

ARTICLE OPEN



Clinical relevance of pathogenic germline variants in mismatch repair genes in Chinese breast cancer patients

Li Hu^{1,4}, Jie Sun^{1,4}, Zhongwu Li^{2,4}, Ziwei Qu³, Yan Liu³, Qiting Wan¹, Jiaming Liu¹, Xinyun Ding¹, Fan Zang¹, Juan Zhang¹, Lu Yao¹, Ye Xu¹, Yin Wang³ and Yuntao Xie¹

The prevalence and clinical relevance of pathogenic germline variants in MMR genes have not been investigated in large series of breast cancers. In this study, we screened the germline variants in MMR genes in 8085 consecutive Chinese breast cancer patients, and investigated the MMR/PD-L1 protein expression and tumor mutation burden (TMB) of breast tumors from MMR variant carriers. We found that 15 of 8085 patients (0.19%) carried a pathogenic germline variant in MMR genes. Compared with non-carriers, MMR variant carriers might have worse recurrence-free survival (unadjusted hazard ratios [HR] = 2.70, 95% CI: 1.12–6.49, $P = 0.027$) and distant recurrence-free survival (unadjusted HR = 3.24, 95% CI: 1.45–7.22, $P = 0.004$). More importantly, some of the breast cancers from MMR carriers displayed MMR protein loss (5/13), TMB-high (2/10), and PD-L1 positive expression (9/13). This study showed that MMR variant carriers were rare in breast cancer. They might have worse survival and part of them might benefit from immunotherapy.

npj Breast Cancer (2022)8:52; <https://doi.org/10.1038/s41523-022-00417-x>

INTRODUCTION

Pathogenic germline variants in four mismatch repair (MMR) genes (MLH1, MSH2, MSH6, and PMS2) have been found to result in hereditary nonpolyposis colorectal carcinoma (HNPCC, also called Lynch syndrome), which is an autosomal dominant inherited disease primarily associated with colorectal, endometrial, gastric, small intestinal, hepatobiliary, renal pelvic, and ureteral cancers¹. Currently, colorectal cancers from MMR variant carriers are associated with poor differentiation, extensive lymphocytic infiltration and a superior survival compared with sporadic cases^{2,3}. More importantly, dysfunction of MMR genes in HNPCC results in MMR protein loss, and induces errors of DNA polymerase during DNA replication, causing the accumulation of mutations in the genome, especially in the repetitive DNA sequences called microsatellites⁴. As a consequence, HNPCCs typically exhibit loss of MMR protein expression, microsatellite instability (MSI) and high tumor mutation burden (TMB), which make them sensitive to immunotherapy^{5,6}.

A recent clinical trial has highlighted the significant responses of solid cancers with MMR deficiency to immunotherapy, regardless of the cancer type⁷, which caused considerable interest in investigating MMR in various cancer types. Whether breast cancer can be recognized as part of the spectrum of HNPCC has been debated. A previous study concluded that breast cancer is not associated with HNPCC⁸. However, multiple studies reported that women from HNPCC families had an increased risk of breast cancer ranging from 2- to 4-fold compared with the general population^{9–14}, and some of the breast cancers from HNPCC families exhibited MMR protein loss or MSI^{15–21}. These results suggested a potential role for MMR deficiency in breast tumorigenesis. Notably, these previous studies focused on breast cancers from HNPCC families, while MMR genes have not been investigated in a large series of consecutive breast cancers.

In this study, we aimed to investigate the frequency of pathogenic germline variants in MMR genes (MLH1, MSH2, MSH6, and PMS2) in a large series of Chinese breast cancer patients. Then we compared the clinicopathological characteristics and survival between MMR variant carriers and non-carriers. Finally, we explored whether these breast cancer patients with germline MMR variants can potentially benefit from immunotherapy.

RESULTS

Prevalence, clinical characteristics and survival of breast cancer patients carrying MMR germline variants

A total of 15 patients in the 8085 Chinese consecutive breast cancer cohort (0.19%) carried a pathogenic germline variant in the four MMR genes, including PMS2 ($n = 6$), MSH6 ($n = 5$), MSH2 ($n = 3$), and MLH1 ($n = 1$) (Table 1). Six patients of the fifteen MMR variant carriers (40.0%) had a personal or family history of HNPCC-related cancers. Gastric cancer ($n = 4$) and colorectal cancer ($n = 3$) were the most common (Table 1). Compared with non-carriers, MMR variant carriers were more likely to have a positive family history of HNPCC-related cancer [40.0% (6/15) vs. 10.9% (882/8070), $p = 0.001$] (Table 2). No difference in age at diagnosis or family history of breast cancer was noted between the MMR variant carriers and non-carriers (Table 2). In terms of the tumor characteristics, most of the breast cancers from MMR variant carriers were invasive ductal carcinomas, while MMR variant carriers might have more medullary [6.7% (1/15) vs. 0.6% (47/8070), $p = 0.08$] and papillary (13.3% (2/15) vs. 0.3% (23/8070), $p < 0.001$) carcinomas compared with non-carriers (Table 2). There was no significant difference in tumor size, grade, lymph node metastasis, or ER/PR/HER2 status between the MMR variant carriers and non-carriers (Table 2). After the median follow-up time of 65.3 months, MMR variant carriers showed significantly worse recurrence-free

¹Familial & Hereditary Cancer Center, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Peking University Cancer Hospital & Institute, 100142 Beijing, P. R. China. ²Department of Pathology, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Peking University Cancer Hospital & Institute, 100142 Beijing, P. R. China. ³Berry Oncology Corporation, 350200 Fujian, P. R. China. ⁴These authors contributed equally: Li Hu, Jie Sun, Zhongwu Li.

✉email: wangyin001@berryoncology.com; zlxty2@bjmu.edu.cn

Table 1. Fifteen breast cancer patients with pathogenic germline variants in MMR genes identified from 8085 breast cancer patients.

Case ID	Germline	Age (y)	Personal history	Family history	Histology	Tumor size	Lymph node	ER status	PR status	HER2 status
P12	PMS2:p.R134X	80	-	-	IDC	1.4*1.2	(-)	(-)	(-)	(-)
P14	PMS2:p.E225X	43	Lymphoma, 55 y	1-1 breast cancer, 42 y; 2-1 ovarian cancer, 27 y; 2-1 gastric cancer, 78 y	IDC	2.5*1.8	(+)	(+)	(+)	(-)
P20	PMS2:p.R315X	38	Cervical cancer, 39 y	-	IDC	2.1*2.0	(+)	(-)	(+)	(-)
P25	PMS2:p.L351X	45	-	1-1 gastric cancer, 45 y; 2-1 gastric cancer, 50 y; 2-1 esophagus cancer; 3-1 esophagus cancer	IDC	5.4*3.8	(+)	(+)	(+)	(-)
P26	PMS2:p.V754fs	34	-	-	IDC	4.6*1.3	(-)	(-)	(-)	(+)
P27	PMS2:p.R151fs	61	-	-	IDC	2.2*2.0	(-)	(+)	(+)	(-)
P5	MSH6:p.R495X	62	-	2-1 pancreatic cancer, 40 y; 2-1 liver cancer, 50 y	IDC	1.1*1.0	(-)	(+)	(+)	(-)
P8	MSH6:p.Y994X	50	-	1-1 breast cancer, 40 y	Invasive papillary	1.7*1.5	(+)	(+)	(+)	(-)
P22	MSH6:p.R922X	57	-	-	IDC	1.7*1.6	(+)	(-)	(-)	(+)
P23	MSH6:p.T716fs	67	-	-	IDC	1.4*1.4	(-)	(+)	(-)	(-)
P31	MSH6:p.T1085fs	53	Endometria cancer, 52 y	1-1 gastric cancer, 60 y; 2-1 gastric cancer, 71 y; 2-1 breast cancer, 50 y	Medullary	2.2*1.5	(-)	(-)	(-)	(-)
P2	MSH2:p.Q4X	54	-	2-1 colorectal cancer, 64 y	IDC	2.4*1.6	(+)	(+)	(+)	(-)
P19	MSH2:p.R877fs	64	-	1-1 brain cancer	IDC	1.6*1.7	(-)	NA	NA	NA
P21	MSH2:p.Q314fs	75	Rectal cancer, 77 y	-	Invasive papillary	2.5*2.2	(-)	(+)	(+)	(-)
P15	MLH1:p.R659X	46	Rectal cancer, 42 y	1-1 rectal cancer, 45 y; gastric cancer, 42 y; 1-1 lung cancer, 47 y	DCIS	2.0*1.8	(-)	(+)	(+)	(-)

NA not accessible, DCIS ductal carcinoma in situ, IDC invasive lobular carcinoma, ILC invasive lobular carcinoma, MMR mismatch repair, ER Estrogen receptor, PR progesterone receptor, HER2 human epidermal growth factor receptor 2.

Table 2. Comparison of clinical characteristics between MMR variant carriers and non-carriers in a large cohort of breast cancer patients.

Characteristic	MMR carriers (n = 15)		Non-carrier (n = 8070)		P value
	N	%	N	%	
Age at diagnosis, years					
Mean ± SD	55.3 ± 13.2		51.1 ± 11.6		0.16
Early-onset					0.88
≤40	2	13.3	1466	18.2	
>40	13	86.7	6604	81.8	
Family history of HNPCC-related cancer					0.001
No	9	60.0	7188	89.1	
Yes	6	40.0	882	10.9	
Family history of breast cancer					0.38
No	12	80.0	7266	90.0	
Yes	3	20.0	804	10.0	
Family history of any cancers					0.39
No	8	53.3	5417	67.1	
Yes	7	46.7	2653	32.9	
Histology					<0.001
Ductal	12	80.0	7195	89.2	
Medullary	1	6.7	47	0.6	
Papillary	2	13.3	23	0.3	
Others	0	0.0	805	10.0	
Tumor size					0.44
≤2 cm	6	40.0	3283	40.7	
>2 cm	7	46.7	4392	54.4	
Unknown	2	13.3	395	4.9	
Tumor grade					0.86
I	0	0.0	625	7.7	
II	10	66.7	4599	57.0	
III	2	13.3	984	12.2	
Unknown	3	20.0	1862	23.1	
Lymph nodes status					0.41
Negative	9	60.0	5514	68.4	
Positive	6	40.0	2084	25.8	
Unknown	0	0.0	472	5.8	
ER status					0.68
Negative	4	26.6	2213	27.4	
Positive	10	66.7	5519	68.4	
Unknown	1	6.7	338	4.2	
PR status					0.84
Negative	4	26.7	2726	33.8	
Positive	10	66.7	4885	60.6	
Unknown	1	6.7	459	5.7	
HER2 status					0.83
Negative	12	80.0	5565	68.9	
Positive	2	13.3	1845	22.9	
Unknown	1	6.7	660	8.2	
Subtype					0.75
ER/PR +, HER2–	10	66.7	5733	71.1	
HER2 +	2	13.3	852	10.5	
ER–, PR–, HER2–	2	13.3	1103	13.7	
Unknown	1	6.7	382	4.7	

P, Carriers vs. Non-carrier.

MMR mismatch repair, HNPCC Hereditary nonpolyposis colorectal cancer, ER estrogen receptor, PR progesterone receptor, HER2 Human epidermal growth factor receptor 2, SD standard deviation.

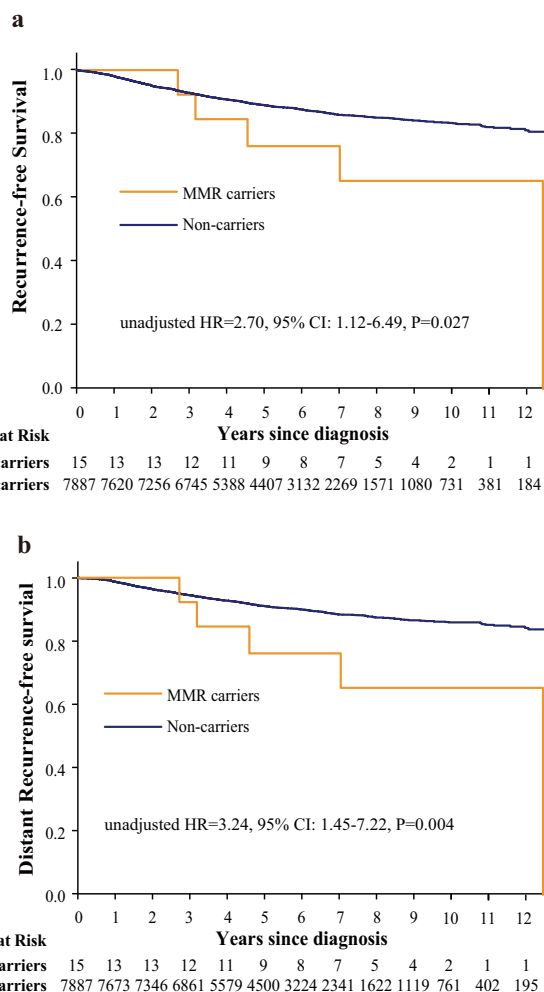


Fig. 1 Kaplan–Meier survival analyses according to germline variants status in MMR genes. **a** and **b** show recurrence-free survival and distant recurrence-free survival, Data are analyzed for statistical significance using univariate Cox proportional hazards models. Two-sided *P* values < 0.05 were considered to be statistically significant.

survival (RFS) (unadjusted HR = 2.70, 95% CI: 1.12–6.49, *P* = 0.027) and distant recurrence-free survival (DRFS) (unadjusted HR = 3.24, 95% CI: 1.45–7.22, *P* = 0.004) than non-carriers (Fig. 1). After adjustment for age, tumor size, lymph node, tumor grade, ER/PR/HER2 status and treatment, pathogenic germline variants in MMR genes had a trend to be associated with worse RFS (adjusted HR = 2.34, 95% CI: 0.97–5.66, *P* = 0.06) and were significantly associated with worse DRFS (adjusted HR = 2.76, 95% CI: 1.14–6.69, *P* = 0.03) in the breast cancer patients in this study (Table 3).

MMR/PD-L1 protein expression in breast cancers from MMR variant carriers

MMR/PD-L1 protein immunostaining was performed on 13 of the 15 breast cancers carrying MMR pathogenic germline variants. Among them, 5 cancers (5/13, 38.5%) showed total loss of at least one MMR protein (Table 4). The losses affected mlh1 and pms2 in one MLH1 variant carrier (P15) (Fig. 2a), msh6 only in two MSH6 variant carriers (P22, P23) (Fig. 2b), and pms2 only in two PMS2 variant carriers (P14, P20) (Fig. 2c). The protein losses were consistent with the underlying germline variants in MMR genes (Table 4, Fig. 2). In addition, PD-L1 positive expression (>1% in

Table 3. Multivariate analyses of recurrence-free survival and distant recurrence-free survival in this cohort.

Variable	RFS		DRFS	
	HR (95% CI)	<i>P</i>	HR (95% CI)	<i>P</i>
Age				
≤40 yr	1.00	–	1.00	–
>40 yr	0.78 (0.66–0.94)	0.007	0.90 (0.73–1.10)	0.28
Tumor size				
≤2 cm	1.00	–	1.00	–
>2 cm	1.57 (1.32–1.86)	<0.001	1.73 (1.43–2.09)	<0.001
Lymph node				
Negative	1.00	–	1.00	–
Positive	3.41 (2.92–3.98)	<0.001	3.59 (3.03–4.25)	<0.001
Grade				
I	1.00	–	1.00	–
II	1.34 (0.99–1.82)	0.06	1.45 (1.03–2.04)	0.03
III	1.38 (0.97–1.98)	0.07	1.49 (1.00–2.21)	0.05
ER status				
Negative	1.00	–	1.00	–
Positive	0.73 (0.56–0.94)	0.01	0.77 (0.58–1.01)	0.06
PR status				
Negative	1.00	–	1.00	–
Positive	0.75 (0.60–0.94)	0.01	0.73 (0.57–0.92)	0.009
HER2 status				
Negative	1.00	–	1.00	–
Positive	1.13 (0.96–1.34)	0.15	1.18 (0.98–1.41)	0.08
Treatment				
No treatment	1.00	–	1.00	–
C vs. no treatment	1.09 (0.77–1.55)	0.63	0.94 (0.65–1.36)	0.74
E vs. no treatment	1.22 (0.86–1.73)	0.28	1.06 (0.73–1.53)	0.78
C + E vs. no treatment	1.03 (0.69–1.56)	0.87	0.88 (0.57–1.36)	0.56
Pathogenic variants				
Non-carriers	1.00	–	1.00	–
MMR carriers	2.34 (0.97–5.66)	0.06	2.76 (1.14–6.69)	0.03

RFS recurrence-free survival, DRFS distant recurrence-free survival, CI confidence interval, C chemotherapy, E endocrine therapy, ER estrogen receptor, PR progesterone receptor, HER2 human epidermal growth factor receptor 2, HR hazard ratio, MMR mismatch repair.

tumor cells (TCs) or immune cells (ICs) was observed in 9 of the 13 MMR variant carriers (9/13, 69.2%) (Table 4, Fig. 2).

Somatic mutation profile, TMB and MSI status of breast cancers from MMR variant carriers

Ten of the 15 MMR variant carriers had FFPE tissues with enough tumor purity (>20%) for DNA extraction and depth sequencing on

654 cancer-related genes. To ensure the accuracy of sequencing of FFPE tissue, target region sequencing was performed on FFPE tumor tissues and fresh-frozen tumor tissues from two MMR variant carriers (P8 and P22). Similar somatic mutation profiles (exactly the same oncogenic mutations) and somatic copy number changes were detected in FFPE tumor tissues and fresh-frozen tumor tissues from the same case (Supplementary Table 1). The TMB estimated from FFPE tumor tissues was slightly lower than that from fresh-frozen tumor tissues (Supplementary Table 1).

Among the 10 breast cancers from MMR variant carriers, PIK3CA (5/10), TP53 (4/10), PTEN (2/10), and ARID1A (2/10) were the most frequently mutated oncogenes or tumor suppressor genes (Fig. 3a). Compared with the somatic mutation profile of general breast cancers reported in the TCGA dataset, breast cancers from MMR variant carriers might have more PTEN mutations [20.0% (2/10) vs. 3.6% (35/982), $P = 0.006$] and ARID1A mutations [20.0% (2/10) vs. 3.0% (29/982), $P = 0.002$] (Fig. 3b). More importantly, 2 of the 10 MMR variant carriers showed TMB-high (>10 Mut/Mb). Second hit events were detected in 1 of the 10 MMR variant carriers (P15: MLH1 carrier, LOH of the wild-type MLH1 allele), and this tumor also exhibited TMB-high and positive PD-L1 expression (Fig. 3a). MSI was estimated in three MMR carriers; only P22 (msh6 protein loss) showed MSI-L, and the other two patients (no MMR protein loss) showed MSS (Table 4).

DISCUSSION

In this study, 0.19% of the Chinese breast cancer patients were found to carry a pathogenic germline variant in the four MMR genes. Compared with non-carriers, MMR variant carriers showed distinct histology and poor survival. In addition, 5 of the 13 breast cancers from MMR variant carriers showed MMR protein loss consistent with the underlying germline variants, and the breast cancers with germline MMR variants showed a relatively high rate of TMB-high (2/10) and PD-L1 positive expression (9/13).

We found that pathogenic germline variants in MMR genes were rare in consecutive breast cancers (15/8085, 0.19%), and germline variants were more common in PMS2/MSH6 genes than in MLH1/MSH2 genes in breast cancers. In addition, we found that MMR germline variants might affect the tumor phenotype and somatic mutation profile. Our study revealed a potential association between MMR germline variants and medullary/papillary histology in breast cancer, which was also reported in breast cancers from HNPCC families in previous studies^{16,17,19}. In addition, PTEN somatic mutations were more common in breast cancers from MMR variant carriers than in those from non-carriers. A similar observation was also reported in endometrioid cancers, in which PTEN loss/somatic mutation were closely associated with MMR deficiency^{22,23}. In contrast, the effect of MMR germline variants on the survival of breast cancer was different from that of colorectal cancer. MMR germline variants predict a better survival in colorectal cancers², while our study and another recent study both found that breast cancer patients carrying MMR germline variants might have worse survival than non-carriers²⁴. Nevertheless, independent breast cancer cohorts are needed to validate the association between MMR germline variants and breast tumor phenotype because of the limited number of samples in our study.

The poor prognosis of breast cancer patients carrying MMR germline variants urged us to explore whether these patients could benefit from immunotherapy. In a previous clinical trial⁷, MMR protein loss was established as a biomarker predicting the response to immunotherapy, regardless of the cancer type. A recent study also reported that one metastatic breast cancer patient with an MMR germline variant achieved a robust and durable response upon immunotherapy²⁵. In this study, we found that one-third of breast cancers (5/13, 38.5%) with MMR germline variants showed MMR protein loss. This rate was similar to the rate

Table 4. MMR/PD-L1 protein expression and TMB of breast cancers with pathogenic germline variants in MMR genes.

Case ID	Germline	MMR proteins	MSI	TMB	PD-L1 expression in TC	PD-L1 expression in IC
P12	PMS2:p.R134X	all positive	NA	3.1	1–4%	1–4%
P14	PMS2:p.E225X	pms2-	NA	1.9	0%	1–4%
P20	PMS2:p.R315X	pms2-	NA	5.6	0%	0%
P25	PMS2:p.L351X	all positive	NA	1.3	0%	0%
P26	PMS2:p.V754fs	all positive	NA	0.6	0%	1–4%
P27	PMS2:p.R151fs	NA	NA	NA	NA	NA
P5	MSH6:p.R495X	all positive	NA	0.6	0%	0%
P8	MSH6:p.Y994X	all positive	MSS	3.1	1–4%	>10%
P22	MSH6:p.R922X	msh6-	MSI-L	2.5	0%	0%
P23	MSH6:p.T716fs	msh6-	NA	19.4	1–4%	>10%
P31	MSH6:p.T1085fs	all positive	NA	NA	80–100%	>10%
P2	MSH2:p.Q4X	all positive	NA	NA	1–4%	1–4%
P19	MSH2:p.R877fs	NA	NA	NA	NA	NA
P21	MSH2:p.Q314fs	all positive	MSS	2.5	0%	10%
P15	MLH1:p.R659X	mlh1-, pms2-	NA	15.6	1–4%	5–10%

MMR mismatch repair, TMB tumor mutation burden, MSI-L low microsatellite instability, MSS microsatellite stability, PD-L1 Programmed cell death ligand 1, IC immune cell, TC tumor cell, NA not accessible.

reported by a recent study²⁵, in which they identified 13 breast cancer patients with MMR germline variants from individuals receiving genetic testing, and found that 42% of them (5/12) displayed MMR protein loss. However, the rate of MMR protein loss in this study was lower than the rate reported in breast cancers from MMR variant carriers with an HNPCC family history (51–71%)^{17,18,26}. This might be explained by that the MMR variant carriers in our study were identified from unselected breast cancers. It is well established that TMB and PD-L1 expression are also biomarkers predicting the response to immunotherapy in breast cancers or other solid cancers^{27–31}. The TMB-high rate in MMR variant carriers in this study (20.0%) was much higher than that in general breast cancers from the TCGA dataset (1.8%). In addition, the rate of PD-L1 positive expression in MMR variant carriers in this study (69.2%) was much higher than the rate in general breast cancers reported in a previous study (23.6%)³². More importantly, we found that the pattern and strength of PD-L1 expression in MMR variant carriers might be different from that in general breast cancer patients. On the one hand, a previous study showed that PD-L1 is usually expressed in ICs instead of TCs in general breast cancers³³, while 46.2% (6/13) of MMR variant carriers in this study showed positive PD-L1 expression in TCs. On the other hand, the percentage of positive cells expressing PD-L1 in a breast cancer was usually lower than 30% in previous studies^{32,34}, while one MMR variant carrier (P31) in this study showed very strong PD-L1 expression (>80% in TCs). In summary, at least 11 of the 15 breast cancers from MMR variant carriers showed MMR protein loss, TMB-high or positive PD-L1 expression, suggesting that the majority of breast cancers from MMR variant carriers might benefit from immunotherapy.

There are several limitations in this study. First, although we have screened MMR germline variants in a large series of breast cancers, the number of MMR variant carriers was very small. Further independent breast cancer cohorts in the further are needed to validate our results. Second, a small portion of breast cancers from MMR variant carriers had genomic DNA with enough quality for MSI detection. Third, MMR variant carriers were associated with poor survival in this study, and caution is needed when interpreting of these results due to small sample size.

In conclusion, 0.19% of the Chinese breast cancer patients carried a pathogenic germline variant in MMR genes, and MMR

variant carriers showed poor survival compared with non-carriers. Breast cancers with germline MMR variants showed relatively high rates of MMR protein loss, TMB-high and PD-L1 positive expression compared with general breast cancers, suggesting that some breast cancer patients with germline MMR variants might benefit from immunotherapy.

METHODS

Patients

A total of 8085 consecutive breast cancer patients who were treated at the Breast Center of Peking University Cancer Hospital from October 2003 to May 2015 were included in this study³⁵. The cohort was unselected for age at diagnosis and family history. Detailed demographic information and tumor characteristics of each patient were collected from medical records and/or telephone interviews. Estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) status were determined using the breast tumor tissue obtained from a core needle biopsy or taken from surgery. ER or PR immunostaining was considered positive when >1% of the TCs showed positive nuclear staining. HER2 positivity was defined as a score of 3+ via immunohistochemical staining or HER2 gene amplification via fluorescence in situ hybridization. Informed written consent was obtained from all participants. This study was approved by the Research and Ethics Committee of Peking University Cancer Hospital.

MMR germline variant classification

Panel sequencing (including four MMR genes: MLH1, MSH2, MSH6, and PMS2) was performed on genomic DNA extracted from the peripheral blood of the 8085 unselected breast cancer patients³⁵. In this study, we reanalyzed the MMR germline variants detected in our previous report. Germline variations were called with GATK (version 3.6). Annotations were defined using ANNOVAR. Only variants with <1% population frequency in the population databases including gnomAD (v3.1.2) and TOPMed (version 20210514) were collected (Supplementary Data 1). Among these, truncating variants (nonsense and frameshift variants) were included in this study, but truncating variants in the last 55 base pairs of the penultimate exon or last exon that potentially avoid nonsense-mediated messenger RNA decay and do not influence known functional domains were excluded. For splice-site, synonymous, nonsynonymous, in-frame, and stop-loss variants, only variants classified as pathogenic or likely pathogenic by ClinVar (version 20210501) were included in the analysis. Variants with conflicting interpretations of pathogenicity in ClinVar were further annotated according to the ACMG/AMP standards and

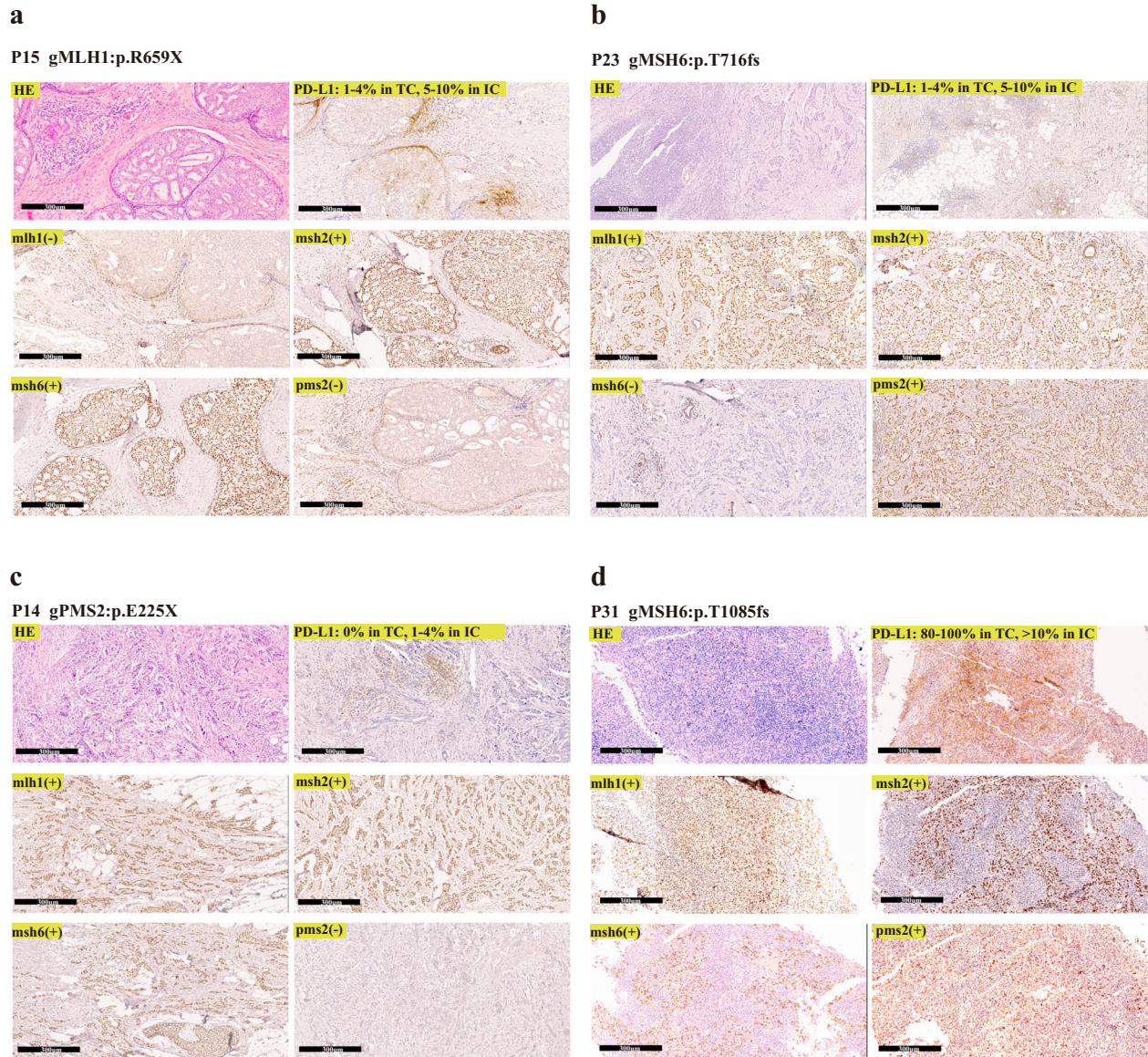


Fig. 2 H&E, PD-L1 protein and MMR protein immunostaining of breast cancers from MMR variant carriers. **a** MLH1 variant carrier, mlh1 and pms2 protein loss; **b** MSH6 variant carrier, msh6 protein loss; **c** PMS2 variant carrier, pms2 protein loss; **d** MSH6 variant carrier, strong pd-l1 protein expression. All scale bars = 300 μ m.

guidelines³⁶, with supporting data from function prediction software, public literature, and curated databases. Variants classified as pathogenic or likely pathogenic were considered as pathogenic in this study.

H&E staining

Hematoxylin and eosin (H&E) staining was performed on 4 μ m sections of formalin-fixed, paraffin-embedded breast cancer tissue from each MMR variant carrier. The staining procedures were as follows: dewaxing, dehydration, hematoxylin, differentiation, bluing, eosin, dehydration, clearing, and cover-slipping. The pathologists histopathologically evaluated the tumor cell area in each H&E section.

Immunohistochemistry (IHC)

Immunohistochemistry was performed on 4 μ m sections of formalin-fixed, paraffin-embedded tissue from MMR variant carriers. MMR protein detection was carried out using primary antibodies against mlh1 (clone GM002, mouse monoclonal antibody, catalogue numbers: GT230407, working solution, Gene Tech), msh2 (clone RED2, rabbit monoclonal antibody, catalogue numbers: GT231007, working solution, Gene Tech), msh6 (clone EP49, rabbit monoclonal antibody, catalogue numbers:

GT219507, working solution, Gene Tech), and pms2 (clone EP51, rabbit monoclonal antibody, catalogue numbers: GT215907, working solution, Gene Tech). MMR protein loss was defined as nuclear immunostaining negative for one or several MMR proteins in the TCs on whole section slides. PD-L1 IHC staining was performed using the primary antibody for the pd-l1 protein (clone SP142, rabbit monoclonal antibody, catalogue numbers: ab228462, dilution ratio 1:400, Abcam). The percentage of PD-L1 was scored as a continuous variable. The threshold for PD-L1 positivity was set at $\geq 1\%$ positive TCs and/or ICs. At least one positive control (embryo tissue section) and one negative control (PBS instead of primary antibody) were used in each IHC assay (Supplementary Fig. 1). The tumor sections were independently scored for PD-L1 expression in TCs/ICs blinded to the germline status, MMR protein status and clinical data.

Target region sequencing on breast cancers from MMR variant carriers

Ten breast cancers with germline variants in MMR genes had enough TCs in the formalin-fixed and paraffin-embedded (FFPE) blocks for target region sequencing. For each block, a H&E stained slide was reviewed by a pathologist to ensure that at least 20% of the nucleated cells in the slide

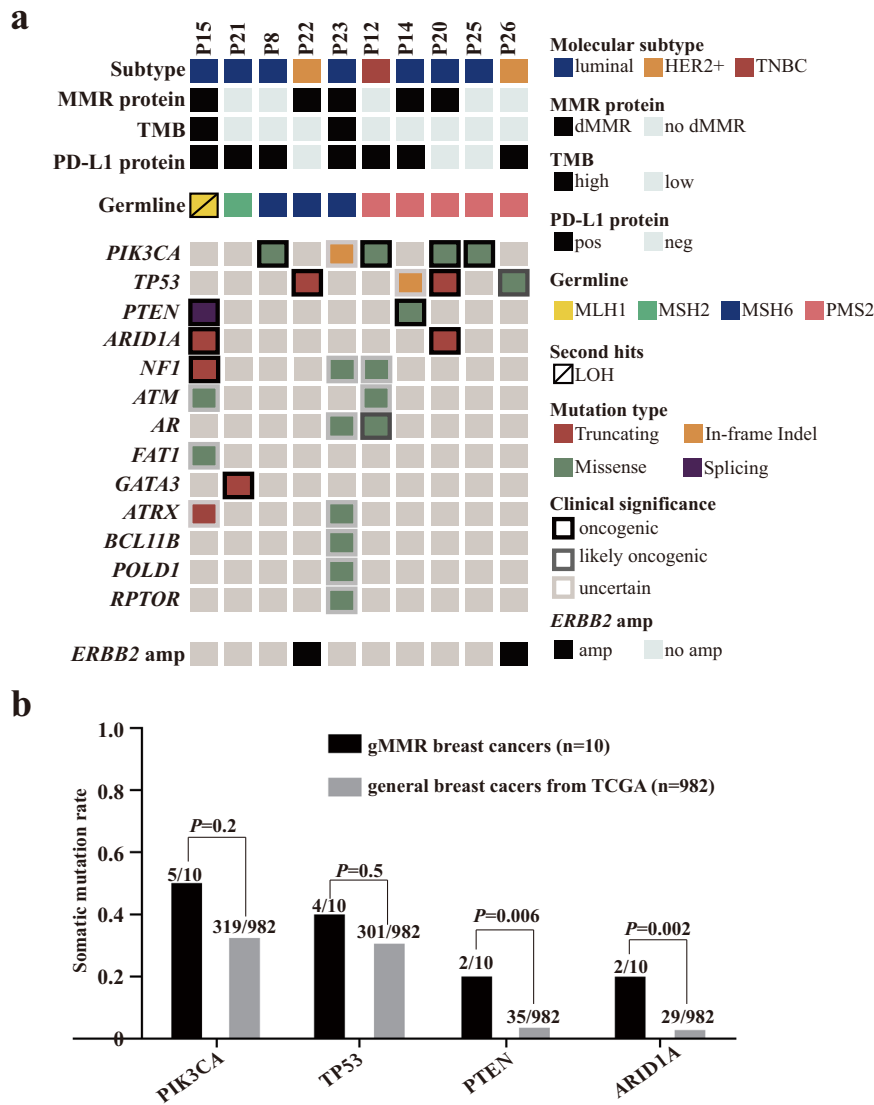


Fig. 3 Somatic mutation profiles of breast cancers from MMR variant carriers. a Landscape of frequently mutated genes in breast cancers from MMR variant carriers. **b** Comparison of the somatic mutation rates of PIK3CA, TP53, PTEN, and ARID1A between breast cancers carrying MMR germline variants and general breast cancers. Data are analyzed for statistical significance using Fisher's exact test. Two-sided P values < 0.05 were considered to be statistically significant.

were derived from the TCs. Genomic DNA was extracted from FFPE sections by using GeneRead DNA FFPE Kit (QIAGEN, Germany). For each sample, 100–300 ng of DNA was prepared to construct a paired-end DNA library. The DNA was subjected to 654 cancer-related genes (including the 4 MMR genes) target capture by using the Solid Tumor Comprehensive Test Kit. The products were sequenced on an Illumina HiSeq 4000 sequencing platform. The tumor samples were sequenced at an average depth of 570 \times on the target region. Matched blood samples were sequenced at an average depth of 1400 \times in the same regions to identify and filter germline variants. In addition, fresh-frozen tumor tissues were available for 2 of the 10 MMR variant carriers (P8 and P22). We extracted DNA from the two fresh-frozen tumor samples by using the DNeasy Blood & Tissue Kit (QIAGEN, Germany), which were also subjected to 654 cancer-related genes target sequencing with at an average depth of 1576 \times on the target region.

Sequenced reads were aligned to the human reference genome (NCBI Build 37) by the Burrows-Wheeler Aligner (version 0.1.22). Somatic indels and single nucleotide variations were called by MutLoc with an additional filter to exclude artificial mutations introduced by FFPE tissue. In brief, duplicates and soft clipped reads removed data was analyzed in MutLoc with these parameters (align quality: 30; strand bias: 0.05; keep the mutation site with highest align quality if more than one mutation sites were examined within 11 bp; keep the mutation sites supported by at least

three different reads). In addition, we filtered out single strand bias based on a read pair orientation of larger than 20:1. Somatic copy number variations were called by GATK (version 3.6). Function annotations were defined using ANNOVAR. All the somatic mutations detected in breast cancers from the ten MMR variant carriers were listed in Supplementary Data 2. Tumor mutation burden (TMB) was defined as the number of non-synonymous somatic mutations (single nucleotide variants and small insertions/deletions) per mega-base in coding regions. The TMB of each tumor was determined on 1.6 Mb of sequenced DNA and reported as mutations/Mb. TMB \geq 10 Mut/Mb was considered as TMB-high.

PCR-based MSI detection

Genomic DNA extracted from the FFPE sections of one case and fresh-frozen tumor tissues of two cases had sufficient quality for MSI detection. Genomic DNA extracted from matched blood samples served as controls to filter germline variants in microsatellite loci. Five mononucleotide loci (NR-21, BAT-26, BAT-25, NR-24, MONO-27) were used for MSI analysis (Promega, USA). Tumors were classified MSI-H (\geq 2 loci exhibit instability), MSI-L (only one locus exhibits instability) or microsatellite stable (MSS, all five loci exhibit stability).

The Cancer Genome Atlas public dataset analysis

The somatic mutation data of 982 breast cancers were downloaded from the TCGA database (<https://portal.gdc.cancer.gov/projects>). The TMB of each tumor was determined on 30 Mb of sequenced DNA and reported as mutations/Mb. TMB \geq 10 Mut/Mb was considered as TMB-high.

Statistical analysis

Categorical variables were compared using the χ^2 test or Fisher's exact test, where appropriate. Continuous variables were tested with a *t*-test, where appropriate. RFS was defined as from the time of diagnosis to first recurrence (local or distant), or death from breast cancer (for patients without a recorded relapse) or date of last follow-up. DRFS was defined as from the time of diagnosis to first distant recurrence, or death from breast cancer (for patients without a record of recurrence), or the date of the last follow up. Survival was estimated using the Kaplan–Meier method. Univariate and multivariate Cox proportional hazards models were used to determine whether a factor was associated with survival. Two-sided *P* values $<$ 0.05 were considered to be statistically significant. All analyses were performed using SPSS 20.0 software.

Reporting summary

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

DATA AVAILABILITY

The next-generation sequencing data analyzed in this study have been uploaded as Supplementary data. The other relevant data are available from the authors upon reasonable request.

Received: 16 November 2021; Accepted: 22 March 2022;

Published online: 21 April 2022

REFERENCES

- Hendriks, Y. M. C. et al. Diagnostic Approach and Management of Lynch Syndrome (Hereditary Nonpolyposis Colorectal Carcinoma): a Guide for Clinicians. *Ca Cancer J. Clin.* **56**, 213–225 (2006).
- Gryfe, R. et al. Tumor Microsatellite Instability and Clinical Outcome in Young Patients with Colorectal Cancer. *N. Engl. J. Med.* **342**, 69–77 (2000).
- Boland, C. R. & Goel, A. Microsatellite Instability in Colorectal Cancer. *Gastroenterology* **138**, 2073–2087.e3 (2010).
- Prolla, T. A. et al. Tumour susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair. *Nat. Genet.* **18**, 276–279 (1998).
- Le, D. T. et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J. Med.* **372**, 2509–2520 (2015).
- Overman, M. J. et al. Nivolumab in patients with DNA mismatch repair deficient/microsatellite instability high metastatic colorectal cancer: Update from Check-Mate 142. *J. Clin. Oncol.* **35**, 519–519 (2017).
- Le, D. T. et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* **357**, 409–413 (2017).
- Müller, A. et al. Exclusion of breast cancer as an integral tumor of hereditary nonpolyposis colorectal cancer. *Cancer Res.* **62**, 1014–1019 (2002).
- Broeke, S. Wten et al. Lynch Syndrome Caused by Germline PMS2 mutations: delineating the cancer risk. *J. Clin. Oncol.* **33**, 319–325 (2015).
- Win, A. K. et al. Colorectal and Other Cancer Risks for Carriers and Noncarriers From Families With a DNA Mismatch Repair Gene Mutation: A Prospective Cohort Study. *J. Clin. Oncol.* **30**, 958–964 (2012).
- Engel, C. et al. Risks of Less Common Cancers in Proven Mutation Carriers With Lynch Syndrome. *J. Clin. Oncol.* **30**, 4409–4415 (2012).
- Goldberg, M. et al. Association between the Lynch syndrome gene MSH2 and breast cancer susceptibility in a Canadian familial cancer registry. *J. Med. Genet.* **54**, 742–746 (2017).
- Roberts, M. E. et al. MSH6 and PMS2 germ-line pathogenic variants implicated in Lynch syndrome are associated with breast cancer. *Genet. Med. Off. J. Am. Coll. Med. Genet.* **20**, 1167–1174 (2018).
- Harkness, E. F. et al. Lynch syndrome caused by MLH1 mutations is associated with an increased risk of breast cancer: a cohort study. *J. Med. Genet.* **52**, 553–556 (2015).
- Rashid, M. U. et al. A novel deleterious c.2656G>T MSH2 germline mutation in a Pakistani family with a phenotypic overlap of hereditary breast and ovarian cancer and Lynch syndrome. *Hered. Cancer Clin. Pr.* **14**, 14 (2016).
- Jensen, U. B. et al. Mismatch repair defective breast cancer in the hereditary nonpolyposis colorectal cancer syndrome. *Breast Cancer Res. Tr.* **120**, 777–782 (2009).
- Walsh, M. D. et al. Lynch syndrome-associated breast cancers: clinicopathologic characteristics of a case series from the colon cancer family registry. *Clin. Cancer Res.* **16**, 2214–2224 (2010).
- Buerki, N. et al. Evidence for breast cancer as an integral part of lynch syndrome. *Genes Chromosom. Cancer* **51**, 83–91 (2011).
- Kanaya, N. et al. Clinicopathological features of breast cancer in Japanese female patients with Lynch syndrome. *Breast Cancer-tokyo* **26**, 359–364 (2018).
- Sorscher, S. Rationale for evaluating breast cancers of Lynch syndrome patients for mismatch repair gene expression. *Breast Cancer Res. Tr.* **178**, 469–471 (2019).
- Porkka, N. K. et al. Does breast carcinoma belong to the Lynch syndrome tumor spectrum? - Somatic mutational profiles vs. ovarian and colorectal carcinomas. *Oncotarget* **11**, 1244–1256 (2020).
- Djordjevic, B., Barkoh, B. A., Luthra, R. & Broaddus, R. R. Relationship between PTEN, DNA mismatch repair, and tumor histotype in endometrial carcinoma: retained positive expression of PTEN preferentially identifies sporadic non-endometrioid carcinomas. *Mod. Pathol.* **26**, 1401–1412 (2013).
- Levine, D. A. et al. Integrated genomic characterization of endometrial carcinoma. *Nature* **497**, 67–73 (2013).
- Cheng, A. S. et al. Mismatch repair protein loss in breast cancer: clinicopathological associations in a large British Columbia cohort. *Breast Cancer Res. Tr.* **179**, 3–10 (2019).
- Schwartz, C. J. et al. Morphological and genomic characteristics of breast cancers occurring in individuals with Lynch Syndrome. *Clin. Cancer Res. Clinanres* **2027**, 2021 (2021).
- Win, A. K., Lindor, N. M. & Jenkins, M. A. Risk of breast cancer in Lynch syndrome: a systematic review. *Breast Cancer Res.* **15**, R27 (2013).
- Kim, S. T. et al. Comprehensive molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric cancer. *Nat. Med.* **24**, 1449–1458 (2018).
- Rizvi, N. A. et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Sci. N. Y* **348**, 124–128 (2015).
- Huang, J. et al. Safety, activity and biomarkers of SHR-1210, an anti-PD-1 antibody, for patients with advanced esophageal carcinoma. *Clin. Cancer Res.* **24**, Clinanres **2439**, 2017 (2018).
- Schmid, P. et al. Atezolizumab plus nab-paclitaxel as first-line treatment for unresectable, locally advanced or metastatic triple-negative breast cancer (IMpassion130): updated efficacy results from a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol.* **21**, 3–5 (2019).
- Schmid, P. et al. Atezolizumab and Nab-Paclitaxel in Advanced Triple-Negative Breast Cancer. *New Engl J. Med* **379**, 2108–2121 (2018).
- Buisseret, L. et al. Tumor-infiltrating lymphocyte composition, organization and PD-1/PD-L1 expression are linked in breast cancer. *Oncoimmunology* **6**, e1257452 (2016).
- Cimino-Mathews, A. et al. PD-L1 (B7-H1) expression and the immune tumor microenvironment in primary and metastatic breast carcinomas. *Hum. Pathol.* **47**, 52–63 (2015).
- Solinas, C. et al. BRCA gene mutations do not shape the extent and organization of tumor infiltrating lymphocytes in triple negative breast cancer. *Cancer Lett.* **450**, 88–97 (2019).
- Sun, J. et al. Germline Mutations in Cancer Susceptibility Genes in a Large Series of Unselected Breast Cancer Patients. *Clin. Cancer Res.* **23**, 6113–6119 (2017).
- Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–423 (2015).

ACKNOWLEDGEMENTS

This study was supported by grants No. 81772824, 81974422, and 81802635 from the National Natural Science Foundation of China and No. PX2021042 from Beijing Municipal Administration of Hospitals Incubating Program.

AUTHOR CONTRIBUTIONS

Y.X. designed and supervised the study; L.H., J.S., and Z.L. analyzed and interpreted the experimental data; L.H., J.L., X.D., and F.Z. performed the immunohistochemistry experiments; L.H., Z.Q., Y.L., Y.W., and Q.W. performed the next-generation sequencing experiments; Y.X., J.Z., L.Y., and Y.X. provided breast cancer patients and clinical data; All authors drafted the paper; L.H. and Y.X. revised the paper; All authors approved the final paper. All authors accounted for all aspects of the work in

ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. L.H., J.S., and Z.L. have contributed equally to this work and share first authorship.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41523-022-00417-x>.

Correspondence and requests for materials should be addressed to Yin Wang or Yuntao Xie.

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

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022