



HAL
open science

Whole-Genome Sequencing Reveals Breast Cancers with Mismatch Repair Deficiency

Helen Davies, Sandro Morganella, Colin A. Purdie, Se Jin Jang, Elin Borgen, Hege Russnes, Dominik Glodzik, Xueqing Zou, Alain Viari, Andrea L. Richardson, et al.

► **To cite this version:**

Helen Davies, Sandro Morganella, Colin A. Purdie, Se Jin Jang, Elin Borgen, et al.. Whole-Genome Sequencing Reveals Breast Cancers with Mismatch Repair Deficiency. *Cancer Research*, 2017, 77 (18), pp.4755-4762. 10.1158/0008-5472.CAN-17-1083. hal-01599736

HAL Id: hal-01599736

<https://inria.hal.science/hal-01599736>

Submitted on 2 Oct 2017

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

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

Whole-Genome Sequencing Reveals Breast Cancers with Mismatch Repair Deficiency

Helen Davies¹, Sandro Morganella¹, Colin A. Purdie², Se Jin Jang³, Elin Borgen⁴, Hege Russnes^{4,5,6}, Dominik Glodzik¹, Xueqing Zou¹, Alain Viari^{7,8}, Andrea L. Richardson^{9,10}, Anne-Lise Børresen-Dale^{4,5}, Alastair Thompson¹¹, Jorunn E. Eyfjord¹², Gu Kong¹³, Michael R. Stratton¹, and Serena Nik-Zainal^{1,14}

Abstract

Mismatch repair (MMR)-deficient cancers have been discovered to be highly responsive to immune therapies such as PD-1 checkpoint blockade, making their definition in patients, where they may be relatively rare, paramount for treatment decisions. In this study, we utilized patterns of mutagenesis known as mutational signatures, which are imprints of the mutagenic processes associated with MMR deficiency, to identify MMR-deficient breast tumors from a whole-genome sequencing dataset comprising a cohort of 640 patients. We identified 11 of 640 tumors as MMR deficient, but only 2 of 11 exhibited germline mutations in MMR genes or Lynch Syndrome. Two additional tumors had a substantially reduced proportion of mutations attributed to MMR deficiency, where the predominant mutational signatures

were related to APOBEC enzymatic activity. Overall, 6 of 11 of the MMR-deficient cases in this cohort were confirmed genetically or epigenetically as having abrogation of MMR genes. However, IHC analysis of MMR-related proteins revealed all but one of 10 samples available for testing as MMR deficient. Thus, the mutational signatures more faithfully reported MMR deficiency than sequencing of MMR genes, because they represent a direct pathophysiologic readout of repair pathway abnormalities. As whole-genome sequencing continues to become more affordable, it could be used to expose individually abnormal tumors in tissue types where MMR deficiency has been rarely detected, but also rarely sought. *Cancer Res*; 77(18); 4755–62. ©2017 AACR.

Introduction

Mismatch repair (MMR) deficiency in human cancer is associated with heterozygous germline mutations in components of MMR machinery (*MSH2*, *MSH6*, *MLH1*, *PMS2* – otherwise

known as Lynch syndrome) or due to somatic mutations in these genes and/or promoter hypermethylation of *MLH1* (1–9). These classical tumor suppressors lose the wild-type allele in the tumor, resulting in biallelic gene inactivation. The MMR pathway plays a pivotal role in reducing error rates during replication, thus MMR deficiency is associated with high levels of mutagenesis—10- to 100-fold more mutations than tumors that have intact MMR pathways (9, 10).

Recognition of MMR-deficient tumors is clinically important. MMR-deficient colorectal cancers, although associated with poor prognoses and intense lymphocytic infiltration, have better clinical responses to therapies that interfere with T-cell immune checkpoints (11–14), particularly of PD-1/PD-L1 proteins. In a phase II clinical trial (15), an objective clinical response to anti-PD1 treatment was demonstrable in MMR-deficient tumors (4/10 colorectal and 5/7 of other tissue-types) in contrast to 0 of 18 of MMR-proficient tumors. This was recently reinforced by an expanded study evaluating anti-PD-1 efficacy across 12 tumor types (14). Sensitivity to immune therapy in MMR-deficient tumors is believed to be mediated through T-cell reactivity against "neoantigens"—novel epitopes that are formed as a consequence of the excessive somatic mutations in these cancers (16). Intriguingly, this phenomenon is not restricted to colon cancer, thus identification of all MMR-deficient tumors, regardless of tissue-type could be beneficial (17).

Diagnostic methods for identification of MMR-deficient tumors have evolved over the last 25 years. Histopathologic criteria preceded PCR-based testing of microsatellite loci, before

¹Wellcome Trust Sanger Institute, Hinxton, United Kingdom. ²Pathology Department, Ninewells Hospital and Medical School, Dundee, United Kingdom. ³Department of Pathology, Asan Medical Center, College of Medicine, Ulsan University, Ulsan, South Korea. ⁴Department of Pathology, Oslo University Hospital, Oslo, Norway. ⁵Department of Cancer Genetics, Institute for Cancer Research, Norwegian Radium Hospital, Oslo University Hospital, Oslo, Norway. ⁶K.G. Jebsen Centre for Breast Cancer Research, Institute for Clinical Medicine, University of Oslo, Oslo, Norway. ⁷Equipe Erable, INRIA Grenoble-Rhône-Alpes, Montbonnot-Saint Martin, France. ⁸Synergie Lyon Cancer, Centre Léon Bérard, Lyon, France. ⁹Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts. ¹⁰Dana-Farber Cancer Institute, Boston, Massachusetts. ¹¹Department of Breast Surgical Oncology, University of Texas MD Anderson Cancer Center, Houston, Texas. ¹²Cancer Research Laboratory, Faculty of Medicine, University of Iceland, Reykjavik, Iceland. ¹³Department of Pathology, College of Medicine, Hanyang University, Seoul, Republic of Korea. ¹⁴East Anglian Medical Genetics Service, Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom.

Note: Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

Corresponding Author: Serena Nik-Zainal, Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, United Kingdom. Phone: 44-0-1223-834244; Fax: 44-0-1223-494919; E-mail: snz@sanger.ac.uk

IHC for loss of MMR protein expression became routine in clinical diagnostics. Sequencing of DNA repair genes and methylation assays are also used to confirm somatic and/or germline pathogenic mutations. However, somatic MMR deficiency is only sought routinely in cancers where it is frequent, like colorectal (~20%) and endometrial (~20%–30%) cancer. Clinical testing is not customarily performed where it is less common [e.g., cervical cancer (8%), or breast cancer (0%–2%); refs. 12, 15]. Indeed, the rare occurrence of MMR-deficient breast tumors has been attributed to germline mutations in Lynch Syndrome, and frequently assessed in that context although the true relationship between MMR-deficient breast tumors and inherited predisposition remains controversial (18).

Modern cancer resequencing experiments, particularly whole-genome sequencing (WGS) approaches, have revealed much more than the handful of driver mutations that are causally implicated in carcinogenesis. They have also exposed the many thousands of passenger mutations that are the products of the mutational processes that have occurred throughout tumorigenesis (19–22). Each mutational process leaves its own characteristic imprint or mutational signature on the genome (23), an indicator of past and ongoing exposures, whether of environmental insults such as ultraviolet radiation, or of endogenous biochemical degradation and deficiencies of DNA repair pathways like MMR deficiency.

Previously, substitution signatures 6, 20, and 26 (among others) were associated with MMR deficiency in colorectal, stomach, and uterine carcinomas (<http://cancer.sanger.ac.uk/cosmic/signatures>). Here, we find that eleven of 640 breast tumors (1.7%) in a WGS study show variable quantities of these substitution signatures. Critically, this subset of tumors would not have been detected as being MMR-deficient using current clinical criteria for assessment of breast cancer. We describe their distinctive and highly individual, pathognomonic genomic profiles, confirm the diagnosis using genetic, epigenetic, and protein IHC methods, and emphasize how genome profiling using mutational signatures could be a powerful additional tool for tumor stratification.

Material and Methods

Detailed methods are provided in Supplementary Materials and Methods.

Accession codes

The overarching accession number for the data on 560 breast cancers previously described in Nik-Zainal and colleagues is EGAS00001001178 (24). For the 80 additional breast cancers described in Davies and colleagues (25), data are deposited under the following EGA accession numbers; bam files in EGAD00001002740 and SNP6 Cel files in EGAD000010001079.

Dataset

Internal review boards of each participating institution approved collection and use of samples of all patients in this study. Informed consent was obtained by the relevant participating institution. Ethical guidelines were in accordance with the Declaration of Helsinki.

DNA was extracted from 640 breast cancer cases along with corresponding normal tissue and subjected to WGS as described previously (24, 25). Resulting BAM files were aligned to the

reference human genome (GRCh37) using Burrows-Wheeler Aligner, BWA (v0.5.9).

Mutation calling was performed as described previously (24). Briefly, CaVEMan (Cancer Variants through Expectation Maximization: <http://cancerit.github.io/CaVEMan/>) was used for calling somatic substitutions. Indels in the tumor and normal genomes were called using a modified Pindel version 2.0 (<http://cancerit.github.io/cgppindel/>) on the NCBI37 genome build. Structural variants were discovered using a bespoke algorithm, BRASS (BReakpoint AnalySiS; <https://github.com/cancerit/BRASS>) through discordantly mapping paired-end reads followed by *de novo* local assembly using Velvet to determine exact coordinates and features of breakpoint junction sequence.

In total, 3,808,160 somatic base substitutions, 399,466 small indels, and 83,191 rearrangements were detected in the 640 samples (24, 25).

Mutational signatures

Mutation signature analysis based on nonnegative matrix factorization (NMF) was performed as described previously (22, 24). Twelve base consensus substitution signatures were identified in the 560 breast whole genomes: signatures 1, 2, 3, 5, 6, 8, 13, 17, 18, 20, 26, and 30 in that analysis where a primary signature extraction using NMF was followed by reintroduction of the 12 signatures mentioned above as currently reported on COSMIC (<http://cancer.sanger.ac.uk/cosmic/signatures>). In the description of previously published data in this manuscript, we have utilized the mutational signatures that were extracted and assigned previously using Alexandrov's method. The activity of these twelve signatures was estimated in a further 80 breast cancer whole genomes (25).

In all new analyses, we did not have access to the Alexandrov method. To maintain consistency, all new analyses used the same new iterative algorithm to identify the set of COSMIC signatures [cancer.sanger.ac.uk/cosmic/signatures] active in each sample (the so called "exposure"; SigFit, Morganello and colleagues manuscript submitted). Each sample was completely described by a vector containing the number of substitutions observed for each mutation and flanking sequence context (defined by the neighboring bases immediately 5' and 3' to the mutated base and by the mutated base itself). Each mutation was orientated with respect to the pyrimidine strand and consequently each vector contained 96 elements. The algorithm started from an initial solution estimated by using a simulated annealing-based method. Then mutations were iteratively reassigned to alternative signatures, cosine similarities were obtained through comparing the reconstructed 96-element vector for each potential reassignment to that of the observed 96-element vector with the aim of identifying the highest possible cosine similarity value. The algorithm stopped when no improvement to the cosine similarity was found. We ran the algorithm controlling the sparsity of the solution (number of signatures simultaneously active in each sample). In particular, we used three different values for the parameter alpha: 0, 0.01, and 0.02 (the sparsity of the solution increases with alpha). In addition, we assessed the consistency of the results for the breast data using two predefined sets of signatures: one containing all thirty COSMIC signatures and one containing only the subset of 12 signatures that were previously associated with breast cancer. As expected different results were obtained for different sets of *a priori* signatures provided (12 or 30).

However, results for different settings of alpha were largely consistent with minor variation, reflecting the degree of constraint between different alpha (Supplementary Table S1) within each set of *a priori* signatures. To maintain consistency in all downstream analyses, the setting of alpha 0.01 was used for all analyses shown in this publication.

Two indel signatures based on the presence of either short tandem repeats or short stretches of identical sequence at the breakpoints (termed overlapping microhomology), were also extracted.

ASCAT copy number analysis

Single-nucleotide polymorphism (SNP) array hybridization using the Affymetrix SNP6.0 platform was performed according to Affymetrix protocols. Allele-specific copy number analysis of tumors was performed using ASCAT (v2.1.1) as performed as described previously (24). ASCAT takes non-neoplastic cellular infiltration and overall tumor ploidy into consideration, to generate integer-based allele-specific copy number profiles for the tumor cells. ASCAT was also applied to next-generation sequencing data directly with highly comparable results. Copy number values and estimates of aberrant tumor cell fraction provided by ASCAT are in put into the CaVEMan substitution algorithm. In addition, ASCAT segmentation profiles were used to establish the presence of loss of heterozygosity across MMR genes and to search for homozygous deletion of these genes.

Detection of variants in MMR genes

We sought to discover both germline and somatic mutations of all classes in the following genes involved in MMR; *MLH1*, *MSH2*, *MSH6*, *PMS2*, *PMS1*, *SETD2*, *MYH11*, *EPCAM*, *TGFBR2*, *MLH3*, and in *MUTYH* in the 640 breast cancer cases. Single base substitutions and small insertions/deletions were interrogated using Caveman and Pindel algorithms, respectively. While large deletions were investigated using a combination of ASCAT copy number data and rearrangement calls by BRASS. Variants affecting the coding regions of these genes were verified by visual inspection to remove common sequencing artefacts and cross-referenced against dbSNP and ClinVar (<http://www.ncbi.nlm.nih.gov/clinvar/>) to identify benign polymorphisms. Variants with evidence in both the tumor and corresponding non-neoplastic tissue were deemed to be germline, while those restricted to the tumor were somatic. ASCAT copy number data was used to determine whether there had been loss of the alternative allele in the tumor sample.

IHC

IHC staining of formalin-fixed paraffin-embedded sections was performed by the center responsible for recruiting the patient. Staining for the *MLH1*, *PMS2*, *MHS2*, and *MSH6* proteins was performed according to standard clinical diagnostic lab procedures for the relevant center.

Methylation analyses for identification of hypermethylated genes

DNA methylation of cancer samples were assayed by Infinium Human Methylation 450k Beadchips (Illumina) according to the manufacturers' protocols. Analysis was performed as described previously (24).

Detection of MMR-deficient cancers in other datasets

See Supplementary Materials and Methods for more details of the analysis of an MMR-deficient Ovarian cancer sample and of whole-exome sequence (WES) breast cancer samples.

Results

Genomic profiles of presumptive MMR-deficient breast tumors

Substitution signature analysis of 640 WGS breast tumors (24, 25) revealed eleven samples (1.7%) with variable quantities of substitution signatures 6, 20, and 26 (Supplementary Table S2; Supplementary Fig. S1). In addition, all eleven had strikingly high numbers of insertions/deletions (indels); orders of magnitude more than other breast cancers (mean of 20,870, range 2,535–66,764 in these 11 compared with mean of 270, range 19–1,512 in other 629 breast cancers). These indels occurred at polynucleotide repeat tracts consistent with the pattern of microsatellite instability associated with MMR deficiency. Altogether, these eleven tumors had a constellation of genomic features that were quite distinct from those identified in the other 629 breast cancers (Fig. 1) and are described in more detail below (Fig. 2).

Of the 11 patients, 5 were ER-positive and 6 were ER-negative tumors. Nine of 11 patients had breast cancers that were hypermutated by MMR substitution signatures, reflected in the absolute number (range 9,232–84,432, average 31,166) and proportions (79%–100%) of Signature 6, 20, and 26 mutations (Fig. 2). The two remaining patients had a large number of mutations attributed to the activity of APOBEC cytidine deaminases in addition to MMR-related signatures. PD5937a had 34% of mutations attributed to MMR signatures while PD23561a had only 13.7% of MMR-related mutations. These two samples thus showed a greater degree of heterogeneity of mutational signatures. It is likely that the predominance of the APOBEC-related signature is indicative of earlier onset of this mutational process in the tumor's clonal evolution, prior to acquiring the MMR signature. See Supplementary Fig. S2 for genome plots of all 11 samples.

Genetic/epigenetic and IHC evidence supporting MMR deficiency in the breast tumors

To support our hypothesis that these genomic profiles were indicative of MMR deficiency, we searched for genetic and/or epigenetic evidence of MMR inactivation in MMR-related genes (*MLH1*, *MSH2*, *MSH6*, *PMS2*, *PMS1*, *SETD2*, *MYH11*, *EPCAM*, *TGFBR2*, *MLH3*) and in *MUTYH*. Inactivating mutations were confirmed in 6 of 11 cancers (55%). Of these 6 cases, two patients had somatic *MLH1* mutations, one had a heterozygous germline *MLH1* truncating mutation, and one had a *PMS2* germline heterozygous Lynch Syndrome mutation (c.137G>T p.S46I, rs121434629; ref. 26). All four tumors had loss of the alternative parental allele making these genes null at the cellular level. The fifth patient had promoter hypermethylation of *MLH1* to a level consistent with complete inactivation of the gene. The remaining tumor had somatic compound heterozygous mutations in *MSH2*. Therefore, genetic/epigenetic confirmation was not achieved in 5 of the 11 (45%) putatively MMR-deficient tumors, although methylation data were not available for these tumors (See Supplementary Table S3 for details of the mutations).

To obtain alternative, corroborative evidence of MMR deficiency, IHC for MMR proteins was performed for 10 of the 11 cases (see Fig. 3 for examples). Reassuringly, 9 of 10 cancer samples

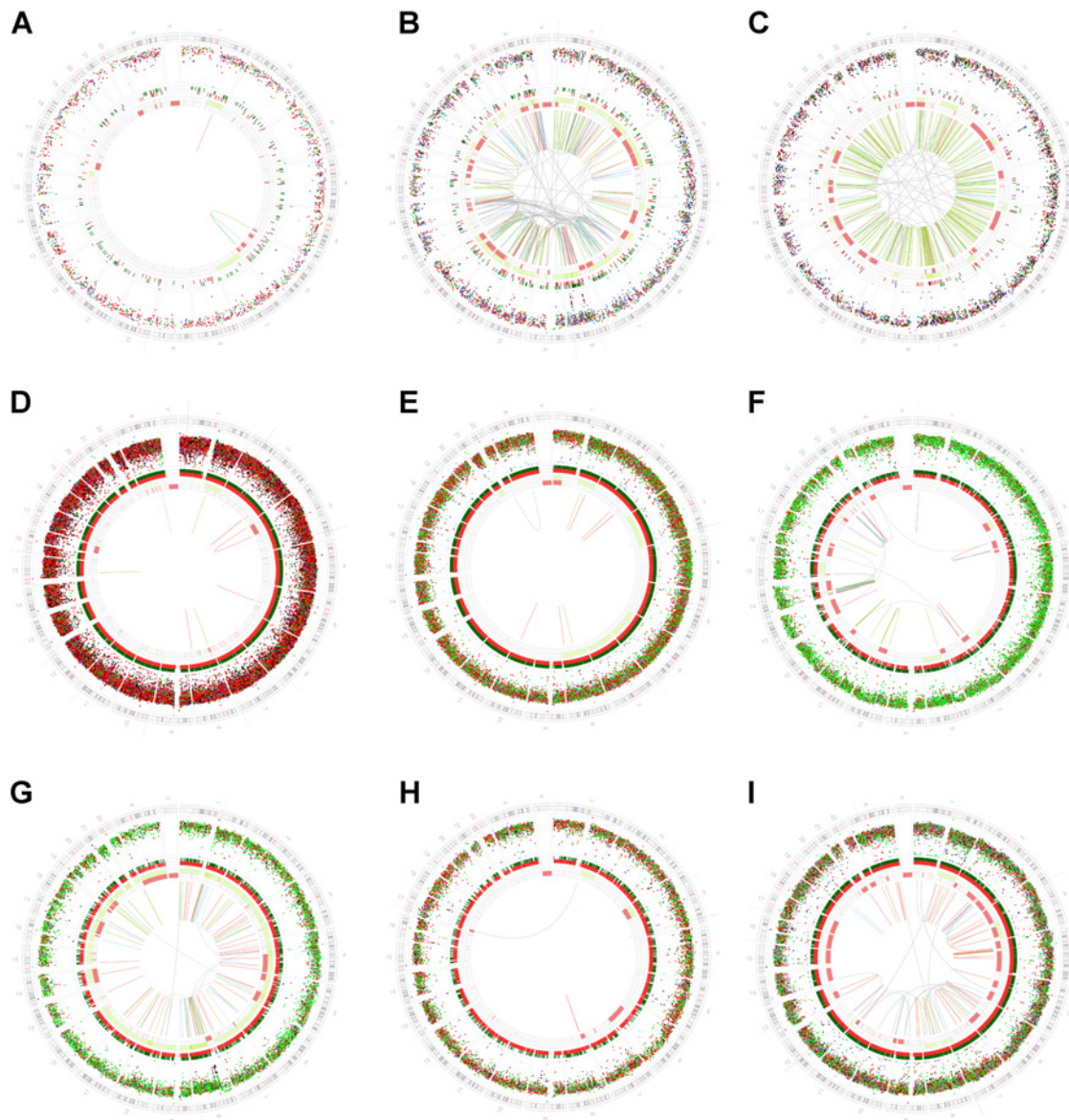


Figure 1.

MMR-deficient breast tumors have highly distinctive whole genome profiles. Whole genome circos plot depicting from outermost rings heading inwards; karyotypic ideogram outermost. Base substitutions next, plotted as rainfall plots (log10 intermutation distance on radial axis; dot colors: blue, C>A; black, C>G; red, C>T; gray, T>A; green, T>C; pink, T>G). Ring with short green lines, insertions; ring with short red lines, deletions. Major copy number allele ring, green, gain; minor copy number allele ring, red, loss. Central lines represent rearrangements (green, tandem duplications; red, deletions; blue, inversions; gray, interchromosomal events). Appearance of MMR-deficient tumors (**D-I**) are in stark contrast to other MMR-proficient tumors. **A**, A typical estrogen receptor-positive breast cancer. **B**, A typical triple-negative breast cancer (or tumors with deficiency of other repair pathways such as homologous recombination-mediated double-strand break repair). **C**, A breast cancer from a *BRCA1* germline mutation carrier. **D**, PD5937a, *MSH2*-deficient cancer also with a large number of APOBEC related mutations. **E** and **F**, PD24189a (**E**) and PD24320a (**F**), MMR-deficient cancers with underlying mutation unknown. **G**, **H**, and **I**, PD9604a (**G**), PD24193a (**H**), and PD6412a (**I**), MMR-deficient cancers with *MLH1* mutations.

were confirmed to exhibit concomitant deficits of *MLH1* and *PMS2* (7), or *MSH2* and *MSH6* (2) IHC staining, in-keeping with expectations in clinical diagnostics of MMR-deficient tumors (Fig. 2). The tenth sample (PD23561a), however, was the exception. This sample was the particular tumor that demonstrated only a small contribution of MMR signatures (13.7% of substitutions). It is therefore possible that the IHC result is a false negative because

MMR deficiency is restricted to only a subpopulation of cells in this tumor. These results emphasize two things. First, knowledge of the precise MMR driver mutation is not essential as the mutational signatures appear to be a dependable read-out of MMR abrogation. Second, for MMR deficiency that has arisen later in cancer evolution or is present only at subclonal levels, genomic signature profiling may be a more faithful reporter than

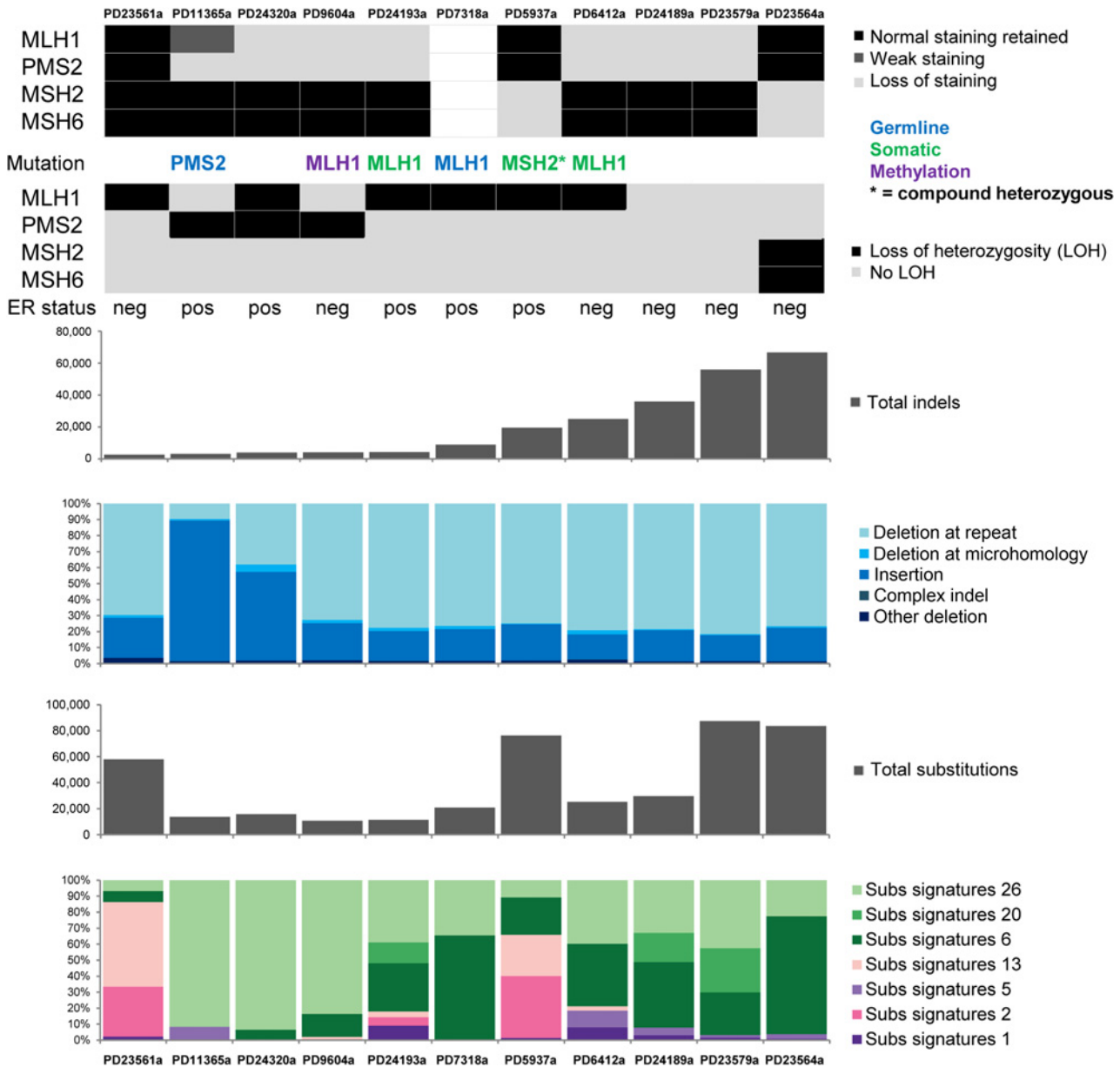


Figure 2.

Genomic characteristics of MMR-deficient breast tumors. From top to bottom, panel showing IHC staining results for 10 of the 11 MMR-deficient cancers. Gray, loss of staining; black, normal staining pattern retained. IHC results were not available for PD7318a. Relevant mutations identified are indicated at bottom. Next, panel showing loss of heterozygosity (LOH) across *MLH1*, *PMS2*, *MSH2*, and *MSH6*. Estrogen receptor (ER) status of each tumor, either negative (neg) or positive (pos) is indicated next. The four histograms below show the total number of indels mutations per sample, the proportion of each indel type present per sample, the total number of substitution mutations, and the proportion of substitution signatures 26, 20, 6, 13, 5, 2, and 1 present in each sample.

IHC, outperforming prevailing protein-based methods used in clinical practice today.

Notably, monoallelic MMR mutations did not result in MMR signatures as somatic *PMS2* (PD22361a), germline *PMS2* (PD13608a), germline *MSH6* (PD18006a), and germline *MLH3* (PD5936a) mutations were observed in the rest of the breast cancer dataset but did not yield genomic features of MMR abrogation. Therefore, biallelic inactivation appears to be a prerequisite for a tumor to exhibit MMR deficiency.

Mutational signatures of MMR deficiency are associated with specific gene defects

In addition, correlations between mutational signatures and particular MMR defects are observed. *MLH1*-inactivated breast cancers had combinations of predominantly mutation types C>T/G>A and T>C/A>G transitions (classified as Signatures 6 and 26, respectively) with overwhelming indel mutagenesis, particularly deletions at polynucleotide repeat tracts. In contrast, *PMS2* inactivation appears to be enriched mainly for

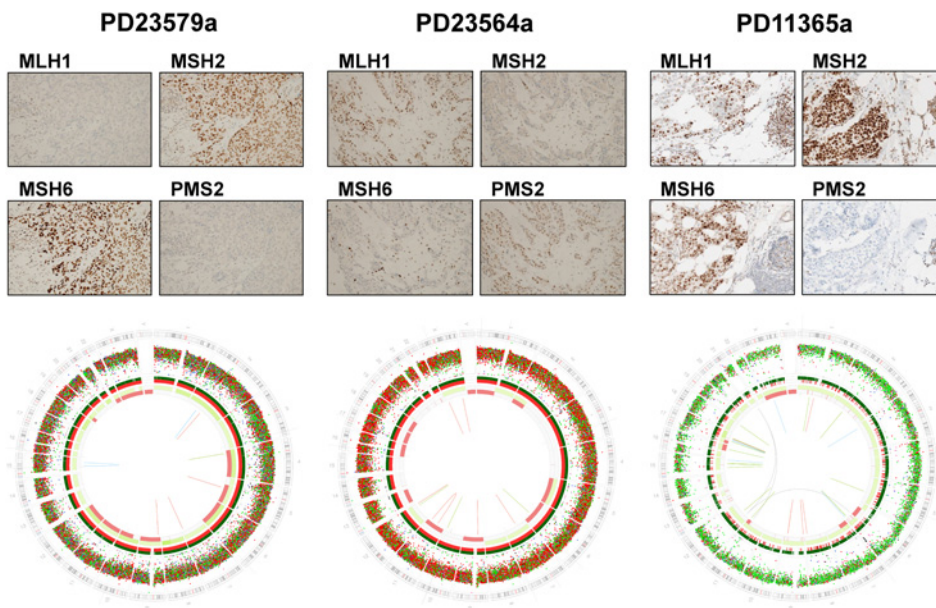


Figure 3.

IHC and corresponding genome profiles of MMR-deficient breast tumors. Top sections: IHC staining of tumor sections for MLH1, MSH2, MSH6, and PMS2 proteins. From left to right, PD23579a shows concomitant loss of MLH1 and PMS2 staining. PD23564a shows concomitant loss of MSH2 and MSH6 staining. No mutations were detected in these tumors. PD11365a shows loss of PMS2 staining and reduced staining for MLH1, consistent with the germline *PMS2* mutation present in this patient. Normal tissue staining included in the right-hand side of the frame. Bottom, genome plots of mutations identified by whole-genome sequencing. Corresponding circo plots depicted below IHC images (description in legend of Fig. 1).

T>C/A>G transitions with some contribution of T>G/A>C transversions (classified as Signature 26 in breast cancer previously) and insertions at polynucleotide repeat tracts. This observation was upheld in another cancer type - where a germline *PMS2* Lynch Syndrome patient, carrying the identical c.137G>T p.S46I mutation, had a strikingly similar mutation profile and enrichment of insertions (Supplementary Fig. S3; ref. 27) in an ovarian cancer. The field of mutational signatures is still in its infancy and there is some variability in the extractions of mutational signatures and their assignments in different tumor types. Signature 26 for example, had not previously been described in ovarian tumors. To avoid bias, we performed an unsupervised mutational signature extraction and found probable contributions from signatures 9 and 12 (of the 30 mutation signatures currently available at COSMIC; <http://cancer.sanger.ac.uk/cosmic/signatures>) rather than signature 26 in both the tumors with the germline *PMS2* mutations (Supplementary Table S1 and Supplementary information for more details). Indeed, inspection of the 96-element mutational profiles of these two tumors suggests that this alternative result may be more suitable than the previous conclusion of Signature 26. However, the numbers are small and the field is still developing; thus further investigations are warranted to draw definitive conclusions. Regardless of what the final attributed "signature/s" will be, there are early hints at associations between mutational signatures and specific genetic defects, at least within the mismatch repair pathway.

Population frequency estimates of MMR deficiency in breast cancer

To investigate whether MMR deficiency could be detected using alternative sequencing strategies in breast cancer, we restricted the analysis to mutations that fell within coding sequences, mimicking a WES experiment. Although MMR signatures were detectable in all of the 11 cases with MMR deficiency, in the absence of any *post hoc* filtering, many other samples were erroneously assigned MMR signatures as well (Supplementary Table S4). Therefore, we applied strict criteria to classifying MMR-deficient tumors (Sup-

plementary for details) on a new pool of data, of approximately 1,100 WES breast cancers (<http://cancer.sanger.ac.uk/cosmic>). We found that approximately 1% of tumors (14 of 1,097) could be classified as MMR-deficient. In contrast, an alternative method, which utilizes WES data to analyze instability at microsatellite loci (28) on a subset overlapping this dataset, did not detect any MMR-deficient breast cancers. WES-based classification is almost certainly conservative (Supplementary Fig. S4; Supplementary Table S4) although here, helps to provide an approximation of population frequency. These 14 of 1,097 samples were not hitherto highlighted as MMR-deficient in mutational signature analyses (21), partially due to lack of previous understanding. Advances in analysis and insights gained from the vast bank of WGS breast cancers sequenced more recently, permit an appraisal of old data, with new awareness.

Discussion

Two important questions arise from this study. First, what is the clinical significance of these observations? The frequency of MMR-deficient breast cancers in this dataset is approximately 1%–2%. Most MMR-deficient breast tumors are not due to Lynch syndrome, as only 2 of the 11 cases carried a germline mutation in an MMR gene. Forty-five percent of cases would not have been identified through sequencing of DNA repair genes in the tumor or in the germline but were instantly recognizable as MMR deficient through mutational signature profiling, increasing the yield of tumors that could have selective sensitivity to specific immune-therapy treatments. Second, is WGS really necessary to detect MMR-deficient tumors? Where the increase in mutational burden is so great, exome sequencing alone could expose tumors where MMR deficiency has arisen early and is clonal, although we demonstrate that there is a reduced specificity using WES within breast cancer. Furthermore, WGS may still be a necessity to detect MMR deficiency that has arisen only in subclones (20), precisely the subpopulations that are potentially selectively targetable, but could go unseen and untreated and may possibly be the source of future recurrence.

Finally, the dramatic improvements in targeted therapeutics on cancer outcomes have not necessarily come from new drugs. Indeed, reappraisal of existing drugs combined with fine-tuned stratification for more specific application has sometimes been at the heart of improved outcomes. Two clinical trials recently investigated sensitivity to PD-L1 checkpoint inhibitors, and used transcriptomic measures of PD-L1 as a biomarker (29). Cutoffs applied in these two trials were startlingly different: Opdivo (nivolumab, CheckMate-026 trial, Bristol-Myers Squibb) used a cutoff of PD-L1 expression exceeding 5% of cells, while Keytruda (pembrolizumab, Keynote-024, Merck) used a 50% cutoff to categorize tumors. Progression-free survival outcomes were contrasting between the two studies and speculated to be due to the choice of biomarker and cutoff, conferring a lack of specificity in tumor classification. Perhaps the application of mutational signatures could make a difference here. Our analyses suggest that because mutational signatures are a direct pathophysiologic read-out of MMR pathway abrogation, they confer a higher degree of sensitivity and specificity for tumor classification, and could possibly out-perform current biomarkers of MMR deficiency.

Disclosure of Potential Conflicts of Interest

H. Davies has ownership interest (including patents) in applied patents. D. Glodzick has ownership interest (including patents) in applied patents. S. Nik-Zainal has ownership interest (including patents) in a patent application and is a consultant/advisory board member for Artios Pharma Ltd. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: H. Davies, A.L. Richardson, A.-L. Børresen-Dale, A. Thompson, M. Stratton, S. Nik-Zainal

Development of methodology: S. Morganella, D. Glodzick, S. Nik-Zainal
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): C.A. Purdie, S.J. Jang, E. Borgen, H.G. Russnes, A. Viari, A.L. Richardson, A.-L. Børresen-Dale, A. Thompson, J.E. Eyfjord, G. Kong, S. Nik-Zainal
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): H. Davies, S. Morganella, D. Glodzick, X. Zou, A.L. Richardson, S. Nik-Zainal
Writing, review, and/or revision of the manuscript: H. Davies, S. Morganella, C.A. Purdie, S.J. Jang, E. Borgen, A.L. Richardson, A.-L. Børresen-Dale, A. Thompson, J.E. Eyfjord, S. Nik-Zainal
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): C.A. Purdie, S. Nik-Zainal
Study supervision: S. Nik-Zainal

Acknowledgments

This work has been performed on data that was previously published. Key Principal Investigators whose vision led to this very fruitful collaboration include Mike Stratton, Alain Viari, Gu Kong, Henk Stenzenberg, Marc van der Vijver, Ewan Birney, Ake Borg, and Anne-Lise Børresen-Dale.

Grant Support

S. Nik-Zainal and X. Zou were supported by the Wellcome Trust via a Wellcome Trust Strategic Award (WT101126/B/13/Z), S. Nik-Zainal is personally funded by a Cancer Research UK Advanced Clinician Scientist Award (C60100/A23916). A. L. Richardson is partially supported by the Dana-Farber/Harvard Cancer Center SPORE in Breast Cancer (NIH/NCI5 P50 CA168504-02). G. Kong is supported by National Research Foundation of Korea (NRF) grants funded by the Korean government (NRF2015R1A2A1A10052578).

References

- Lynch HT, Snyder CL, Shaw TG, Heinen CD, Hitchins MP. Milestones of Lynch syndrome: 1895–2015. *Nat Rev Cancer* 2015;15:181–94.
- Geurts-Giele WR, Leenen CH, Dubbink HJ, Meijssen IC, Post E, Sladdens HF, et al. Somatic aberrations of mismatch repair genes as a cause of microsatellite-unstable cancers. *J Pathol* 2014;234:548–59.
- Bronner CE, Baker SM, Morrison PT, Warren G, Smith LG, Lescoe MK, et al. Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature* 1994;368:258–61.
- Gazzoli I, Loda M, Garber J, Syngal S, Kolodner RD. A hereditary non-polyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. *Cancer Res* 2002;62:3925–8.
- Leach FS, Nicolaidis NC, Papadopoulos N, Liu B, Jen J, Parsons R, et al. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell* 1993;75:1215–25.
- Miyaki M, Konishi M, Tanaka K, Kikuchi-Yanoshita R, Muraoka M, Yasuno M, et al. Germline mutation of MSH6 as the cause of hereditary non-polyposis colorectal cancer. *Nat Genet* 1997;17:271–2.
- Nicolaidis NC, Papadopoulos N, Liu B, Wei YF, Carter KC, Ruben SM, et al. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature* 1994;371:75–80.
- Papadopoulos N, Nicolaidis NC, Wei YF, Ruben SM, Carter KC, Rosen CA, et al. Mutation of a mutL homolog in hereditary colon cancer. *Science* 1994;263:1625–9.
- Parsons R, Li GM, Longley M, Modrich P, Liu B, Berk T, et al. Mismatch repair deficiency in phenotypically normal human cells. *Science* 1995;268:738–40.
- Parsons R, Li GM, Longley MJ, Fang WH, Papadopoulos N, Jen J, et al. Hypermutability and mismatch repair deficiency in RER+ tumor cells. *Cell* 1993;75:1227–36.
- Llosa NJ, Cruise M, Tam A, Wicks EC, Hechenbleikner EM, Taube JM, et al. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. *Cancer Discov* 2015;5:43–51.
- Dudley JC, Lin MT, Le DT, Eshleman JR. Microsatellite instability as a biomarker for PD-1 blockade. *Clin Cancer Res* 2016;22:813–20.
- Lee V, Murphy A, Le DT, Diaz LA Jr. Mismatch repair deficiency and response to immune checkpoint blockade. *Oncologist* 2016;21:1200–11.
- Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, et al. Mismatch-repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 2017;357:409–13.
- Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med* 2015;372:2509–20.
- Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Science* 2015;348:69–74.
- Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 2015;348:124–8.
- Win AK, Lindor NM, Jenkins MA. Risk of breast cancer in Lynch syndrome: a systematic review. *Breast Cancer Res* 2013;15:R27.
- Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, et al. Mutational processes molding the genomes of 21 breast cancers. *Cell* 2012;149:979–93.
- Nik-Zainal S, Van Loo P, Wedge DC, Alexandrov LB, Greenman CD, Lau KW, et al. The life history of 21 breast cancers. *Cell* 2012;149:994–1007.
- Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. *Nature* 2013;500:415–21.
- Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR. Deciphering signatures of mutational processes operative in human cancer. *Cell Rep* 2013;3:246–59.

23. Helleday T, Eshtad S, Nik-Zainal S. Mechanisms underlying mutational signatures in human cancers. *Nat Rev Genet* 2014;15: 585–98.
24. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 2016;534:47–54.
25. Davies HR, Glodzik DG, Morganella S, Yates LR, Staaf J, Zou X, et al. HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational-signatures. *Nat Med* 2017;23:517–25.
26. Auclair J, Leroux D, Desseigne F, Lasset C, Saurin JC, Joly MO, et al. Novel biallelic mutations in MSH6 and PMS2 genes: gene conversion as a likely cause of PMS2 gene inactivation. *Hum Mutat* 2007;28: 1084–90.
27. Patch AM, Christie EL, Etemadmoghadam D, Garsed DW, George J, Fereday S, et al. Whole-genome characterization of chemoresistant ovarian cancer. *Nature* 2015;521:489–94.
28. Hause RJ, Pritchard CC, Shendure J, Salipante SJ. Classification and characterization of microsatellite instability across 18 cancer types. *Nat Med* 2016;22:1342–50.
29. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med* 2012;366:2443–54.

