EMT Transcription Factor ZEB1 Represses the Mutagenic POLθ-Mediated End-Joining Pathway in Breast Cancers

Mélanie K. Prodhomme[1,2], Roxane M. Pommier[1,2,3], Camille Franchet[4], Frédérique Fauvet[1,2], Valérie Bergoglio[5], Pierre Brousset[4], Anne-Pierre Morel[1,2], Anne-Cécile Brunac[4], Mojgan Devouassoux-Shisheboran[1,2], Virginie Petrilli[6], Caroline Moyret-Lalle[1,2], Jean-Sébastien Hoffmann[4], Alain Puisieux[1,2,7], and Agnès Tissier[1,2]

ABSTRACT

A characteristic of cancer development is the acquisition of genomic instability, which results from the inaccurate repair of DNA damage. Among double-strand break repair mechanisms induced by oncogenic stress, the highly mutagenic theta-mediated end-joining (TMEJ) pathway, which requires DNA polymerase theta (POLθ) encoded by the *POLQ* gene, has been shown to be overexpressed in several human cancers. However, little is known regarding the regulatory mechanisms of TMEJ and the consequence of its dysregulation. In this study, we combined a bioinformatics approach exploring both Molecular Taxonomy of Breast Cancer International Consortium and The Cancer Genome Atlas databases with CRISPR/Cas9-mediated depletion of the zinc finger E-box binding homeobox 1 (ZEB1) in claudin-low tumor cells or forced expression of ZEB1 in basal-like tumor cells, two triple-negative breast cancer (TNBC) subtypes, to demonstrate that ZEB1 represses *POLQ* expression. ZEB1, a master epithelial-to-mesenchymal transition–inducing transcription factor, interacted directly with the *POLQ* promoter. Moreover, downregulation of *POLQ* by ZEB1 fostered micronuclei formation in TNBC tumor cell lines. Consequently, ZEB1 expression prevented TMEJ activity, with a major impact on genome integrity. In conclusion, we showed that ZEB1 directly inhibits the expression of *POLQ* and, therefore, TMEJ activity, controlling both stability and integrity of breast cancer cell genomes.

Significance: These findings uncover an original mechanism of TMEJ regulation, highlighting ZEB1 as a key player in genome stability during cancer progression via its repression of *POLQ*.

See related commentary by Carvajal-Maldonado and Wood, p. 1441

Introduction

Chromosomal instability (CIN) is a hallmark of cancer (1), arising notably through the error-prone repair of double-strand breaks (DSB), ultimately resulting from oncogenic activation (2). Indeed, the usage of unfaithful pathways eventually leads to inappropriate end-joining events at the origin of genomic instability (3). Alongside the well-documented homologous recombination (HR) and canonical nonhomologous end-joining (c-NHEJ) DSB repair pathways (4, 5), mammalian cells also rely on an independent highly mutagenic theta-mediated end-joining (TMEJ) pathway, corresponding to one of the initial alternative end-joining pathways defined as a Ku-independent c-NHEJ, and then renamed microhomology-mediated end joining on the basis of the mechanism or TMEJ, on the basis of the key actor (6–9). In TMEJ, DSBs are sealed by microhomology-mediated base pairing of DNA single strands, yielding products systematically associated with short DNA deletions and insertions, potentially generating chromosomal translocations and mutagenic rearrangements (10–13). Key TMEJ actors in human cells include the A-family DNA polymerase theta (POLθ) encoded by the *POLQ* gene, PARP1 (14–16), and DNA ligase IIIα (encoded by the *LIG3* gene; refs. 7, 17). Thus far, the mechanisms regulating TMEJ in normal and cancer cells are totally unknown. In this work, we gain insight into the regulatory process of triple-negative breast cancers (TNBC). TNBCs are aggressive breast malignancies that are characterized by the lack of estrogen and progesterone receptor expression and the absence of HER2 overexpression. TNBCs represent up to 20%–25% of all breast carcinomas. According to molecular classifications based on gene expression profiles, TNBCs are essentially composed of basal-like and claudin-low subtypes (18). Claudin-low tumors display a low level of expression of cell–cell adhesion molecules, such as claudins or E-cadherin, encoded by the *CDH1* gene. In addition, they are highly enriched in mesenchymal traits and stem cell features, and are therefore, considered to be the most primitive breast cancers with poor survival outcomes compared with many other breast cancer subtypes (19, 20).

ZEB1, a transcription factor inducer of the epithelial-to-mesenchymal transition (EMT), modulates breast cancer cell plasticity by conferring stemness properties to the cells via numerous mechanisms, including the transcriptional repression of epithelial actors, such as E-cadherin (*CDH1* gene; ref. 21), or miRNAs like miR-200 (22).

[1]Université de Lyon, Université Claude Bernard Lyon 1, INSERM 1052, CNRS 5286, Centre Léon Bérard, Cancer Research Centre of Lyon, Équipe Labellisée Ligue contre le Cancer, Lyon, France. [2]LabEx DEVweCAN, Université de Lyon, Lyon, France. [3]Gilles Thomas Bioinformatics Platform, Centre Léon Bérard, Cancer Research Centre of Lyon, Lyon, France. [4]Laboratoire de Pathologie, Laboratoire d'excellence Toulouse Cancer, Institut Universitaire du Cancer-Toulouse, Oncopole, Toulouse Cedex, France. [5]Cancer Research Centre of Toulouse, INSERM UMR 1037, Toulouse, France. [6]Université de Lyon, Université Claude Bernard Lyon 1, INSERM 1052, CNRS 5286, Centre Léon Bérard, Cancer Research Centre of Lyon, Lyon, France. [7]Institut Curie, PSL Research University, Paris, France.

Prodhomme et al.

ZEB1 expression was shown to promote malignant transformation while maintaining genome stability (23). In particular, high *ZEB1* expression is causally associated with claudin-low tumors, characterized by a subnormal genomic landscape (24). ZEB1 has been implicated both in preventing the formation of oncogene-induced DNA damage and in increasing the clearance of DNA breaks. On one hand, ZEB1 is able to protect mammary stem cells against oncogene-induced damage through the activation of a preemptive antioxidant program, and favors tumorigenesis in the absence of gross genomic instability (24). On the other hand, the kinase ATM, critical player in DNA damage response, phosphorylates and stabilizes ZEB1, triggering the cell-cycle checkpoint CHK1 stabilization at the origin of treatment resistance (25). Owing to these pleiotropic effects, ZEB1 is considered to be the central factor in providing cancer cells with a high level of plasticity and may thus, be pivotal in the development of therapeutically resistant cancers.

Here, we show that ZEB1 directly controls the expression of *POLQ* and influences not only genome stability of breast cancer cells, but also genome integrity.

Materials and Methods

Bioinformatics and statistical analyses of public databases (The Cancer Genome Atlas and Molecular Taxonomy of Breast Cancer International Consortium)

All bioinformatics and statistical analyses were carried out with the R software (version 3.5.1; ref. 26). Figures were created using either the R software or GraphPad Prism 6.0 (GraphPad Software Inc.; RRID:SCR_002798). OncoPrint plots were generated using the ComplexHeatMap open source software (https://github.com/jokergoo/ComplexHeatmap; RRID:SCR_017270; ref. 27).

Expression data processing

Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) microarray expression data from discovery and validation sets were extracted from the EMBL–EBI archive (EGA; http://www.ebi.ac.uk/ega/; accession no. EGAS00000000083; RRID: SCR_ 004944; "normalized expression data" files; ref. 28). The expression levels of different probes associated with the same Entrez Gene ID were averaged for each sample to obtain a single expression value by gene.

The Cancer Genome Atlas (TCGA; RRID:SCR_003193) breast invasive carcinoma (BRCA) RNA sequencing expression data were extracted as fragments per kilobase of transcript per million mapped reads (FPKM) values from the Genomic Data Commons (GDC) data portal (https://portal.gdc.cancer.gov/). FPKM data by gene were converted to transcripts per kilobase million (TPM) as follows: for each gene, g ∈ G and each sample, s ∈ S,

$$TPM(g,s) = \left(\frac{FPKM(g,s)}{\sum_{i=1}^{G} FPKM(i,s)} \right) \times 10^6$$

Expression data by gene from TCGA and METABRIC (discovery and validation sets independently) were finally merged in a common file, keeping all genes present in both datasets, and batch normalization was performed using the R function ComBat from sva package (29, 30). The final expression dataset was composed of 18,845 genes and 3,083 breast tumor samples.

Copy-number data processing

METABRIC segmented copy-number data from discovery and validation sets were extracted from the EMBL–EBI archive [EGA, http:// www.ebi.ac.uk/ega/; accession no. EGAS00000000083; RRID:SCR_ 004944; "segmented (CBS) copy-number aberrations" files; ref. 28].

TCGA (RRID:SCR_003193) BRCA segmented copy-number data were extracted from the GDC data portal repository (files corresponding to alignments on the hg19 version of the human genome without germline copy-number variation were chosen).

As described previously (24, 31), fraction of genomic alterations (FGA) was evaluated from TCGA and METABRIC segmented copy-number data (both generated from Affymetrix SNP6.0 arrays) as follows:

$$FGA = \frac{\sum_{CNi > WM+ T} L(i)}{(\sum L(i))} + \frac{\sum_{CNi < WM- T} L(i)}{(\sum L(i))}$$

For each segment i, CNi is the mean log R ratio along segment i, $L(i)$ is the length of segment i, WM is the weighted median of CNi by $L(i)$ for each sample I, and T is the threshold value of the CNi above which the segments are considered to be altered. In other words, FGA is the ratio of the sum of the lengths of all segments with signal above the threshold to the sum of all segment lengths, that is, FGA is the percentage of the genome displaying an aberrant copy number (deletion and amplification). For METABRIC and TCGA TNBCs analysis, T was set as 0.1, taking into account that TNBCs were not sorted by cellularity.

Triple negative status

Estrogen, progesterone, and HER2 receptors' statuses were determined through expression analysis of the *ESR1, PGR*, and *ERBB2* genes, respectively. Using global distribution of each gene expression, samples were classified into positive and negative subgroups using mclust R package (version 5.4.2), which decomposes the global distribution into Gaussian mixture models to classify samples (32).

Breast cancer subtype assignment

Breast cancer molecular subtype attribution [basal-like, luminal A, luminal B, Her2, normal-like, and integrative clusters (IntClust)] was performed using the R package "genefu," version 3.8 (33). Basal-like, luminal A, luminal B, Her2, and normal-like subtype assignments were computed with five different algorithms (PAM50, AIMS, SCMGENE, SSP2006, and SCMOD2; refs. 34–38). An assignment was considered final if defined by at least three different algorithms. In case of divergence between classifiers, PAM50 subtype attribution was conserved.

The claudin-low subtype classification was defined by nearest centroid method. To achieve this, we computed the Euclidean distance between each sample and the previously described claudin-low and non-claudin-low centroids, using the 1,667 genes defined by Prat and colleagues as significantly differentially expressed between claudin-low tumors and all other molecular subtypes (20).

Cell cultures

HMEC-hTERT was generated in the laboratory as described previously (23). HMEC-hTERT was cultured in 1:1 DMEM/Ham F12 medium with 1% GlutaMAX (Gibco-Thermo Fisher Scientific; catalog no. 31331093) supplemented with 10 ng/mL Human EGF (PromoCell; catalog no. C-60170), 0.5 mg/mL Hydrocortisone (Sigma-Aldrich; catalog no. H0888), and 10 mg/mL insulin (Actrapid, Novonordisk).

BT-20 (human mammary carcinoma cells, ATCC) was maintained in minimum essential medium with 1% GlutaMAX (Gibco-Thermo Fisher Scientific; catalog no. 41090093). MDA-MB-468 (human mammary adenocarcinoma cells, derived from metastatic site: pleural effusion, DSMZ) was maintained in Leibovitz L15 with 1% GlutaMAX (Gibco-Thermo Fisher Scientific; catalog no. 31415086). HCC70

(human mammary primary ductal carcinoma, ATCC) was cultured in RPMI1640 with 1% GlutaMAX (Gibco-Thermo Fisher Scientific; catalog no. 61870044) supplemented with 1.5 g/L sodium bicarbonate (Gibco-Thermo Fisher Scientific; catalog no. 25080060), 10 mmol/L HEPES (Gibco-Thermo Fisher Scientific; catalog no. 15630056), and 1 mmol/L sodium pyruvate (Gibco-Thermo Fisher Scientific; catalog no. 11360039). HCC1937 (human mammary carcinoma cells, ATCC) and BT-549 (human mammary carcinoma cells, ATCC) were cultured in RPMI1640 with 1% GlutaMAX (Gibco-Thermo Fisher Scientific; catalog no. 61870044). SUM159 cells were a gift from Hasan Korkaya's laboratory at Augusta University (Augusta, GA). SUM159 was cultured in Ham-F12 with glutamine (Gibco-Thermo Fisher Scientific; catalog no. 21765037) supplemented with 3.2 μg/mL Gentamicin (Gibco-Thermo Fisher Scientific; catalog no. 15710049), 5 mg/mL insulin (Actrapid, Novonordisk), and 2 mg/mL hydrocortisone (Sigma-Aldrich; catalog no. H0888). MDA-MB-231 (human mammary adenocarcinoma cells, ATCC), Hs 578T (human mammary carcinoma cells, ATCC), CAL-120 (human mammary adenocarcinoma cells, DSMZ), Phoenix, Plat-E, and HEK293T were maintained in DMEM with 1% GlutaMAX (Gibco-Thermo Fisher Scientific; catalog no. 31966047).

All media were supplemented with 10% FCS (Sigma-Aldrich or Eurobio) and penicillin–streptomycin (100 μg/mL, 100 U/mL, respectively, Gibco-Thermo Fisher Scientific; catalog no. 15140130), except for MDA-MB-468, which was supplemented with 20% FCS (Sigma-Aldrich).

All cell lines were kept at 37°C in a 5% CO$_2$/95% air incubator, and were routinely tested negative for *Mycoplasma* contamination using the Lonza MycoAlert PLUS Mycoplasma Detection Kit (Lonza; catalog no. LT07-318).

Lentiviral and retroviral infections

To produce lentiviral particles, 2×10^6 HEK293T cells were transfected using the GeneJuice Transfection Reagent (Sigma-Aldrich; catalog no. 70967-4), according to the manufacturer's instructions, with 13 μg of total lentiviral expression vectors (5.1 μg pCMVdeltaR8.91, 1.3 μg phCMVG-VSVG, and 6.6 μg plasmid of interest). The pCMVdeltaR8.91 and phCMVG-VSVG vectors were gifts from D. Nègre (International Centre for Infectiology Research, INSERM U1111–CNRS UMR5308–ENS de Lyon–UCB Lyon1, EVIR Team, Lyon, France).

To produce retroviral particles, 2×10^6 Phoenix cells were transfected by GeneJuice Transfection Reagent (Sigma-Aldrich; catalog no. 70967-4), according to the manufacturer's instructions, with 10 μg of plasmid of interest.

For both infections, the supernatant was collected, filtered, supplemented with 5 or 10 μg/mL Polybrene (Sigma-Aldrich; catalog no. H9268) 48 hours after transfection and combined with the targeted cells for 12 hours.

The ZEB1-depletion model in MDA-MB-231 (MDA-MB-231 ZEB1$^{-/-}$ clones) using the CRISPR-cas9 gene editing technology was generated. Scrambled sgRNA/Cas9 All-in-One Lentivector (Applied Biological Materials; catalog no. K010) and *ZEB1* sgRNA/Cas9 All-in-One Lentivector (Human) (target 1: 5′-CACCTGAAGAGGACCAG-3′; Applied Biological Materials; catalog no. K2671006) lentiviral particles were used to infect MDA-MB-231 cells. Scrambled sgRNA/Cas9 and *ZEB1* sgRNA/Cas9 cells were selected with Puromycin (InvivoGen; catalog no. ant-pr-1) at 1 μg/mL 48 hours after infection. After cloning by limiting dilution, single cells were grown for approximately 3 weeks and colonies were screened for knockouts by quantitative PCR and genomic DNA sequencing and Western blot-

ting. Genomic DNA sequencing was performed using the Sanger sequencing method with the following primers for amplification and sequencing: 5′-TGAACTGAACGTCAGAGTGGT-3′ (forward) and 5′-TCACGTGCAGTGGCATTACT-3′ (reverse).

To generate the model with forced overexpression of ZEB1 in none-ZEB1–expressing cells, BT-20 and HCC70 basal-like cells were infected with retroviral pBabe expression vectors containing *ZEB1*. Cells were selected with neomycin at 100 μg/mL for BT-20 (Gibco-Thermo Fisher Scientific; catalog no. 10131027) and with puromycin (InvivoGen; catalog no. ant-pr-1) at 1 μg/mL 48 hours after infection.

For luciferase assay, MDA-MB-231 cells were infected with four lentiviral reporter pEZX-LvPG04 plasmids (GeneCopoeia). Two *POLQ* promoter constructs, a *CDH1* promoter (GeneCopoeia; catalog no. HPRM45458-LvPG04) and a negative promoter (GeneCopoeia; catalog no. NEG-LvPG04) as a control, were independently transduced. The *POLQ* promoter −691 bp was generated from the *POLQ* promoter −1,280 bp (GeneCopoeia; catalog no. HPRM54321-LvPG04-01) digested by *Eco*R1 and *Spe*1 restriction enzymes and reconstituted (New England Biolabs).

To generate the model with forced overexpression of *POLQ* in MDA-MB-231, lentiviral pCDH-EF1-FHC-*POLQ* vector containing human *POLQ* cDNA (Addgene; catalog no. 64875 for *POLQ* and catalog no. 64874 for empty control) was used. Cells were selected with Puromycin (InvivoGen; catalog no. ant-pr-1) at 1 μg/mL 48 hours after infection.

Drugs and siRNA

6-Thioguanine (6-TG; Sigma-Aldrich; catalog no. A4882) stock solution was dissolved in NaOH 0.1 N and cytochalasin B (Sigma-Aldirch; catalog no. C6762-1MG) stock solution was dissolved in DMSO.

Transient siRNA-mediated knockdown was performed with INTERFERin Reagent (PolyPlus-transfection; Ozyme; catalog no. POL409-50) according to the manufacturer's protocol during the time of the experiment (kinetic or single point). siRNAs were used at a final concentration of 2 nmol/L for MDA-MB-231 or 8 nmol/L for BT-549 and SUM159, and cells were treated every day. siRNA sequences (Eurogentec) are as follows: siRNA nontargeted siNT, 5′-GGU-UUGGCUGGGGUGUUAU-3′; si*ZEB1*#1, 5′-GGUAGAUGGUAA-UGUAAUA-3′; si*ZEB1*#2, 5′-GCAACAGGGAGAAUUAUUATT-3′; and si*POLQ*, 5′-CAAACAACCCUUAUCGUAAA-3′.

Western blot analysis and antibodies

Equal amounts of protein of each sample were analyzed using SDS-PAGE, electrophoretic transfer, immunoblotting, and chemiluminescence detection. Briefly, cells were washed and scratched with ice-cold PBS (Eurobio; catalog no. CS1PBS01K-BP) supplemented with Protease Inhibitor Cocktail (PIC, Sigma-Aldrich, catalog no. 11836145001), Phenyl-methane Sulfonyl Fluoride (PMSF, Sigma-Aldrich; catalog no. 93482), and Phosphatase Inhibitor Cocktail (PhoIC 2 and 3, Sigma-Aldrich; catalog no. P5726-5ML and catalog no. P0044-5ML, respectively) on ice. The cell pellets were lysed in 2% SDS, 125 mmol/L Tris, pH 6.8, PIC, PMSF, and PhoIC 2 and 3 on ice. After sonication, proteins were separated by SDS-PAGE and transferred to Polyvinylidene Difluoride Membranes (Bio-Rad; catalog no.1620177). Antibodies and dilutions were as follows: anti-ZEB1, 1:1,000 (polyclonal; Sigma Life Science; catalog no. HPA027524); anti-POLθ, 1:10,000 (as described previously; ref. 39); anti-PARP1, 1:1,000 [polyclonal PARP-1 (H-250); Santa Cruz Biotechnology; catalog no. sc-7150]; anti-LIG3, 1:1,000 (clone 6G9; Santa Cruz Biotechnology; catalog no. sc-56089); anti-CAS9, 1:1,000 (clone 7A9-3A3; Cell Signaling Technology; catalog

no. 14697); and anti-Human DNA Topoisomerase I, 1:1,000 (clone C-21; Cell Signaling Technology; catalog no. 556597). Species-specific secondary horseradish peroxidase (HRP)-coupled antibodies (Santa Cruz Biotechnology; goat anti-mouse catalog no. SC 2005 and mouse anti-rabbit catalog no. SC 2357) were used. Protein bands were visualized using Clarity or Clarity-max (Bio-Rad; catalog no. 1705061 and catalog no. 1705062, respectively) and the ChemiDoc MP System (Bio-Rad).

Quantitative real-time PCR

Total RNA was extracted with the RNeasy Mini Kit (Qiagen; catalog no. 74106) according to the manufacturer's recommendations. Reverse transcription was performed from 1 μg total RNA with the Dynamo cDNA Synthesis Kit (Thermo Fisher Scientific; catalog no. F-470L). The reverse transcription product was diluted 1:10 and used as cDNA template for qPCR analysis. TaqMan Quantitative PCR (Bioline Meridian Bioscience Europe; catalog no. BIO-86050) was used in the detection of PCR products in real-time in a CFX96 Real-time PCR Detection System (Bio-Rad) according to the manufacturer's instructions. Quantitative reverse-transcription PCR (qRT-PCR) was performed using 200 nmol/L of specific primers and DNA probes (**Table 1**; design on universal probe library by Roche Life Science). Conditions for the TaqMan method were 2 minutes at 50°C, 20 seconds at 95°C, and then 40 cycles, each consisting of 3 seconds at 95°C and 30 seconds at 60°C. The housekeeping genes used were *UBB* and glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*). The comparative C_t method was used to quantify the expression of the gene of interest. Refer to the list of primers and probes in **Table 1**.

Human tumor samples and IHC analyses

Ten rare positive and 22 negative previously characterized ZEB1 tumors (24) were analyzed for *POLQ* expression by RNAscope. RNA ISH for Hs-*POLQ* mRNA was performed on the Ventana Discovery Ultra Automated Slide Staining System (Roche Diagnostics) using RNAscope VS Universal HRP Reagent Kit (Brown; Advanced Cell Diagnostics, Inc.; catalog no. 323220) and RNAscope probe specific to the region of the gene encoding *Homo sapiens* DNA polymerase θ mRNA (catalog no. 465519) according to the manufacturer's instructions. Briefly, 5 μm formalin-fixed, paraffin-embedded tissue sections were pretreated at 96°C for 16 minutes prior to hybridization with the target probes. Preamplifier, amplifier, and HRP-labeled oligos were then hybridized sequentially, followed by chromogenic precipitate development. RNA integrity was controlled by the use of a RNAscope probe specific to Hs-PPIB RNA (catalog no. 313909). A negative control with a probe specific to bacterial dapB RNA was also performed (catalog no. 312039). Specific RNA staining signal was identified as brown dots. Slides were digitized using Panoramic 250 Flash II Slide Scanner (3DHISTECH) with 40× objective (resolution, 0.12154 μm/pixel) and extended focus algorithm. For each case, three representative images were acquired at ×60 magnification with CaseViewer (3DHISTECH). In the images, positive cells (cells with at least one

POLQ RNAscope signal in the nucleus) were manually annotated using dedicated annotation layers with the counter tool of the viewer Aperio ImageScope version 12.3.2 (Leica Biosystems). Percentage of stained cells, mean number of *POLQ* RNAscope signal per stained cell, and H-score (product of the two previous scores) were calculated for each case. *POLQ* expression was considered low or high with respect to a threshold for H-score at 30. This method for RNAscope quantification relies on an excellent interobserver correlation (R^2 = 0.96; $P < 0.0001$, between two independent pathologists).

Chromatin immunoprecipitation assay

Chromatin immunoprecipitation (ChIP)-IT High Sensitivity Kit (Active Motif; catalog no. 53040) was used to determine the association of the transcription factor ZEB1 with *POLQ*- or *CDH1*-specific genomic regions. MDA-MB-231 and MDA-MB-231 ZEB1$^{-/-}$ cells were subjected to cell fixation, 1% Formaldehyde (Sigma-Aldrich, catalog no. 252549-1L) on ice to cross-link the proteins bound to chromatin DNA. After washing, chromatin DNA was sheared by sonication to produce DNA fragments of around 200–1,000 bp. The same amounts of sheared DNA were used for immunoprecipitation using anti-ZEB1 antibody (GeneTex; catalog no. GTF105278) or an equal amount of pre-immune Rabbit IgG (Bio-Rad; catalog no. PRABP01, 10 mg). The immunoprecipitate was then incubated with protein G magnetic beads, and the antibody–protein G magnetic beads complex was collected for subsequent reverse cross-linking. The same amount of sheared DNA without antibody precipitation was processed for reverse cross-linking and served as an input control. DNA recovered from reverse cross-linking was used for qPCR to determine the abundance of the target DNA sequence(s) relative to the input chromatin. ChIP-qPCR primers for the *POLQ* promoter were: primers #1: 5′-ACGTTCAGAACTCGTTCGCT-3′ (forward) and 5′-CCCCAGGGATCGTTATGAGC-3′ (reverse); primers #2: 5′-CCGGCGAGATCTCTTTTATT-3′ (forward) and 5′-GTCAGT-TAATGAAGTGTGCCA-3′ (reverse); and ChIP-qPCR primers for *CDH1* promoter: 5′-GGCCGGCAGGTGAACCCTCA-3′ (forward) and 5′-GGGCTGGAGTCTGAACTGA-3′ (reverse).

Luciferase assay

For luciferase assay, MDA-MB-231 was transfected using an siRNA against *ZEB1* or control siRNA. Cells were treated each day during 96 hours. After 48 hours of siRNA treatment, MDA-MB-231 was infected with four different lentiviral GLuc-ON Promoter Reporter Plasmids pEZX-LvPG04 (GeneCopoeia). Two *POLQ* promoter constructs, a *CDH1* promoter (GeneCopoeia; catalog no. HPRM45458-LvPG04) and a negative promoter (GeneCopoeia; catalog no. NEG-LvPG04) as a control, were independently transduced. The *POLQ* promoter −691 bp was generated from the *POLQ* promoter −1,280 bp (GeneCopoeia; catalog no. HPRM54321-LvPG04-01) digested by *Eco*R1 and *Spe*1 restriction enzymes and reconstituted (New England Biolabs). pEZX-LvPG04 contains the Gaussia luciferase reporter gene under the control of the indicated promoter, and SEAP as an internal

Table 1. Primer and probe list for qPCR.

Primers	DNA probe	Forward	Reverse
GAPDH	60	5′-AGCCACATCGCTCAGACAC-3′	5′-GCCCAATACGACCAAATCC-3′
LIG3	7	5′-GCTGGCCACAAAGTCTTCTC-3′	5′-CCAGTGAAGATGTCCAGCAA-3′
PARP1	22	5′-TCTTTGATGTGGAAAGTATGAAGAA-3′	5′-GGCATCTTCTGAAGGTCGAT-5′
POLQ	40	5′-GATTGAGCCAGAGTCTGTTGG-3′	5′-TCCATAAATGATCCCATAGCTT-3′
UBB	39	5′-AGGATCCTGGTATCCGCTAAC-3′	5′-TCACATTTTCGATGGTGTCACT-3′
ZEB1	57	5′-AACTGCTGGGAGGATGACAC-3′	5′-TCCTGCTTCATCTGCCTGA-3′

control. Forty-eight hours after infection, cells were seeded onto 96-well plates (20,000 cells/well). After 48 hours, the supernatants were collected to reveal the luciferase signal using the Secrete-Pair Dual Luminescence Assay Kit (GeneCopoeia, TEBU; catalog no. LF032), according to the manufacturer's recommendations. Changes in transcription activity of *POLQ* promoter were normalized with respect to the corresponding negative control samples. *CDH1* promoter and negative promoter were used as controls.

Hypoxanthine-guanine phosphoribosyltransferase gene mutation assay

Cells were infected with the indicated HPRT1 sgRNA CRISPR/Cas9 All-in-One Constructs (Applied Biological Materials; catalog no. K0986605) and cultured for an additional 7 days (cells were passaged twice in this period) with Puromycin (InvivoGen; catalog no. ant-pr-1) for selection. After 7 days, cells were trypsinized, counted, and seeded at low density (500 cells for untreated conditions and 3,000 cells for 6-TG-treated conditions). For each sample, seven plates were seeded: three were left untreated to determine the cloning efficiency (−6TG condition), whereas 5 μg/mL 6-TG (Sigma-Aldrich; catalog no. A4882) was added to the four other plates to select hypoxanthine-guanine phosphoribosyltransferase (HPRT)-deficient cells. Two weeks after the addition of 6-TG, plates were washed with PBS 1 × and stained with solution containing 50% ethanol, 5% acetic acid, and 0.5% brilliant blue R (Sigma-Aldrich; catalog no. B7920). Surviving colonies were scored, and the HPRT mutation frequency was calculated as follows:

$$\text{Mutation frequency} = \frac{\text{Number of } 6-\text{TG resistant clones}}{\text{Number of cells plated} \times \text{cloning efficiency}}$$

$$\left(\text{Cloning efficiency} = \frac{\text{Number of survived clones on untreated plates}}{\text{Number of cells plated}}\right)$$

Micronuclei

Cells were plated onto removable Chamber Slide (IBIDI; catalog no. 81201) and treated for 36 hours for MDA-MB-231 or 72 hours for BT-20 and HCC70 with cytochalasin B in the culture media at a final concentration of 3 μg/mL. When the majority of cells were binucleated, they were fixed in 4% formaldehyde in 1× PBS and 1% BSA (Sigma-Aldrich; catalog no. A8412) for 20 minutes at room temperature. Cells were stained with phalloidin-TRITC and Hoechst. Finally, the silicone chambers were removed and the slides were mounted with Fluoromount Aqueous Mounting Medium (Sigma-Aldrich; catalog no. F4680-25ML). A minimum of 200 binucleated cells was counted per condition.

Statistical analysis

Statistical analysis and graphs were performed using GraphPad Prism 6.0 (RRID:SCR_002798). Data are expressed as mean ± SEM of at least three independent experiments and were analyzed using unpaired two-tailed Student *t* tests with a Welch correction. Significance is represented with asterisks; $P < 0.05$ was considered to be significant. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

Results

POLQ expression according to genomic instability in breast tumors

To explore the importance of TMEJ in breast tumorigenesis and determine its impact on cancer genome stability, we first evaluated the expression of *POLQ, PARP1*, and *LIG3* in primary breast tumors according to their genomic landscape. We analyzed the FGA, as a consequence of episodes of CIN accumulation over the course of tumor development and progression, in 3,083 primary breast cancers from combined databases, namely TCGA-BRCA and the METABRIC. We observed that the mRNA expression of *POLQ* was significantly positively correlated with FGA (**Fig. 1A**), whereas the correlation with *PARP1* and *LIG3* expression was less significant (Supplementary Fig. S1A). These results highlighted POLθ as a putative marker of CIN. We then assessed the variation in *POLQ, PARP1*, and *LIG3* expression in primary tumors with distinct genomic landscapes by comparing their abundance at the transcript level in 10 IntClust from a molecular classification of breast cancers based on genomic and transcriptomic analyses (28). The IntClust10, mostly characterized as the high-genomic instability subgroup (**Fig. 1B**), was greatly enriched in tumors expressing high levels of *POLQ* (**Fig. 1C**). Conversely, the expression of *POLQ* was reduced in IntClust4 and IntClust3 breast tumors, characterized by a paucity of FGA, and IntClust4 being termed the copy-number alterations (CNA)-devoid subgroup (28, 40).

In agreement with previous observations, we found that 52% of all claudin-low tumors belonged to the genomically stable IntCluster-4 (24), while 72% of basal-like tumors were mostly found in IntClust10, characterized by a large number of genomic aberrations (**Fig. 1D**). We have shown previously (24, 31), and confirmed here with additional data, that the IntClust4 was enriched in TNBCs with higher levels of *ZEB1* expression (**Fig. 1E**; Supplementary Fig. S1B). Although *POLQ, PARP1*, and *LIG3* expression within the subtypes displayed similar trends (Supplementary Fig. S1B), *POLQ* showed the most significant differences according to subtype. Having shown that these TMEJ actors were all underexpressed in the CIN^low IntClust with high *ZEB1* expression, we then examined whether their regulation was coordinated. To address this, we first analyzed the steady-state levels of ZEB1, POLθ, PARP1, and DNA ligase IIIα (LIG3) proteins in four basal-like and four claudin-low TNBC cell lines. Immunoblot analyses revealed low level of POLθ protein in most of the four claudin-low ZEB1-expressing cell lines compared with basal-like cells that do not express ZEB1 (**Fig. 1F**). No such correlation was seen for PARP1 or LIG3 (**Fig. 1F**). Compared with immortalized human mammary epithelial cell (HMEC-hTERT), POLθ protein appeared mostly increased in none-ZEB1–expressing cells (Supplementary Fig. S1C). Previous studies reported variable levels of *POLQ* expression in breast cancer cell lines (41), but the putative connection with ZEB1 was not explored. In addition, we show, using qRT-PCR, that *ZEB1* expression was elevated in claudin-low cell lines, whereas *POLQ* was largely poorly expressed in most of the basal-like cells (Supplementary Fig. S1D). No major variation was observed at the mRNA level for *PARP1* and *LIG3* in the tested cell lines, as shown previously for the proteins. Taken together, these data indicate that *POLQ* expression is lower in ZEB1-expressing tumors and cancer cell lines.

ZEB1-positive tumors are usually *POLQ* negative

We next performed cooccurrence analyses for gene expression across a set of 530 primary TNBCs from the combined TCGA and METABRIC databases. Statistical analysis for mutual exclusivity using the OR calculation revealed a significant likelihood OR of 0.072 ($P = 0.005574$), which implies that the changes found in the expression of these two genes were mostly mutually exclusive (**Fig. 2A**). Owing to the lack of available POLθ antibodies for IHC detection, we then used RNAscope to monitor *POLQ* expression in 10 TNBCs characterized previously for ZEB1 expression (24), by *in situ* IHC staining (**Fig. 2B**). *POLQ* expression was low in seven, high in two, and intermediate in
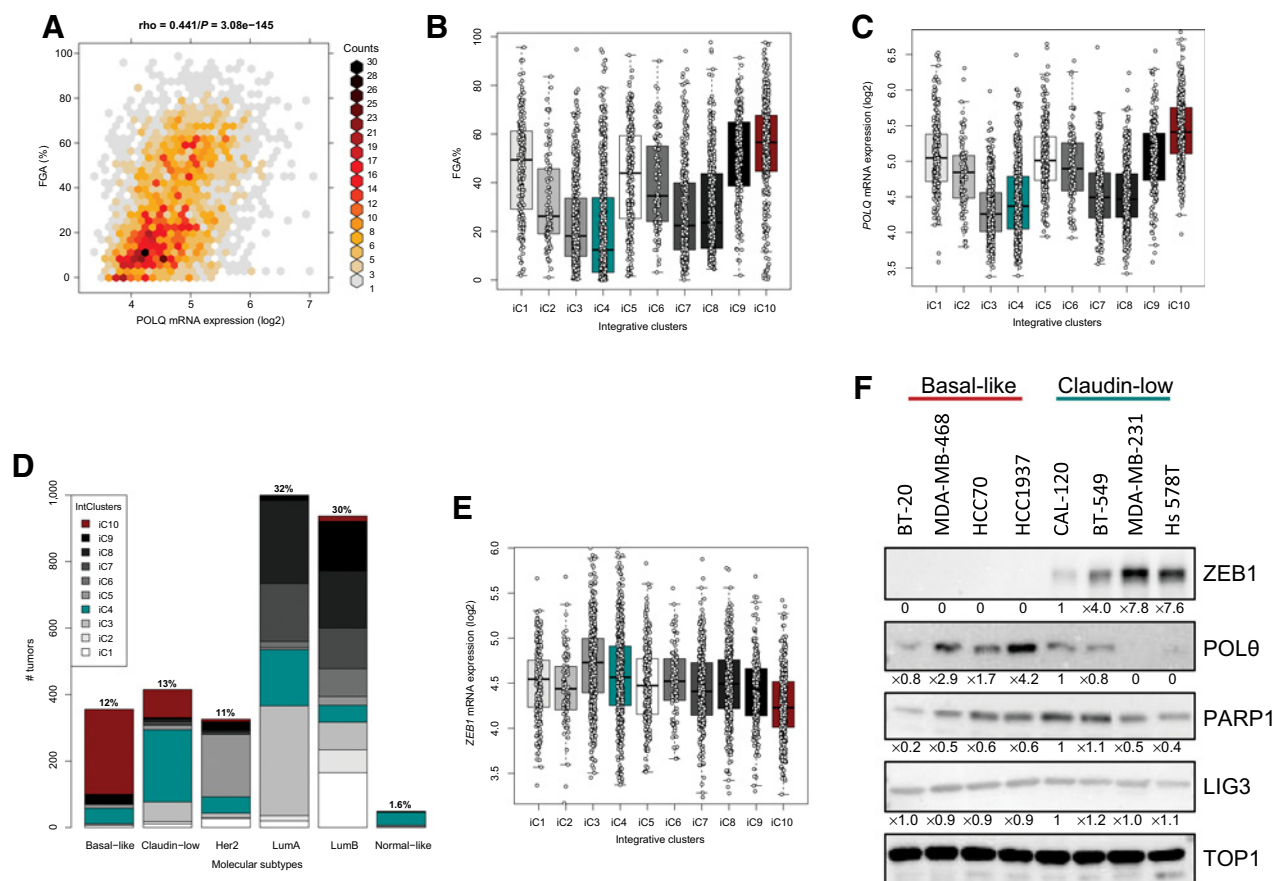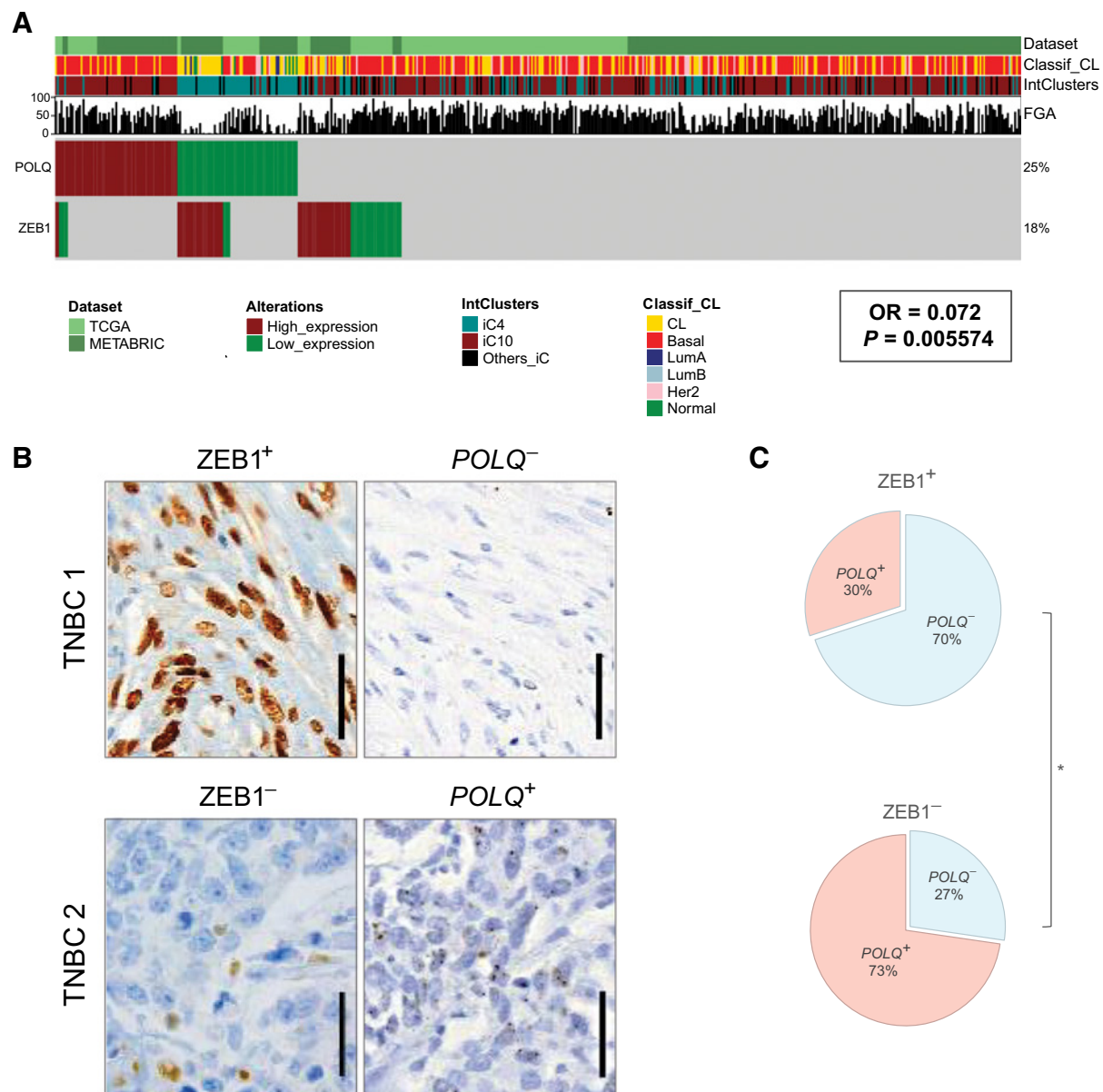
Figure 1.
POLQ expression according to genomic instability in breast tumors. **A,** Pearson correlation between the FGA and *POLQ* mRNA expression in 3,083 breast tumors for the combined TCGA and METABRIC datasets. The color code indicates the number of samples (counts) corresponding to each color. **B, C,** and **E,** FGA (**B**), *POLQ* mRNA expression (**C**), and *ZEB1* mRNA expression (**E**) in subgroups (IntClust, iC) of breast tumors. Analyses were performed for the combined METABRIC and TCGA breast cancer datasets. Results are expressed as log$_2$ of mRNA expression. Significance was determined using the Mann–Whitney U test. Central line, median; points, all data points. **D,** Distribution of IntClust among breast molecular subtypes (PAM50 + claudin-low). **F,** Immunoblot showing the protein levels of ZEB1, POLθ, PARP1, and DNA ligase III (LIG3) in TNBC cell lines normalized to CAL-120. DNA topoisomerase I (TOP1) served as a loading control. Noted at the bottom are the mean levels of protein (normalized to correct for loading differences using the TOP1 as control) for the representative Western blot analysis presented; $n = 4$ independent experiments.

one of the 10 ZEB1-expressing tumors (**Fig. 2C**; Supplementary Fig. S2A). Conversely, *POLQ* expression was high in 16 of the 22 non-ZEB1–expressing TNBCs. It is worth mentioning that a single tumor sample displayed a ZEB1-positive staining in one part of the tumor with a *POLQ* negative one, whereas the opposite was observed in another part (Supplementary Fig. S2B), therefore, highlighting the mutual exclusivity of *POLQ* and ZEB1 expression in tumor presenting intratumor heterogeneity. Altogether, these results revealed an overall mutual exclusivity of ZEB1 and *POLQ* expression.

ZEB1 regulates *POLQ* expression

Given the lower expression of *POLQ* in ZEB1-expressing claudin-low cancer cell lines, compared with basal-like cell lines, we hypothesized that the ZEB1 EMT-inducing transcription factor (EMT-TF) may regulate *POLQ* expression. To address this, we first engineered ZEB1$^{-/-}$ cell lines using the CRISPR-Cas9 gene editing technology in the claudin-low MDA-MB-231 cell line (Supplementary Fig. S3A) and demonstrated that ZEB1 depletion resulted in a significant increase in

POLθ protein (**Fig. 3A**) and mRNA levels (Supplementary Fig. S3B). Second, a depletion of ZEB1 using an siRNA approach (**Fig. 3B**) over 4 days resulted in an increase in POLθ protein and mRNA levels in this model (Supplementary Fig. S3C). Equivalent results were obtained by using two ZEB1 siRNAs for depletion in the MDA-MB-231, BT-549, and SUM159 cell lines (Supplementary Fig. S3D). Conversely, over-expression of ZEB1 in the BT-20 basal-like cell line decreased the POLθ protein (**Fig. 3C**) and mRNA levels (Supplementary Fig. S3E), as well as in HCC70 cells (Supplementary Fig. S3F). Collectively, these findings suggest that *POLQ* expression is negatively regulated by ZEB1 in claudin-low cell lines. To explore whether this repression involved a direct interaction with the *POLQ* promoter, we analyzed the JASPAR database (http://jaspar.binf.ku.dk). We identified three putative ZEB1 binding sites within this region between positions −1279/−1269, −1191/−1181, and −712/−704 upstream of the transcription start site, conforming with the optimal recognition sequence of ZEB1 (CACCTG). Two putative E-boxes, binding sites for all of the EMT-TFs and other transcription factors (CANNTG), were identified
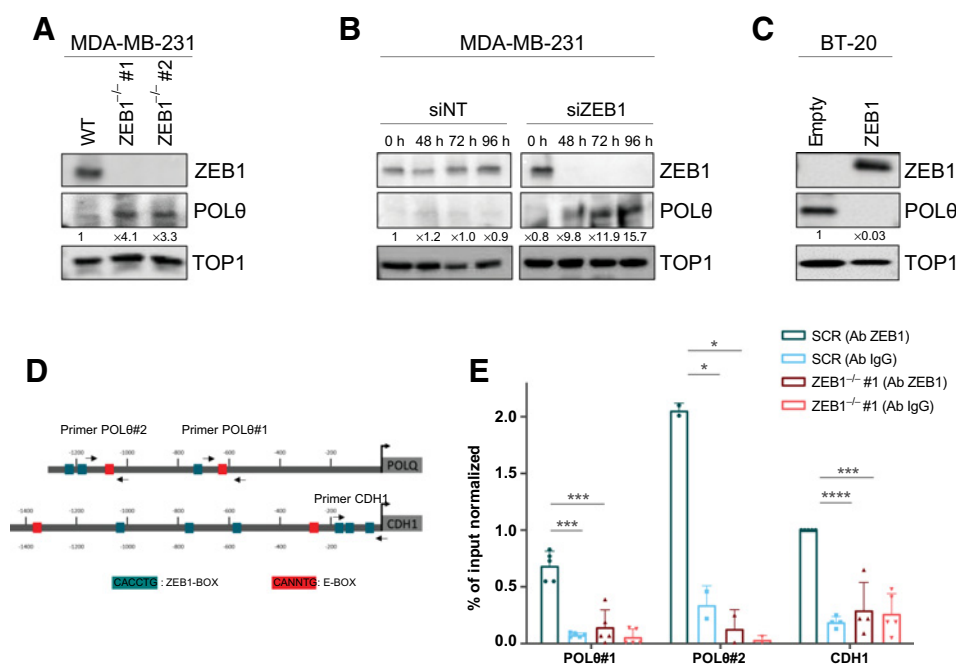
Figure 2.
ZEB1-positive tumors are usually *POLQ* negative. **A,** OncoPrint of high and low expression for *POLQ* and *ZEB1* genes. The OncoPrint provides an overview of genomic alterations (legend) in particular genes (rows) affecting particular individual samples (columns). Gene expression of *POLQ* and *ZEB1* is mutually exclusive in TNBC as evidenced by the OR. Molecular breast cancer subtype, FGA, as well as IntClust for each sample are depicted. $N = 3,083$. **B,** Representative IHC of two serial sections of a negative and a positive case of TNBC for ZEB1 (brown) and RNAscope for *POLQ* (brown dots) showing mutual exclusion in cancer cells. Scale bars, 40 μm for high magnification pictures. **C,** Distribution of *POLQ* expression staining in 10 ZEB1-positive and 22 ZEB1-negative TNBCs. *, $P < 0.05$; unpaired two-tailed Student t test with Welch correction.

at positions $-1076/-1071$ and $-662/-654$ (**Fig. 3D**; Supplementary Fig. S3G and S3H). We then performed ChIP analyses and demonstrated that ZEB1 was able to bind to the *POLQ* promoter (**Fig. 3E**) compared with ZEB1 depleted and IgG controls, as demonstrated previously for the *CDH1* promoter used as a positive control (42). To determine the significance of the ZEB1 binding to the promoter of *POLQ*, the promoter region of *POLQ* was cloned into luciferase-expressing plasmids (Supplementary Fig. S3I). The model was then validated by analyzing *CDH1* mRNA levels after knocking down ZEB1 expression in the MDA-MB-231 breast cancer cell line using siRNA (Supplementary Fig. S3J). ZEB1 depletion resulted in a significant increase in *POLQ* promoter activity for the -1280 bp construct (Supplementary Fig. S3K), suggesting a downregulation of the *POLQ* promoter by ZEB1. We then generated a -691-bp construct lacking the ZEB1 boxes of the promoter. We observed significantly less activity of the *POLQ* promoter. The stimulation following ZEB1 depletion was

Prodhomme et al.

Figure 3.
ZEB1 regulates *POLQ* expression. **A,** Immunoblot showing the protein levels of ZEB1 and POLθ in stable depletion of ZEB1 in MDA-MB-231 cells (ZEB1$^{-/-}$ #1 or #2); $n = 4$ independent experiments. **B,** Kinetics of POLθ protein level after transient knockdown with si*ZEB1*#1. Western blot analysis of noted cell types assessing expression of endogenous ZEB1 and POLθ proteins. **C,** Immunoblot showing the protein levels of ZEB1 and POLθ in stable overexpression of ZEB1 in BT-20 cells; $n = 4$ independent experiments. DNA topoisomerase I (TOP1) served as a loading control. Noted at the bottom are the mean levels of protein (normalized to correct for loading differences using the TOP1 as control) for the representative Western blot presented. $n = 4$ independent experiments. **D,** Schematic representation of the promoter region of human *POLQ* (top) and *CDH1* (bottom) genes on chromosomes 3q13.33 and 16q22.1, respectively. The potential ZEB1 boxes (ZEB1-box CACCTG in blue) and E-box (CANNTG in red) and the regions amplified after ChIP (arrows represent primer pairs) are depicted. Numbers indicate positions in bp on chromosomal DNA relative to the transcription start site (+1). **E,** qPCR analysis after ChIP for endogenously expressed ZEB1 in ZEB1-depleted (ZEB1$^{-/-}$ #1) or control (SCR) MDA-MB-231 cells, using ZEB1 antibody (Ab ZEB1) or control IgG (Ab IgG). *CDH1* was used as positive control; $n = 3$ independent experiments. Mean ± SEM. *, $P < 0.05$; ***, $P < 0.001$; ****, $P < 0.0001$; unpaired two-tailed Student t test with Welch correction.

no longer detected, suggesting a role for the −1280/−691 *POLQ* promoter in ZEB1 repression. Together, these results indicate that *POLQ* is a direct transcriptional target of ZEB1 and that ZEB1 represses *POLQ* expression.
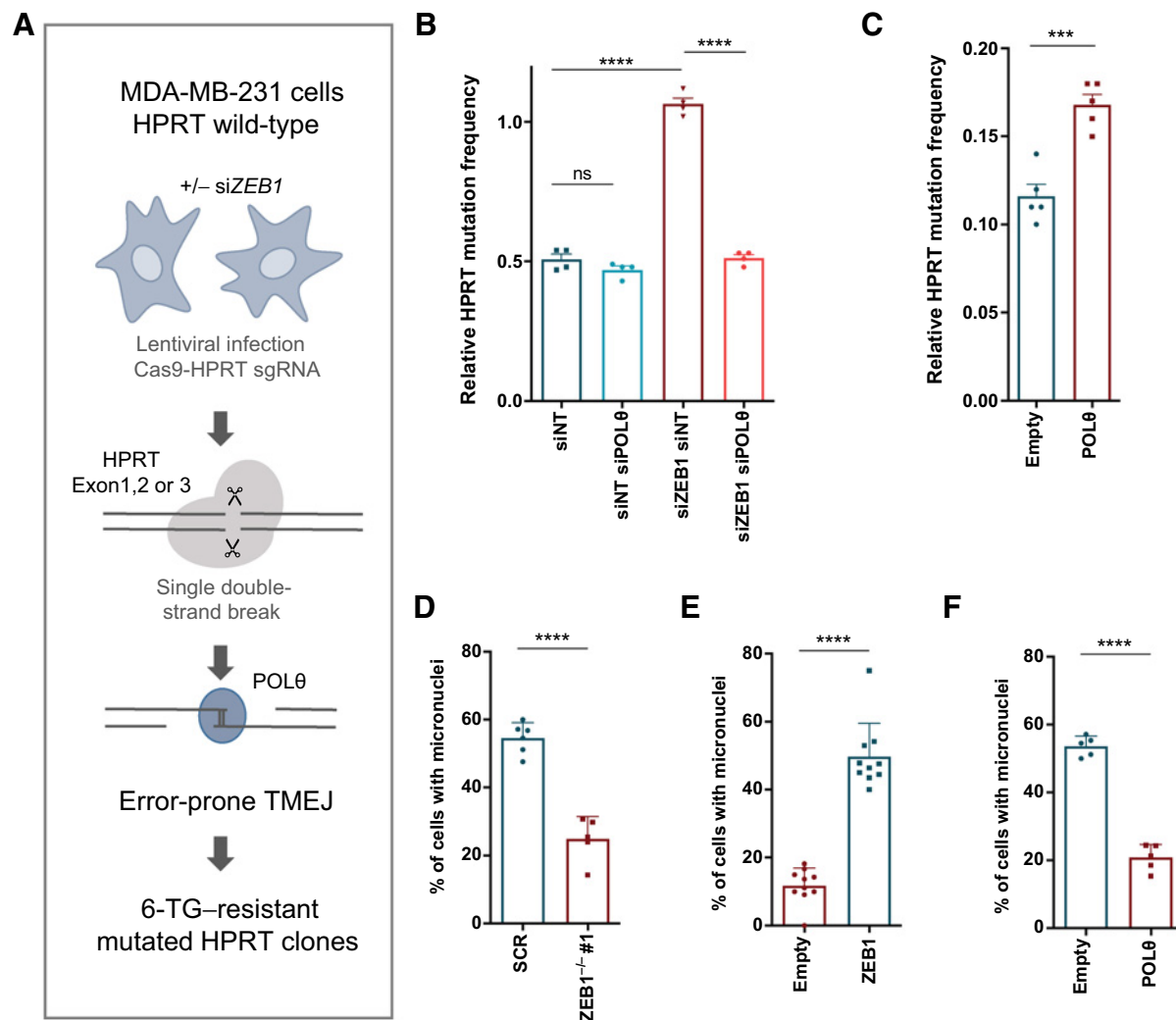
ZEB1 modulates TMEJ

We next focused on the functional interaction between ZEB1 and *POLQ* to mechanistically unravel the potential, as yet unreported, role of ZEB1 in TMEJ regulation. We directly assessed the intrinsic mutagenic TMEJ activity by evaluating the repair of a single genomic DSB in ZEB1-expressing compared with nonexpressing tumor cells. To this end, we used a selection-based assay that captures mutagenic end-joining described previously (43), in MDA-MB-231 claudin-low cell lines and in their ZEB1-depleted counterparts. Briefly, a unique site-specific DSB in the selectable *HPRT* marker gene was induced by CRISPR-Cas9 using a guided RNA directed against *HPRT* exon 2 (or exon 1 and 3 to evaluate the impact of the DSB location). HPRT enzymatic activity converts 6-TG drug into toxic nucleotides inducing cell death. Mutagenic repair of the targeted DSB in the *HPRT* gene leads to loss of HPRT protein expression or expression of an inactive HPRT protein, and renders cells resistant to 6-TG treatment (**Fig. 4A**). Thus, the frequency of *HPRT* mutations reflects the efficacy of mutagenic DSB repair and TMEJ activity. Colony-forming assay analysis revealed a significant increase in mutation frequency in

ZEB1-depleted cells (**Fig. 4B**), reflecting the increased level of POLθ (Supplementary Fig. S4A). As expected, no change in mutational frequency was observed following *POLQ* knockdown (**Fig. 4B**) because ZEB1-expressing cells displayed very low levels of POLθ. Yet, combined si*ZEB1*/si*POLQ* rescued the increase in *HPRT* mutation observed in si*ZEB1*, confirming that the HPRT assay mirrors TMEJ activity (43). A significant increase in the mutational frequency in ZEB1-depleted cells compared with wild-type cells for the DSB induced both in exon 1 and 3 (Supplementary Fig. S4B–S4E) was observed, suggesting no location effect of the single DSB generated. These data led us to propose that mutagenic repair of the DSB was alleviated in claudin-low cells because of the reduction in POLθ steady-state protein levels by ZEB1. Next, we generated an MDA-MB-231 model ectopically overexpressing POLθ and observed that it recapitulated the effects of ZEB1 depletion, that is, a significant increase in the mutational frequency (**Fig. 4C**; Supplementary Fig. S4F), further validating our hypothesis.

ZEB1-mediated micronuclei formation by *POLQ* repression

It has been reported previously in mouse models that either *Polq* mutation (44) or *Polq* depletion (45) leads to increased numbers of spontaneous micronuclei. Various molecular mechanisms contribute to micronuclei formation, including impaired DNA repair response and persistence of DSBs during mitosis associated with a defect in

Figure 4.
ZEB1 modulates TMEJ and ZEB1 mediates micronuclei formation by *POLQ* repression. **A,** Schematic outline of HPRT mutagenic test. **B,** Relative HPRT mutation frequency after the indicated siRNA treatment with Cas9-WT targeting exon 2. **C,** MDA-MB-231 cells with forced *POLQ* expression. The data shown represent the mean $\pm$ SEM; $n = 3$ independent experiments; and are expressed as a fraction of the mutation frequency observed in control cells expressing a nontargeted siRNA. **D,** Percentage of cells with micronuclei in ZEB1-depleted (ZEB1$^{-/-}$ #1) or control (SCR) MDA-MB-231 cells. **E,** BT-20 cells with forced ZEB1 expression. **F,** MDA-MB-231 cells with forced *POLQ* expression stained with Hoechst after cytochalasin B treatment. Statistical significance was calculated via unpaired two-tailed Student *t* test with Welch correction. ***, $P < 0.001$; ****, $P < 0.0001$; ns, not significant.

DNA repair pathways (46). To address the consequences of a reduction of *POLQ* expression by ZEB1 in breast cancer cells, we investigated whether ZEB1 was associated with elevated levels of micronuclei in our previously described models. Consistent with our hypothesis, more than 50% of ZEB1-expressing cells, characterized by low POLθ levels, presented micronuclei (**Fig. 4D**). Importantly, ZEB1 depletion promoted a significant decrease in the number of micronuclei (lower than 25%), while the forced expression of ZEB1 in BT-20 resulted in a significant increase in micronuclei (**Fig. 4E**), as well as in HCC70 (Supplementary Fig. S4G). Moreover, the forced expression of POLθ in ZEB1-expressing cells also led to a marked decrease in their number (**Fig. 4F**). These data suggest that the negative regulation of POLθ by ZEB1 in claudin-low breast cancer cells contributes to micronuclei formation. Collectively, these findings argue in favor of an essential role for POLθ during the error-prone

TMEJ mechanism in preserving genomic integrity at the cost of enhancing genetic alterations.

Discussion

In conclusion, we provide evidence of the mechanism of TMEJ regulation involving the EMT-inducing transcription factor, ZEB1, and show how the interplay between ZEB1 and *POLQ* may impact cancer genome stability and integrity (**Fig. 5**).

First, we highlight a relatively generalized mutual exclusivity of *POLQ* and ZEB1 expressions in TNBCs. High *POLQ* gene expression was enriched in the most genomically rearranged breast cancer subtype, IntClust10, containing HR-deficient tumors (31), confirming the previously described dependency of HR-deficient tumors on POLθ and TMEJ repair for survival in ovarian and breast tumors (14, 15).

Prodhomme et al.
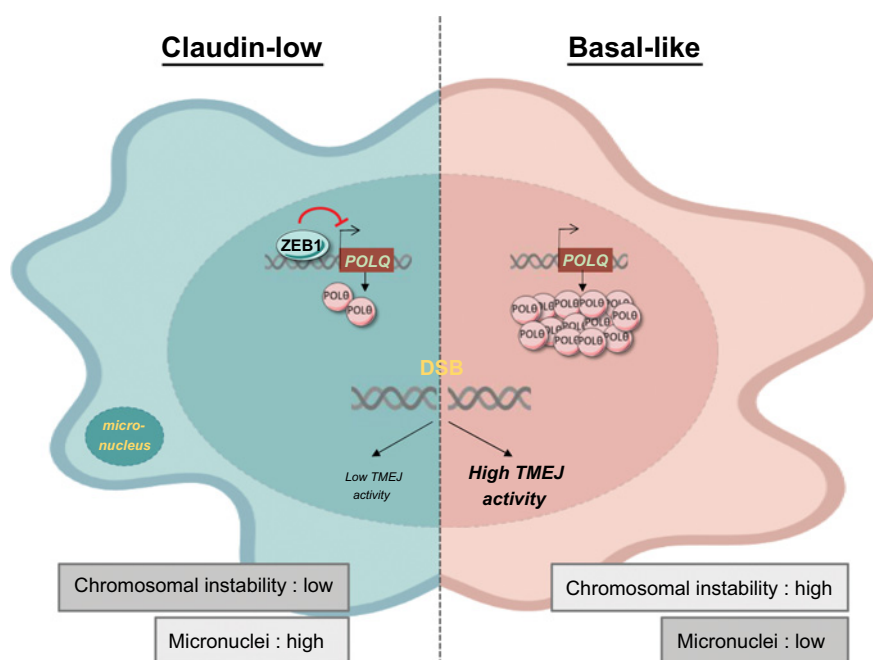
Figure 5.
Model of potential implication of TMEJ modulation by ZEB1. Among TNBC cell lines, *POLQ* expression was likely upregulated in the majority of basal-like tumor cell lines, especially in HR-deficient breast cancers recurrently depicted as highly genomically unstable. Nevertheless, some claudin-low cell lines exhibit low POLθ due to ZEB1 repression of *POLQ* gene expression. This inhibition, which we propose here modulates TMEJ activity, may participate in the low CIN observed in claudin-low tumors, but is surprisingly associated with high micronuclei in the cancer cell lines. Many mechanisms may be at the origin of micronuclei formation, but it can, however, not be excluded that low POLθ protein level, associated with low TMEJ activity, may induce residual DNA damage at the origin of micronuclei formation.

Our findings also show that some TNBCs displayed lower *POLQ* expression. Indeed, lower *POLQ* expression was observed in the IntClust4 subgroup encompassing high-ZEB1–expressing claudin-low tumors that exhibit a paucity of genomic aberrations (24, 31). Our findings supported the notion that POLθ plays a role in the onset of some type of chromosomal alterations in breast tumors. Nevertheless, not all of the chromosomal alterations can be ascribed to POLθ within TMEJ. For instance, some large deletions have been shown to arise in *Caenorhabditis elegans* genome in the absence of POLθ (47), and articles have reported that, in c-NHEJ-deficient cells, POLθ could also protect against CIN in noncancer cells (45, 48). However, instability of the genome is inherent to the great majority of human cancers (1), and POLθ is largely upregulated in highly unstable human cancers, including breast (49), ovarian (14, 50), lung, gastric, and colorectal (51). As we have shown that *POLQ* expression is highly correlated with genome instability, we suggest that the low CIN instability observed in claudin-low tumors is partly due to *POLQ* repression by ZEB1.

Second, we confirmed that micronuclei are observed in *POLQ*-deficient cells, and we showed that their number increased in a ZEB1-dependent manner. We proposed that increased micronuclei number occurred likely as a result of unrepaired break accumulation due to TMEJ defect. It is unclear how chromosomally unstable cancer cells cope with the presence of micronuclei. Surprisingly, in claudin-low tumors, the loss of genome integrity, that is, micronuclei increase, seems to preserve genome stability. Nevertheless, these observations are consistent with the concept that ZEB1 acts in cancer progression by protecting the genome against instability. We have previously shown that ZEB1 withstands an oncogenic activation by driving an antioxidant program, leading to a process of malignant transformation in the absence of exacerbated genomic instability (24). Similarly, it was suggested that ZEB1 is required for HR-mediated DNA damage repair and the clearance of DNA breaks (25). We report here that ZEB1 prevents the highly mutagenic TMEJ in claudin-low cells displaying micronuclei. In cancer cells, TMEJ would, therefore, be a full-fledged repair pathway, that

is, a factor assisting cancer cell survival similarly to replication stress (52). As suggested previously (24), ZEB1 may foster plasticity through cell adaptability rather than genomic variation. Future approaches need to address how TMEJ thus participates in genome integrity in human cells.

Finally, to further our understanding of the biological complexity of claudin-low tumors and ultimately improve the outcomes of patients with breast cancer, our data may have clinical implications. The TMEJ downregulation that we characterized in claudin-low cells, may trigger a compensatory increase in the c-NHEJ. Although this hypothesis remains to be tested, assessment of the sensitivity of claudin-low tumors to NHEJ inhibitors may be a strategy for treating patients with claudin-low cancers.

Authors' Disclosures

No disclosures were reported.

Authors' Contributions

M.K. Prodhomme: Conceptualization, resources, data curation, formal analysis, validation, investigation, visualization, methodology, writing–original draft, writing–review and editing. **R.M. Pommier:** Resources, data curation, software, formal analysis, validation, writing–original draft, writing–review and editing. **C. Franchet:** Data curation, formal analysis, visualization, methodology, writing–original draft. **F. Fauvet:** Resources, data curation, formal analysis, validation. **V. Bergoglio:** Conceptualization, resources, validation, methodology, writing–original draft. **P. Brousset:** Conceptualization, resources, supervision. **A.-P. Morel:** Resources, supervision, funding acquisition, writing–original draft, project administration. **A.-C. Brunac:** Data curation, formal analysis, visualization, methodology, writing–original draft. **M. Devouassoux-Shisheboran:** Resources, data curation, formal analysis, validation. **V. Petrilli:** Resources, formal analysis, validation, methodology. **C. Moyret-Lalle:** Conceptualization, supervision, validation, writing–original draft, writing–review and editing. **J.-S. Hoffmann:** Conceptualization, resources, supervision, validation, writing–original draft, writing–review and editing. **A. Puisieux:** Conceptualization, formal analysis, supervision, funding acquisition, validation, investigation, writing–original draft, project administration, writing–review and editing. **A. Tissier:** Conceptualization, formal analysis, supervision, funding acquisition, validation, investigation, methodology, writing–original draft, writing–review and editing.

Acknowledgments

References

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011; 144:646–74.
2. Halazonetis TD, Gorgoulis VG, Bartek J. An oncogene-induced DNA damage model for cancer development. Science 2008;319:1352–5.
3. Ceccaldi R, Rondinelli B, D'Andrea AD. Repair pathway choices and consequences at the double-strand break. Trends Cell Biol 2016;26:52–64.
4. O'Driscoll M, Jeggo PA. The role of double-strand break repair—insights from human genetics. Nat Rev Genet 2006;7:45–54.
5. Scully R, Panday A, Elango R, Willis NA. DNA double-strand break repair-pathway choice in somatic mammalian cells. Nat Rev Mol Cell Biol 2019;20: 698–714.
6. Seol J-H, Shim EY, Lee SE. Microhomology-mediated end joining: good, bad and ugly. Mutat Res 2018;809:81–7.
7. Schimmel J, van Schendel R, den Dunnen JT, Tijsterman M. Templated insertions: a smoking gun for polymerase theta-mediated end joining. Trends Genet 2019;35:632–44.
8. Patterson-Fortin J, D'Andrea AD. Exploiting the microhomology-mediated end-joining pathway in cancer therapy. Cancer Res 2020;80:4593–600.
9. Brambati A, Barry RM, Sfeir A. DNA polymerase theta (Polθ)–an error-prone polymerase necessary for genome stability. Curr Opin Genet Dev 2020; 60:119–26.
10. Simsek D, Jasin M. Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4-ligase IV during chromosomal translocation formation. Nat Struct Mol Biol 2010;17:410–6.
11. Chiarle R, Zhang Yu, Frock RL, Lewis SM, Molinie B, Ho Yu-J, et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. Cell 2011;147:107–19.
12. Hwang T, Reh S, Dunbayev Y, Zhong Yi, Takata Y, Shen J, et al. Defining the mutation signatures of DNA polymerase θ in cancer genomes. NAR Cancer 2020;2:zcaa017.
13. Carvajal-Garcia J, Cho J-E, Carvajal-Garcia P, Feng W, Wood RD, Sekelsky J, et al. Mechanistic basis for microhomology identification and genome scarring by polymerase theta. Proc Natl Acad Sci U S A 2020;117:8476–85.
14. Ceccaldi R, Liu JC, Amunugama R, Hajdu I, Primack B, Petalcorin MIR, et al. Homologous-recombination-deficient tumours are dependent on Polθ-mediated repair. Nature 2015;518:258–62.
15. Mateos-Gomez PA, Gong F, Nair N, Miller KM, Lazzerini-Denchi E, Sfeir A. Mammalian polymerase θ promotes alternative NHEJ and suppresses recombination. Nature 2015;518:254–7.
16. Seki M, Wood RD. DNA polymerase θ (POLQ) can extend from mismatches and from bases opposite a (6-4) photoproduct. DNA Repair 2008; 7:119–27.
17. Wood RD, Doublié S. DNA polymerase θ (POLQ), double-strand break repair, and cancer. DNA Repair 2016;44:22–32.
18. Perou CM. Molecular stratification of triple-negative breast cancers. Oncologist 2010;5:39–48.
19. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. Mol Oncol 2011;5:5–23.
20. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res 2010;12:R68.
21. Puisieux A, Brabletz T, Caramel J. Oncogenic roles of EMT-inducing transcription factors. Nat Cell Biol 2014;16:488–94.
22. Wellner U, Schubert J, Burk UC, Schmalhofer O, Zhu F, Sonntag A, et al. The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. Nat Cell Biol 2009;11:1487–95.
23. Morel A-P, Hinkal GW, Thomas C, Fauvet F, Courtois-Cox S, Wierinckx A, et al. EMT inducers catalyze malignant transformation of mammary epithelial cells and drive tumorigenesis towards claudin-low tumors in transgenic mice. PLos Genet 2012;8:e1002723.
24. Morel A-P, Ginestier C, Pommier RM, Cabaud O, Ruiz E, Wicinski J, et al. A stemness-related ZEB1-MSRB3 axis governs cellular pliancy and breast cancer genome stability. Nat Med 2017;23:568–78.
25. Zhang P, Wei Y, Wang Li, Debeb BG, Yuan Y, Zhang J, et al. ATM-mediated stabilization of ZEB1 promotes DNA damage response and radioresistance through CHK1. Nat Cell Biol 2014;16:864–75.
26. R Core Team. R: a language and environment for statistical computing. Available at: http://www.R-project.org/.
27. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016;32:2847–9.
28. Curtis C, Shah SP, Chin S-F, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012;486:346–52.
29. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007;8: 118–27.
30. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 2012;28:882–3.
31. Pommier RM, Sanlaville A, Tonon L, Kielbassa J, Thomas E, Ferrari A, et al. Comprehensive characterization of claudin-low breast tumors reflects the impact of the cell-of-origin on cancer evolution. Nat Commun 2020;11:3431.
32. Scrucca L, Fop M, Murphy TB, Raftery AE. mclust 5: clustering, classification and density estimation using gaussian finite mixture models. The R Journal 2016;8: 289–317.
33. Gendoo DMA, Ratanasirigulchai N, Schröder MS, Paré L, Parker JS, Prat A, et al. Genefu: an R/Bioconductor package for computation of gene expression-based signatures in breast cancer. Bioinformatics 2016;32:1097–9.
34. Haibe-Kains B, Desmedt C, Loi S, Culhane AC, Bontempi G, Quackenbush J, et al. A three-gene model to robustly identify breast cancer molecular subtypes. J Natl Cancer Inst 2012;104:311–25.
35. Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, et al. Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. Breast Cancer Res 2008;10:R65.
36. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009;27:1160–7.
37. Hu Z, Fan C, Oh DS, Marron Js, He X, Qaqish BF, et al. The molecular portraits of breast tumors are conserved across microarray platforms. BMC Genomics 2006;7:96.
38. Paquet ER, Hallett MT. Absolute assignment of breast cancer intrinsic molecular subtype. J Natl Cancer Inst 2015;107:357.
39. Fernandez-Vidal A, Guitton-Sert L, Cadoret J-C, Drac M, Schwob E, Baldacci G, et al. A role for DNA polymerase Î in the timing of DNA replication. Nat Commun 2014;5:4285.
40. Dawson S-JJ, Rueda OM, Aparicio S, Caldas C. A new genome-driven integrated classification of breast cancer and its implications. EMBO J 2013;32:617–28.
41. Wang Zi, Song Y, Li S, Kurian S, Xiang R, Chiba T, et al. DNA polymerase theta (POLQ) is important for repair of DNA double-strand breaks caused by fork collapse. J Biol Chem 2019;294:3909–19.
42. Cieply B, Farris J, Denvir J, Ford HL, Frisch SM. Epithelial-mesenchymal transition and tumor suppression are controlled by a reciprocal feedback loop between ZEB1 and Grainyhead-like-2. Cancer Res 2013;73:6299–309.
43. Schimmel J, Kool H, van Schendel R, Tijsterman M. Mutational signatures of non-homologous and polymerase theta-mediated end-joining in embryonic stem cells. EMBO J 2017;36:3634–49.
44. Shima N, Munroe RJ, Schimenti JC. The mouse genomic instability mutation chaos1 is an allele of POLQ that exhibits genetic interaction with atm. Mol Cell Biol 2004;24:10381–9.

45. Yousefzadeh MJ, Wyatt DW, Takata K-I, Mu Y, Hensley SC, Tomida J, et al. Mechanism of suppression of chromosomal instability by DNA polymerase POLQ. PLos Genet 2014;10:e1004654.

46. Ye CJ, Sharpe Z, Alemara S, Mackenzie S, Liu G, Abdallah B, et al. Micronuclei and genome chaos: changing the system inheritance. Genes 2019;10:3661. DOI: 10.3390/genes10050366.

47. Koole W, van Schendel R, Karambelas AE, van Heteren JT, Okihara KL, Tijsterman M. A Polymerase theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat Commun 2014;5:3216.

48. Wyatt DW, Feng W, Conlin MP, Yousefzadeh MJ, Roberts SA, Mieczkowski P, et al. Essential roles for polymerase θ-mediated end joining in the repair of chromosome breaks. Mol Cell 2016;63:662–73.

49. Lemee F, Bergoglio V, Fernandez-Vidal A, Machado-Silva A, Pillaire M-J, Bieth A, et al. DNA polymerase up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. Proc Natl Acad Sci U S A 2010;107:13390–5.

50. Konstantinopoulos PA, Ceccaldi R, Shapiro GI, D'Andrea AD. Homologous recombination deficiency: exploiting the fundamental vulnerability of ovarian cancer. Cancer Discov 2015;5:1137–54.

51. Kawamura K, Bahar R, Seimiya M, Chiyo M, Wada A, Okada S, et al. DNA polymerase theta is preferentially expressed in lymphoid tissues and upregulated in human cancers. Int J Cancer 2004;16:9–16.

52. Feng W, Simpson DA, Carvajal-Garcia J, Price BA, Kumar RJ, Mose LE, et al. Genetic determinants of cellular addiction to DNA polymerase theta. Nat Commun 2019;10:4286.

Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

AACR American Association for Cancer Research

EMT Transcription Factor ZEB1 Represses the Mutagenic POLθ -Mediated End-Joining Pathway in Breast Cancers

Mélanie K. Prodhomme, Roxane M. Pommier, Camille Franchet, et al.

Updated version	Access the most recent version of this article at: doi:10.1158/0008-5472.CAN-20-2626
Supplementary Material	Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2020/11/25/0008-5472.CAN-20-2626.DC1

Cited articles	This article cites 51 articles, 11 of which you can access for free at: http://cancerres.aacrjournals.org/content/81/6/1595.full#ref-list-1
Citing articles	This article has been cited by 1 HighWire-hosted articles. Access the articles at: http://cancerres.aacrjournals.org/content/81/6/1595.full#related-urls

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/81/6/1595. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.