clinical benefit to PARP inhibition was also similar in heterozygous and biallelic patients ( $n = 30$  and 80, respectively; hazard ratio 0.44 and 0.47,  $P = 6.7 \times 10^{-4}$  and  $3.6 \times 10^{-5}$ , respectively; Fig. 4c, 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 heterozygous *BRCA1/2* mutations in non-BRCA-associated cancer types in whom HRD was largely absent remained on PARP inhibitor therapy for more than four 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 that the rate of *BRCA1/2* mutations increased with increasing mutational burden and these tumours generally lacked the phenotypic evidence of BRCA dependence, we postulated that despite suggested links<sup>24,25</sup>, *BRCA1/2* alterations would not further discriminate patients likely to benefit from immune checkpoint blockade beyond the influence of increased tumour mutational burden. We found no association between LoF *BRCA1/2* alterations and event-free survival of patients with pan-cancer from the start of immune checkpoint blockade therapy in 2,246 treated patients (hazard ratio 0.99, 95% confidence interval 0.78–1.3,  $P = 0.9$ ; Fig. 4d), even after adjusting for tumour mutational burden and affected cancer type or when considering the individual BRCA genes, mutational origin, or zygosity (Fig. 4e).

Our results indicate that mutant *BRCA1* and *BRCA2* have pleiotropic effects that are tumour lineage-dependent and that most *BRCA1/2* alterations in non-BRCA associated cancers may be incidental findings unrelated to tumour 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 tumour lineages. Further mechanistic study is required to determine the mechanisms that mediate differences in tolerance to defects in homologous recombination across various tumour lineages. Despite the recent emergence of tissue-agnostic biomarkers of cancer therapy response such as NTRK fusions<sup>26</sup> and microsatellite instability<sup>27</sup>, our data suggest that mutant *BRCA1/2* is unlikely to be of similar therapeutic relevance in all cancer types in which it is found. Although the identification of germline carriers will continue to be important for broader reduction of cancer risk, we caution that the integration of both germline and somatic *BRCA1/2* mutational status with broader somatic tumour profiles will ultimately be necessary to identify the subset of non-BRCA-associated cancers with true phenotypic dependence on mutant BRCA.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1382-1>.

Received: 3 January 2019; Accepted: 12 June 2019;  
Published online 10 July 2019.

1. Wooster, R. et al. Identification of the breast cancer susceptibility gene *BRCA2*. *Nature* **378**, 789–792 (1995).
2. Miki, Y. et al. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* **266**, 66–71 (1994).
3. Roy, R., Chun, J. & Powell, S. N. *BRCA1* and *BRCA2*: different roles in a common pathway of genome protection. *Nat. Rev. Cancer* **12**, 68–78 (2011).
4. Kuchenbaecker, K. B. et al. Risks of breast, ovarian, and contralateral breast cancer for *BRCA1* and *BRCA2* mutation carriers. *J. Am. Med. Assoc.* **317**, 2402–2416 (2017).
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** (suppl 5), v103–v110 (2016).
6. Maxwell, K. N. et al. *BRCA* locus-specific loss of heterozygosity in germline *BRCA1* and *BRCA2* carriers. *Nat. Commun.* **8**, 319 (2017).
7. Lord, C. J. & Ashworth, A. *BRCA*ness revisited. *Nat. Rev. Cancer* **16**, 110–120 (2016).
8. Yu, V. P. et al. Gross chromosomal rearrangements and genetic exchange between nonhomologous chromosomes following *BRCA2* inactivation. *Genes Dev.* **14**, 1400–1406 (2000).
9. Moynahan, M. E., Pierce, A. J. & Jasins, M. *BRCA2* is required for homology-directed repair of chromosomal breaks. *Mol. Cell* **7**, 263–272 (2001).
10. Moynahan, M. E., Chiu, J. W., Koller, B. H. & Jasins, M. *Brca1* controls homology-directed DNA repair. *Mol. Cell* **4**, 511–518 (1999).
11. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
12. Davies, H. et al. HRDetect is a predictor of *BRCA1* and *BRCA2* deficiency based on mutational signatures. *Nat. Med.* **23**, 517–525 (2017).
13. Marquard, A. M. 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).
14. Moore, K. et al. Maintenance olaparib in patients with newly diagnosed advanced ovarian cancer. *N. Engl. J. Med.* **379**, 2495–2505 (2018).
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).
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).
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. *J. Am. Med. Assoc.* **318**, 825–835 (2017).
18. Cheng, D. T. 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).
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).
20. Levy-Lahad, E. & Friedman, E. Cancer risks among *BRCA1* and *BRCA2* mutation carriers. *Br. J. Cancer* **96**, 11–15 (2007).
21. Mersch, J. et al. Cancers associated with *BRCA1* and *BRCA2* mutations other than breast and ovarian. *Cancer* **121**, 269–275 (2015).
22. Scully, R. & Livingston, D. M. In search of the tumour-suppressor functions of *BRCA1* and *BRCA2*. *Nature* **408**, 429–432 (2000).
23. Timms, K. M. 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).
24. Mouw, K. W., Goldberg, M. S., Konstantinopoulos, P. A. & D'Andrea, A. D. DNA damage and repair biomarkers of immunotherapy response. *Cancer Discov.* **7**, 675–693 (2017).
25. Nolan, E. et al. Combined immune checkpoint blockade as a therapeutic strategy for *BRCA1*-mutated breast cancer. *Sci. Transl. Med.* **9**, eaal4922 (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).
27. Le, D. T. et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* **357**, 409–413 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

## METHODS

**Data reporting.** No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

**Study cohort and prospective sequencing.** The study cohort consisted of 18,392 tumour 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 (IRB) and all patients provided written informed consent for tumour sequencing and review of medical records for detailed demographic, pathology, and treatment information. Genomic sequencing was performed on tumour DNA extracted from formalin-fixed paraffin-embedded 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 previously described<sup>19,28</sup>. 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 tumours, and the remainder were primary tumour 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 tumour 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 and expert review of medical records. 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 were included in analyses. In total, 199 of these patients had at least one sequenced specimen before the start of therapy. Patients treated with one or more lines of immune checkpoint blockade therapy included those receiving 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; NCT01775072). For the remainder of patients, their genomic and clinical data were anonymized before analysis with a deterministic one-way hash function. In these patients, germline variant calling was performed using the clinically validated pipeline used for the identified analysis described above<sup>18</sup>. Those germline variants that were suspected to have resulted from clonal haematopoiesis or circulating cell-free DNA (cfDNA) from the tumour were excluded (as described in P.S. et al., manuscript in preparation). In brief, excluded variants were those with insufficient aligned read coverage (less than 100- or 50-fold in the matched blood normal or tumour, respectively) unless spanned by a somatic homozygous deletion, or those occurring in a clonal-haematopoiesis-associated gene<sup>29</sup> with a variant allele frequency (VAF) in the blood normal  $< 0.35$  or 0.25 for single-nucleotide variants and indels, respectively, or a somatic VAF  $< 0.25$ . Germline variants were retained for further analysis if their observed VAF in the tumour specimen was within the 95% confidence interval of the expected VAF for the variant inferred using the tumour 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 tumour specimens, respectively. Finally, known oncogenic variants of low VAF in normal blood specimens and with a ratio of tumour-to-normal sample VAFs greater than three were presumed to have originated in tumour-derived circulating cfDNA and excluded. The remaining germline variants were annotated with Variant Effect Predictor (VEP, v.88) using vcf2maf v.1.6.10 (<https://github.com/mskcc/vcf2maf>) as well as ClinVar and myvariant.info (accessed September 2017), and subsequently filtered to include only rare variants by excluding common variants with a minor allele frequency above 2% using population frequencies from the Genome Aggregation Database (gnomAD)<sup>30</sup>.

Variant classification as pathogenic or likely pathogenic was determined using a random forest classifier trained on expert curation by medical geneticists. In brief, the classifier used a feature matrix, which consisted of multiple independently and partly overlapping feature categories including variant and gene type (oncogene versus tumour 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. Tenfold cross-validation achieved 94% ( $\pm 6\%$  confidence interval) and 89% ( $\pm 7\%$  confidence interval) average precision and recall, respectively (P.S. 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 consortium<sup>31</sup>. 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 study<sup>32</sup>. These variants (*BRCA1* C24Y and E85K) were classified as LoF by experimental methods but not predicted pathogenic by our classification. Of note, neither tumour containing 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 samples<sup>28</sup>. After correction for GC-bias, genic and intragenic gains and losses at gene and individual exon level were called at fold-change thresholds of 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 tumours of patients (comparing tumour 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 of single-nucleotide polymorphisms covered by the MSK-IMPACT assay design. Subpopulations were identified using self-reported race for the patients who consented to germline testing. Individuals of Ashkenazi Jewish ancestry were identified within the cluster comprising people who self-reported as white and confirmed as those individuals carrying any of the 60 alleles enriched in the gnomAD Ashkenazi Jewish subpopulation compared with the non-Finnish European subpopulation. To test for associations between mutant *BRCA1* and *BRCA2* and cancer types, we developed an ancestry-controlled permutation-based framework (P.S. et al., manuscript in preparation). In brief, for every cancer type of sufficient sample size (50 or more), we tested for the enrichment of germline mutations in *BRCA1* or *BRCA2* relative to a background distribution of frequencies generated from all other tumour 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 rate-corrected (Benjamini and Hochberg)  $q < 0.15$ .

**Somatic mutational analyses.** Somatic mutations (substitutions and small insertions and deletions), gene-level focal copy number alterations, and structural rearrangements were detected with a clinically validated pipeline as previously described<sup>19,28</sup>. Somatic alterations were classified as oncogenic or likely oncogenic using OncoKB<sup>33</sup>. Any somatic mutations not otherwise classified known or likely oncogenic were considered variants of uncertain significance (VUS). 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. Tumours were classified as hypermutated and thus analysed separately if they had LoF mutations owing 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 was determined for all tumour samples using MSIsensor as previously clinically validated<sup>34,35</sup>. Additional DNA mismatch-repair-mediated hypermutated tumours were those for whom univariate  $k$ -means clustering established an increased tumour mutational burden compared with other tumours of the given cancer type of which 50% or greater of all somatic mutations were attributed to mismatch repair or microsatellite instability mutational signatures as indicated from signature decomposition analysis (see below). POLE ultra-mutated tumours 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 tumour mutational burdens by cancer type as described above, with at least 50% of the mutations attributable to the *POLE*-associated mutational signature. Tumours 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 temozolomide<sup>36</sup>. Hypermutated tumours 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 tumour purity and ploidy using FACETS (v.0.5.6, <http://github.com/mskcc/facets>)<sup>37</sup>. Each tumour 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 (P.S. et al., manuscript in preparation). A subset of tumour-normal pairs was manually reviewed for fit accuracy. Tumour samples with a concordance less than 60% between observed and expected mutant allele frequencies for heterozygous single-nucleotide polymorphisms in the tumour in regions of total copy number less than or equal to three were excluded from zygosity analyses. For tumours for which FACETS was unable to estimate a tumour 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 tumour sample that qualified for zygosity analysis based on these criteria.

**Zygosity of germline and somatic variants.** The tumour-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 whether the observed VAF for clonal events in the tumour was consistent with the expected VAF given the tumour 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, in which the tumour purity is  $\Phi$ , and  $N$  and  $n$  are the locus-specific and allele-specific copy number, respectively. For variants with allelic imbalance in favour of the mutant allele, LOH was considered present if the somatic VAF was within the 95% binomial confidence interval of the expected VAF of the lesser allele having a copy number of zero. LOH favouring the mutant allele was those variants for which the observed VAF exceeded the lower bound of the 95% confidence interval of the expected VAF, and conversely, loss of the mutant (favouring 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 tumour VAF was estimated, the variant was homozygous in the normal, or the tumour depth at the variant site was less than 50.

Although specimens of low tumour cell content will limit the sensitivity of LOH inference based on somatic VAFs, we previously estimated that low tumour cell content affects sensitivity of clinical sequencing with MSK-IMPACT in less than 8.5% of all tumours profiled<sup>19</sup>. Moreover, we routinely called LOH in *BRCA1/2*-mutant tumour specimens with tumour cell content <20–30% (Extended Data Fig. 2a, b). Indeed, although tumour cell content will affect the somatic VAF of germline variants, this had no effect on the sensitivity for LOH detection, and only those somatic mutations detected in samples of tumour cell content less than approximately 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 tumour cell content more than 30% nor is tumour purity a statistically significant predictor of LOH for somatic mutations in such cases. Nevertheless, we cannot formally exclude the possibility that some heterozygous mutant tumours 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 tumours via monoallelic promoter methylation, we analysed 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 tumours (see below and Fig. 3).

**Enrichment of LOH above background.** To assess the selective pressure for LOH leading to somatic wild-type 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 tumour 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 wild-type patients by cancer type and that zygosity changes in tumours of patients with these benign variants selected for loss of the wild-type 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 modelled neutral selection with this background distribution, we proceeded to test for the enrichment for biallelic inactivation (loss of wild type) 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 (P.S. et al., manuscript in preparation). We constructed a similar background distribution of zygosity changes spanning somatic *BRCA1/2* VUSs (all variants that are not designated as LoF here) and tested for the enrichment of biallelic inactivation for somatic LoF *BRCA1/2* mutations in the same manner. Multiple hypothesis correction was performed using the Bonferroni method.

**Estimating the clonality of somatic mutations.** For all somatic *BRCA1* and *BRCA2* mutations, we estimated clonality in each affected tumour specimen (fraction of tumour cells harbouring the indicated mutation) as described previously<sup>38,39</sup>. In brief, we inferred the cancer cell fraction for all mutations using the mutant allele fraction, locus-specific read coverage, and an analytical estimate of tumour purity using a binomial distribution and maximum likelihood estimation to generate posterior probabilities. A cancer cell fraction 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  $\mu = (\text{VAF}/\Phi) \times (N \times \Phi + 2 \times (1 - \Phi))$ , in which  $n_E = 1$  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% confidence interval of its cancer cell fraction 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 tumours, respectively, as well as 522 *BRCA1/2* wild-type tumours. Although 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 was 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 was captured by hybridization using either the SureSelectXT Human All Exon V4 (Agilent 5190-4632) or xGen Exome Research Panel v1.0 (IDT) according to the 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 100-bp or 125-bp paired-end run using the HiSeq 3000/4000 SBS Kit or HiSeq Rapid SBS Kit v2 (Illumina) or on NovaSeq 6000 in a 100-bp paired-end run using the NovaSeq 6000 SBS v1 Kit and an S2 flow cell (Illumina).

In brief, demultiplexed FASTQ files were trimmed of adaptors and short reads with TrimGalore (v.0.2.5mod) and aligned to the b37 assembly of the human reference genome with BWA mem v.0.7.5a. Read groups were annotated and PCR duplicates were marked with Picard Tools v.2.9. Indel realignment was performed using the Assembly Based ReAligner (ABRA) v.2.12<sup>40</sup> and base quality recalibration was performed with GATK v.3.3-0<sup>41</sup>. Somatic mutations (point mutations and small insertions and deletions) were identified in tumour-normal pairs using MuTect v.1.1.4<sup>42</sup> and Vardict v.1.5.1<sup>43</sup>. 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>)<sup>44,45</sup> or a known non-truncating oncogenic variant per OncoKB literature curation<sup>33</sup>. 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](http://hgdownload.cse.ucsc.edu/goldenPath/hg19/database/rmsk.txt.gz) and [genome.ucsc.edu/cgi-bin/hgFileUi?db=wgEncodeMapability](http://genome.ucsc.edu/cgi-bin/hgFileUi?db=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 tumour or greater than 3 reads in the matched normal, tumour and matched normal coverage less than 20 and 10, respectively). Any mutation flagged by any of these filters in three or more independent tumour samples was 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 tumour 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  $P(\text{SSF}) > 0.05$ ; (2) with MQ < 60 and SSF > 0.01; (3) spanning repetitive regions of length greater than 10; or (4) were 1-bp 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 artefacts 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 whole-exome sequencing samples and those MSK-IMPACT sequenced samples with five or more such mutations. The fraction of mutations attributable to each of 30 known mutational signatures<sup>46</sup> 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 phenotype<sup>13</sup> at the level of DNA copy number alterations. Scores for large-scale transitions, number of telomeric allelic imbalances, and HRD deficiency (HRD-LOH) were inferred as previously described<sup>13</sup> using total or allele-specific copy number segmentation data from FACETS analysis of exome and targeted sequencing data (Supplementary Table 6). Owing 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 large-scale transitions among DNA copy number changes. We first standardized each of the individual scores into a Z-score (mean-centred 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 ( $\rho = 0.89$ ,  $P = 10^{-270}$ ).

To establish a control population of presumed homologous-recombination-intact tumours for comparison, we inferred the mutational signatures of HRD in tumours that excluded any case with a germline or somatic mutation in an expanded set of genes presumed to be homologous-recombination effector genes: *ATM*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CHEK1*, *CHEK2*, *FANCA*, *FANCC*, *MRE11A*, *NBN*, *PALB2*, *RAD50*, *RAD51*, *RAD51B*, *RAD51C*, *RAD51D*, *RAD52* and *RAD54L*, and all serous ovarian cancers given work indicating that the majority of such tumours harbour evidence of HRD<sup>47</sup>. 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 aetiology.

**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](http://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<sup>48,49</sup>. 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 set<sup>49</sup>. Germ-cell tumours were excluded from analysis owing to the dynamic nature of DNA methylation changes in primordial germ cells that are not observed in somatic tissues.

**Outcome analyses.** Clinical end points were determined as follows. For PARP inhibitor therapy, outcomes were the percentage of patients on treatment relative to treatment duration. For assessing clinical benefit of immune checkpoint blockade therapy, the clinical end point 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.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data 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 are available as Extended Data.

## Code availability

Source code for these analyses is available at <https://github.com/taylor-lab/BRCA>.

28. Cheng, D. T. 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).
29. Coombs, C. C. 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 (2017).
30. Karczewski, K. J. et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. Preprint at <https://www.bioRxiv.org/content/10.1101/531210v2> (2019).
31. Spurdle, A. B. 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).
32. Findlay, G. M. et al. Accurate classification of *BRCA1* variants with saturation genome editing. *Nature* **562**, 217–222 (2018).
33. Chakravarty, D. et al. OncoKB: a precision oncology knowledge base. *JCO Precis. Oncol.* <https://doi.org/10.1200/PO.17.00011> (2017).
34. Niu, B. et al. MSIsensor: microsatellite instability detection using paired tumor-normal sequence data. *Bioinformatics* **30**, 1015–1016 (2014).
35. Middha, S. et al. Reliable pan-cancer microsatellite instability assessment by using targeted next-generation sequencing data. *JCO Precis. Oncol.* <https://doi.org/10.1200/PO.17.00084> (2017).
36. Johnson, B. E. et al. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science* **343**, 189–193 (2014).
37. Shen, R. & Seshan, V. E. FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res.* **44**, e131 (2016).
38. Bielski, C. M. et al. Genome doubling shapes the evolution and prognosis of advanced cancers. *Nat. Genet.* **50**, 1189–1195 (2018).
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, L. E., Wilkerson, M. D., Hayes, D. N., Perou, C. M. & Parker, J. S. ABRA: improved coding indel detection via assembly-based realignment. *Bioinformatics* **30**, 2813–2815 (2014).
41. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
42. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
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).
44. Chang, M. T. et al. Accelerating discovery of functional mutant alleles in cancer. *Cancer Discov.* **8**, 174–183 (2018).
45. Chang, M. T. et al. Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. *Nat. Biotechnol.* **34**, 155–163 (2016).
46. Alexandrov, L. B. et al. Clock-like mutational processes in human somatic cells. *Nat. Genet.* **47**, 1402–1407 (2015).
47. Wang, Y. K. et al. Genomic consequences of aberrant DNA repair mechanisms stratify ovarian cancer histotypes. *Nat. Genet.* **49**, 856–865 (2017).
48. Huang, K.-L. et al. Pathogenic germline variants in 10,389 adult cancers. *Cell* **173**, 355–370 (2018).
49. Knijnenburg, T. A. et al. Genomic and molecular landscape of DNA damage repair deficiency across the cancer genome atlas. *Cell Rep.* **23**, 239–254 (2018).

**Acknowledgements** We thank our patients and their families for participating in this study and the members of Taylor laboratory 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 (NIH) awards P30 CA008748, U54 OD020355 (D.B.S., B.S.T.), R01 CA207244 (D.M.H., B.S.T.), P50 CA092629 (D.B.S., 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.).

**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, and 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. L.Z. reports receiving honoraria from Future Technology Research LLC, Roche Diagnostics Asia Pacific, BGI, and Illumina. L.Z. has a family member with a leadership position and ownership interest in Shanghai Genome Center. J.B. is an employee of AstraZeneca, serves on the Board of Directors of Foghorn and is a past board member of Varian Medical Systems, Bristol-Myers Squibb, Grail, Aura Biosciences and Infinity Pharmaceuticals. He has performed consulting and/or advisory work for Grail, PMV Pharma, ApoGen, Juno, Lilly, Seragon, Novartis and Northern Biologics. He

has stock or other ownership interests in Tango and Venthera, for which he is a co-founder. He has previously received honoraria or travel expenses from Roche, Novartis, and Lilly. 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, and Loxo Oncology and personal fees from Atara Biotherapeutics, Chugai Pharma, Boehringer Ingelheim, AstraZeneca, Pfizer, Bayer, Debiopharm Group, and Genetech. 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 Intezyne. All stated activities were outside of the work described herein. No other disclosures were noted.

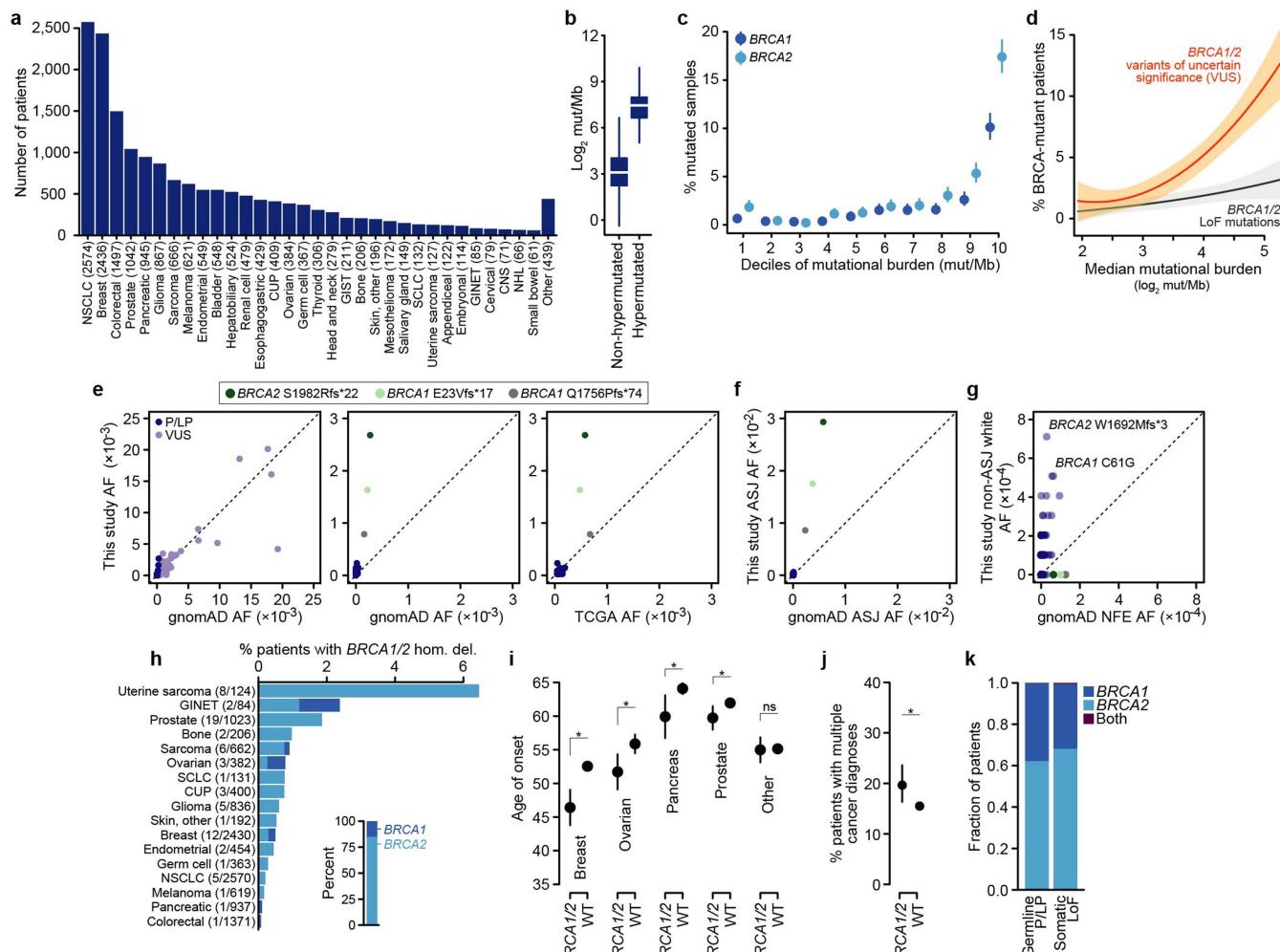
#### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1382-1>.

**Correspondence and requests for materials** should be addressed to M.F.B., D.B.S. or B.S.T.

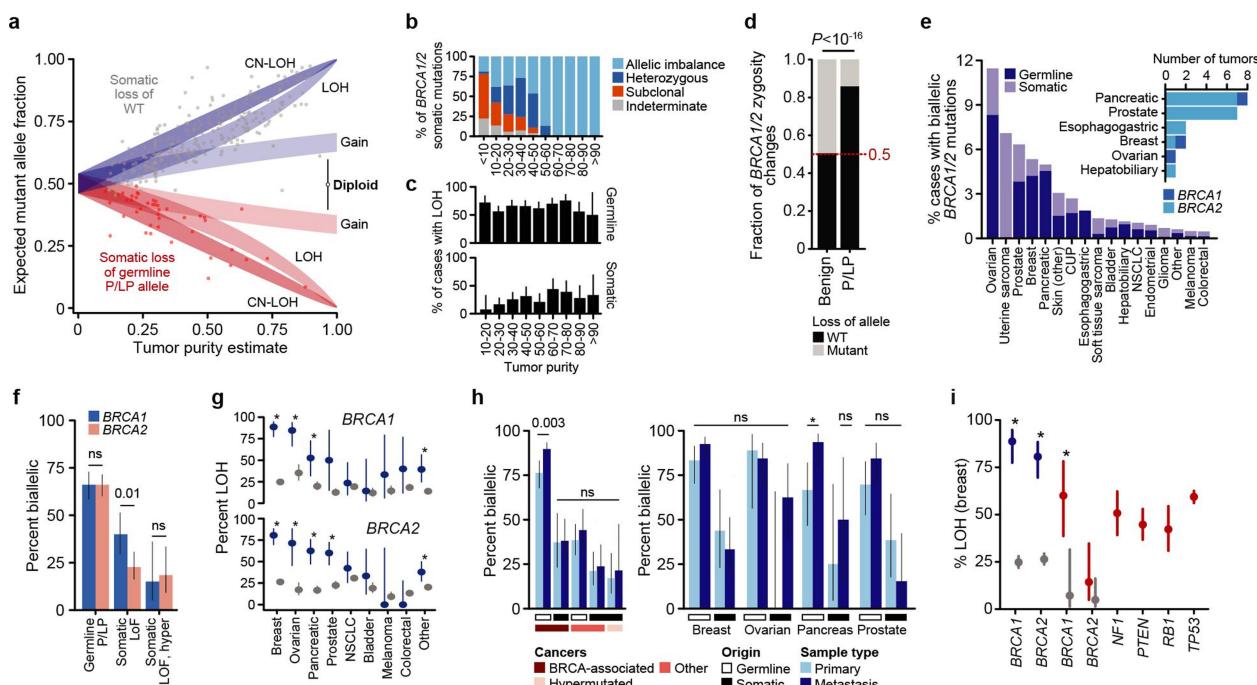
**Peer review information** *Nature* thanks Clare Turnbull and the other anonymous reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



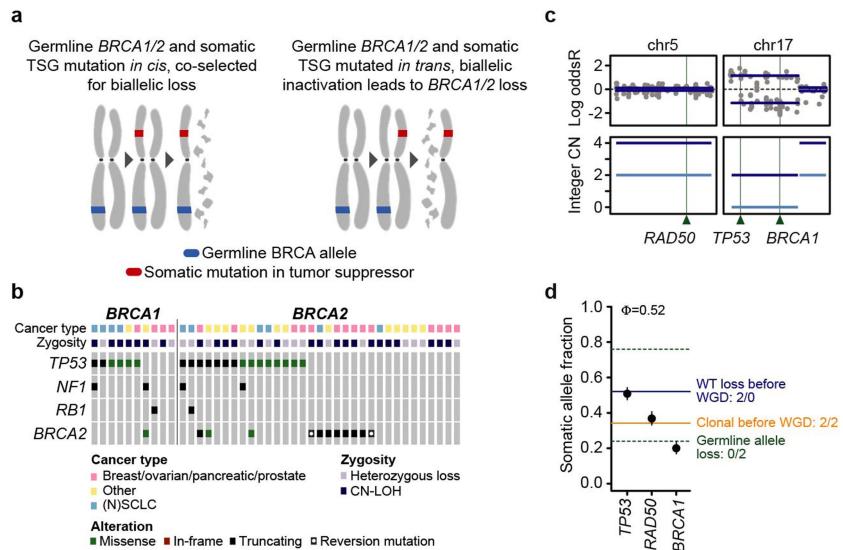
**Extended Data Fig. 1 | Study cohort and BRCA1/2 germline and somatic mutation distribution.** **a**, The number of tumour and matched normal specimens are shown by cancer type. CNS, non-glioma central nervous system tumours; CUP, cancer of unknown primary; GINET, gastrointestinal neuroendocrine tumour; GIST, gastrointestinal stromal tumour; NHL, non-Hodgkin's lymphoma; NSCLC, non-small-cell lung cancer; SCLC, small-cell lung cancer. **b**, Somatic mutational burden ( $\log_2(\text{mutations per megabase})$ ) in tumours defined as non-hyperpermuted or hyperpermuted (via microsatellite instability, DNA polymerase epsilon mutations, or alkylating therapy-induced; see Methods). Data are shown as median and interquartile range. **c**, *BRCA1* and *BRCA2* somatic mutation rates in deciles of increasing tumour mutational burden. The highest mutational burden tumours also had the highest rate of *BRCA1/2* mutations. Error bars are binomial confidence intervals. **d**, The percentage of tumours in each tumour type containing either somatic VUS or LoF *BRCA1/2* mutations as a function of the median somatic mutational burden of that cancer type (excluding hyperpermuted cases). Overall, the rate of somatic LoF *BRCA1/2* mutations by cancer type modestly increased with their increasing mutational burden, and 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; centre, only P/LP alleles; right, comparison between the study cohort and the germline results from the TCGA cohort. Ashkenazi Jewish (ASJ) founder *BRCA1/2* alleles are shown. **f**, As in **e** for only the ASJ sub-populations. **g**, As in **e** and **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 tumours 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 wild type) in *BRCA*-associated cancer types and all other cancer types. Error bars are 95% confidence intervals. \* $P < 0.01$ , two-sided Wilcoxon test. **j**, Percentage of *BRCA1/2* germline carriers with several independent cancer diagnoses compared to germline wild-type patients. Error bars are 95% confidence intervals. \* $P = 0.02$ ,  $\chi^2$  test. **k**, The fraction of all germline pathogenic or somatic LoF alterations in *BRCA1* versus *BRCA2* (in non-hyperpermuted tumours).



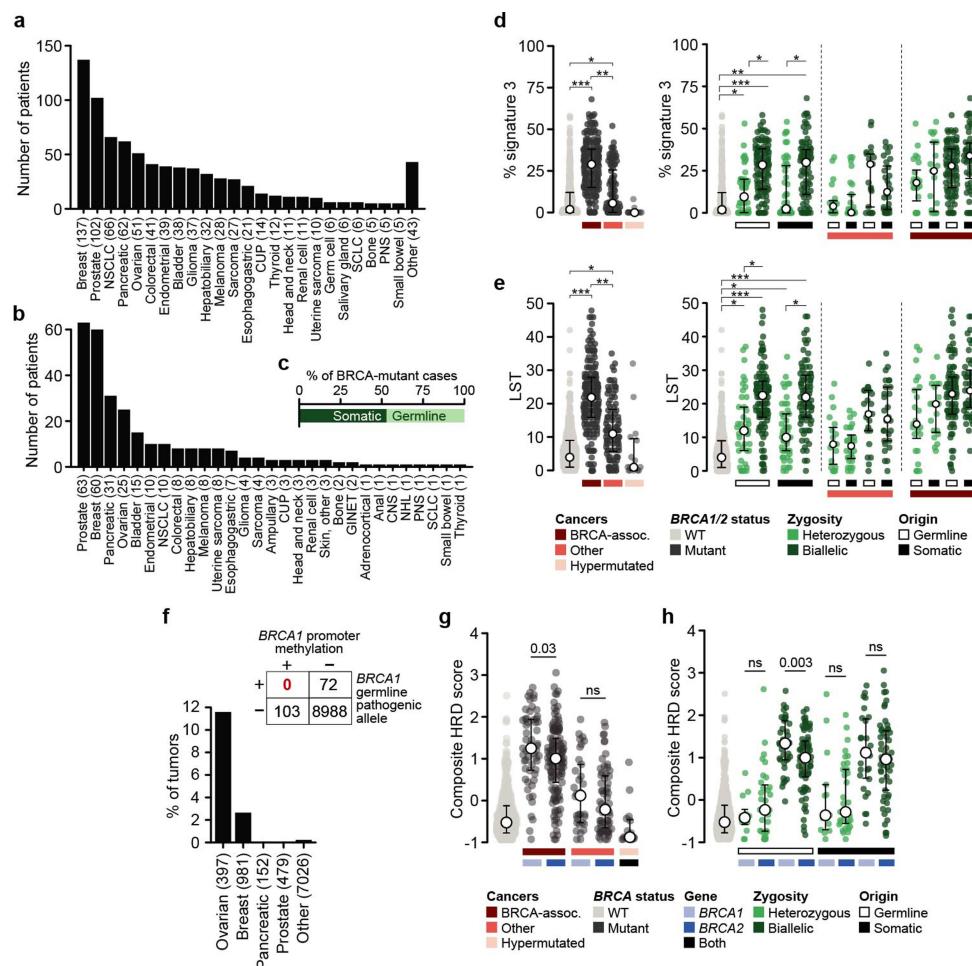
**Extended Data Fig. 2 | BRCA1/2 zygosity.** **a**, Diagrammatic representation of the integration of allele-specific copy number with purity and mutant allele frequencies to determine the zygosity of the germline pathogenic allele in the corresponding tumour (and the mechanism of its selection). CN-LOH, copy-neutral LOH. **b**, In only a subset of cases of low tumour cell content (<30%) does the LOH inference become increasingly analytically challenging (increasing rate of indeterminant calls). **c**, The percentage of cases with LOH affecting the germline pathogenic or somatic LoF BRCA1/2 mutations (as labelled) as a function of tumour purity (see Methods). Although somatic mutant allele frequencies are affected by tumour purity, this does not affect the sensitivity for LOH detection for germline variants and only affects sensitivity for LOH of somatic mutations in tumours of less than 30% purity. Error bars are 95% confidence intervals in all panels. **d**, In tumours with benign germline variants in BRCA1 and BRCA2, the ratio of zygosity changes affecting the wild-type or mutant BRCA1/2 allele is approximately 0.5, indicating neutral selection. By contrast, the rate of zygosity changes leading to loss of the wild-type 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 wild-type BRCA1/2), the percentage of tumours by cancer type containing 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 hyperpermuted and non-hyperpermuted tumours). *P* values determined by two-sided Fisher's exact test. **g**, The rate of loss of wild-type BRCA1 or BRCA2 (LOH) in patients with germline deleterious BRCA1 or BRCA2 mutations compared with rare benign variants in either gene in BRCA-associated cancer types and in those not conventionally associated with BRCA germline carriers. *P* values determined by 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. Right, the four BRCA-associated cancer types are shown individually. *P* values determined by two-sided Fisher's exact test. ns, not significant. **i**, The rate of LOH spanning germline or somatic mutant BRCA1 and BRCA2 in breast cancers (coloured as in Fig. 2b, c) as well as other somatically mutated tumour-suppressor genes.



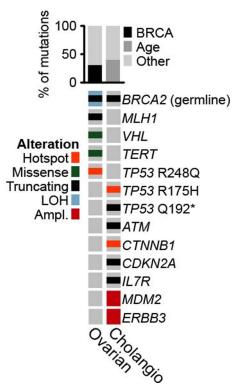
**Extended Data Fig. 3 | Somatic loss of the pathogenic germline *BRCA1* or *BRCA2* allele.** **a**, Schematic representation of the different allelic configurations that would lead to the retention or loss of a germline allele in the presence of a somatically mutated tumour-suppressor gene (TSG) responsible for driving biallelic inactivation. **b**, Among tumours 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, whereas *RB1* is encoded on chromosome 13 on which *BRCA2* also appears) arising in the same tumours 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 **c**) are consistent with deletion of the allele containing the *BRCA1* founder mutation as compared to the observed and expected values for clonal heterozygous somatic mutations (*RAD50*) or biallelic inactivation of mutant *TP53* (tumour purity is  $\phi$ ). The selective pressure for biallelic *TP53* inactivation driven by the initial R248Q mutation probably precipitated the subsequent heterozygous loss of the wild-type *TP53* allele, leading to deletion of the *BRCA1* pathogenic mutation and retention of the wild-type *BRCA1* allele, indicating that mutant *BRCA1* was dispensable for its pathogenesis. Error bars are binomial confidence intervals.

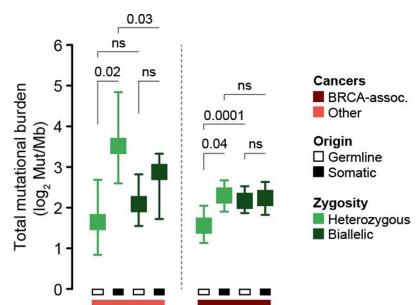


**Extended Data Fig. 4 | HRD phenotype in BRCA1/2-mutant cancers characterized by whole-exome sequencing.** **a**, Total number of prospectively sequenced cases by cancer type for which exome re-sequencing was obtained. PNS, peripheral nervous system; other abbreviations are as in Extended Data Fig. 1. **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, e**, The somatic single-nucleotide mutational signature 3 of HRD (**d**), and the DNA copy number-based large-scale transitions metric of HRD as inferred from exome sequencing data (**e**) are shown as a function affected cancer types (left) and BRCA1/2 mutational origin and zygosity (right) as in Fig. 3. \* $P < 0.01$ , \*\* $P < 1 \times 10^{-10}$ , \*\*\* $P < 1 \times 10^{-20}$ , two-sided Student's  $t$ -test. Circles and horizontal lines denote median and lower and upper quartiles, respectively. The individual metrics are highly correlated with the composite HRD score ( $\rho = 0.89$ ,  $P = 1 \times 10^{-270}$ ; see Methods) and consequently the qualitative results based on lineage, mutational origin and zygosity are similar. **f**, The rate of BRCA1 promoter

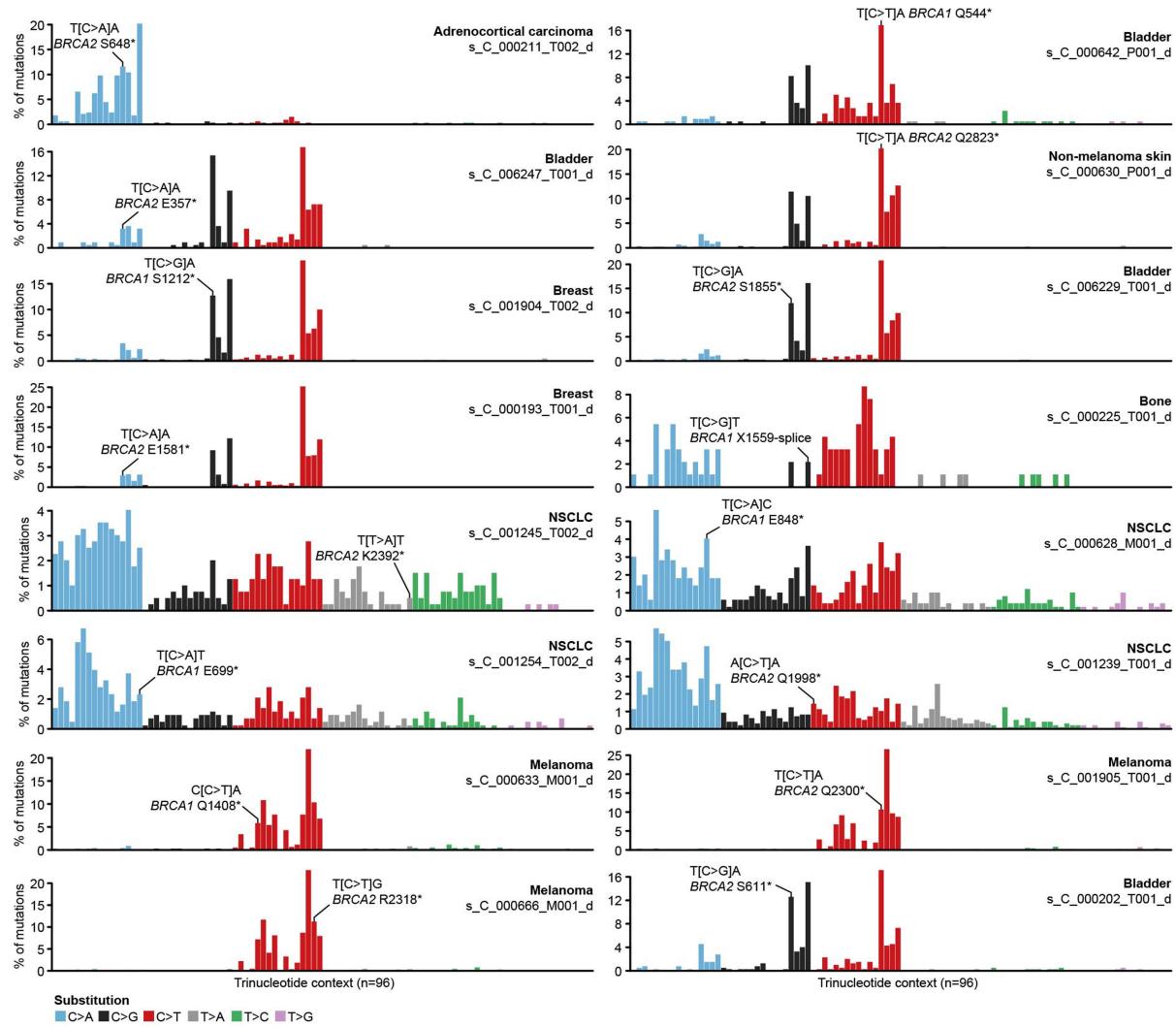
methylation in ovarian, breast and other cancer types (no evidence of BRCA2 silencing via promoter methylation was apparent). Inset, BRCA1 germline mutations and promoter methylation leading to BRCA1 silencing are mutually exclusive in affected cancers, indicating that heterozygous BRCA1-mutant tumours 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 tumours (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 homologous-recombination-wild-type tumours (light grey) and in tumours with either germline or somatic BRCA1 or BRCA2 mutations (dark grey) grouped by BRCA-associated cancer types (dark red; breast, ovary, pancreas and prostate) versus other cancer types (red), and tumours with somatic hypermutation (light red).  $P$  values determined by two-sided Student's  $t$ -test. **h**, As in **g** and Fig. 3c, grouped by a combination of lineage, origin and zygosity.


**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 cancers—one a HRD serous ovarian cancer (left) with loss of wild-type *BRCA2*; the other a co-incident cholangiocarcinoma with intact wild-type *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.

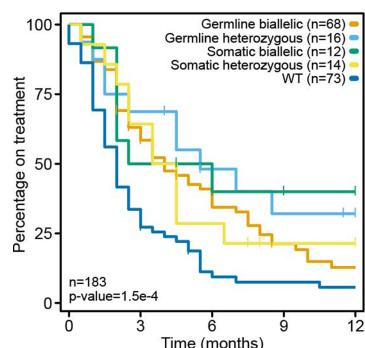


**Extended Data Fig. 6 | Tumour mutational burden by *BRCA1/2* genotype.** The somatic mutational burden of tumours as a function of cancer type, *BRCA1/2* mutational origin, and somatic *BRCA1/2* zygosity. *P* values determined by two-sided Student's *t*-test. Boxes and error bars denote median and 95% confidence intervals, respectively.



**Extended Data Fig. 7 | Mutations in *BRCA1* or *BRCA2* are attributable to other mutational signatures.** The somatic mutations in each of the evaluable cancers in 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* mutational 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 Fig. 4). All four classes of *BRCA*-mutant patients (as indicated) achieve significantly greater clinical benefit to PARP inhibitor therapy than do treated patients with wild-type *BRCA* tumours (*BRCA1/2*-mutant classes: germline carrier somatic heterozygous (hazard ratio = 0.39, 0.21–0.72,  $P = 0.003$ ); germline carrier somatic biallelic (hazard ratio = 0.5, 0.35–0.72,  $P = 2 \times 10^{-4}$ ); somatic heterozygous LoF (hazard ratio = 0.5, 0.26–0.95,  $P = 0.03$ ); and somatic LoF biallelic (hazard ratio = 0.34, 0.16–0.72,  $P = 0.005$ )).

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

## Software and code

Policy information about [availability of computer code](#)

Data collection

No code used for data collection.

Data analysis

Variant Effect Predictor v88 (variant annotation); vcf2maf v1.6.10 (variant annotation); FACETS v0.5.6 (copy-number analysis); TrimGalore v0.2.5mod (adaptor trimming); Picard tools v2.9 (sequencing read PCR de-duplication); ABRA v2.12 (local insertion/deletion re-alignment); GATK v3.3-0 (base quality recalibration); MuTect v1.1.4 (somatic SNV calling); Vardict v1.1.5 (somatic SNV and indel calling); and [www.github.com/mskcc/mutation-signatures](https://github.com/mskcc/mutation-signatures) (mutational signature inference). All other analyses were performed using the R environment for statistical computing and all code will be publically available via <https://github.com/taylor-lab/BRCA>.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The whole-exome sequencing data as well as germline variant calls have been deposited in the NCBI dbGaP archive under accession numbers phs001783.v1.p1 and phsXXXXXX, respectively. All other genomic and clinical data accompanies the manuscript and is available as Extended or Supplementary Data.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences     Behavioural & social sciences     Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was not pre-determined; all available specimens were utilized.
Data exclusions	No exclusion criteria were specified for the study population.
Replication	No experimental replication was performed.
Randomization	No randomization of data was performed.
Blinding	No blinding of data was performed.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies
<input type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology
<input checked="" type="checkbox"/>	Animals and other organisms
<input type="checkbox"/>	Human research participants
<input checked="" type="checkbox"/>	Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	Not applicable
Authentication	Not applicable
Mycoplasma contamination	Not applicable
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	Not applicable

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	Age at first diagnosis: median 57 years Male/female: 47%/53% Extended data available in extended data table 2.
Recruitment	Patients who underwent prospective sequencing as part of their clinical care at Memorial Sloan Kettering Cancer Center (MSKCC) from February 2014 to July 2017. All such patients whose tumor sequencing was performed with a matched normal sample were included.
Ethics oversight	MSKCC Institutional Review Board

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	NCT01775072
Study protocol	Not applicable.
Data collection	Not applicable.
Outcomes	Not applicable.