AWARD NUMBER: W81XWH-17-1-0239

TITLE: DNA Polymerase Zeta Inactivation in Prostate Cancer

PRINCIPAL INVESTIGATOR: Richard Wood, Ph.D.

CONTRACTING ORGANIZATION: The University of Texas MD Anderson Cancer Center
Houston, TX

REPORT DATE: November 2021

TYPE OF REPORT: Final Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
November 2021	Final	15Jul2017 – 14Jul2021

4. TITLE AND SUBTITLE

DNA Polymerase Zeta Inactivation in Prostate Cancer

5a. CONTRACT NUMBER
W81XWH-17-1-0239

5b. GRANT NUMBER
GRANT12258091

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Richard Wood, Ph.D.

E-Mail: rwood@mdanderson.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The University of Texas MD
Anderson Cancer Center
1515 Holcombe Blvd., Unit 116
Houston, TX 77030-4009

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Not all prostate cancer patients respond in the same way to therapies. For example, some cancers respond well to hormone therapies, and others to radiation therapy. A major reason for these differences is that different genetic changes underlie individual cancers. In order to personalize therapy and make it much more effective, it is important to take advantage of genetic analyses and determine, as early as possible during treatment, the therapeutic strategies that will be most effective to cure or control the cancer. Although it is the most common cancer in American men, more than a quarter of primary prostate cancers of both good and poor clinical prognosis are driven by unknown molecular changes in the genome. Recently, in the course of our studies of DNA repair, we analyzed prostate cancer genome data and discovered that the gene for an important DNA repair enzyme called DNA polymerase zeta (abbreviated "pol zeta") is deleted in 13% of primary prostate cancers. This is very significant because identification of cancers with deletion of the pol zeta gene is highly likely to be useful in diagnosis and therapy. This is because suppressing pol zeta sensitizes cells to DNA damage. The absence of pol zeta is likely very important for improving therapy in these cancers, but it has never been investigated.

15. SUBJECT TERMS

pol ζ, pol δ, pol ϵ, pol η, TAG, MMC, ICL, HR, EJ WT, Het, KO, Puro, µg, µl, IHC, STR, ATCC, NOD-SCID

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	96	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

1. INTRODUCTION:

The major objective of the research is to develop the idea that prostate cancers with pol zeta deletions are specifically sensitive to therapeutic DNA damaging agents, including radiation. This could lead to individualized treatment of an important group of prostate cancers. It will also indicate the usefulness of DNA damaging chemotherapy for a previously unrecognized major group of prostate cancers. Because we know that normal cells do not grow well in the absence of pol zeta, we also intend in this research to identify genetic alterations that allow cells to proliferate in the absence of pol zeta. This will be a practically important advance because it will help identify the pol zeta-deleted class of cancers.

2. KEYWORDS:

DNA polymerase, DNA repair, mitomycin C, cisplatin, radiation, cell lines, gene deletion

3. ACCOMPLISHMENTS:

What were the major goals of the project?

Specific Aim 1: Determine the DNA damage sensitivity conferred by disruption of pol ζ in prostate cancer cells		
Major Task 1: *REV3L* will be inactivated by targeted genetic deletion in prostate cancer cell lines and specific mutations will be tested	Months	% comple-tion
Inactivate REV3L	1-6	100
Toxicity measurements	7-12	100
Make specific mutations in cDNA	7-12	100
Complementation assays using mouse MEFs	13-20	100

Specific Aim 2: Identify suppressor mutations that allow cells to proliferate in the absence of pol ζ		
Major Task 1: Candidate suppressor genes suggested by preliminary studies will be tested in REV3L-deficient human prostate cancer cells.	Months	MDA
Make targeted deletions in cell lines	1-12	70
Toxicity measurements	6-18	100

Major Task 2: Identify genes that, when downregulated, alleviate the growth defects in REV3L-defective human cancer cells.		
Subtask 1: Make targeted deletions in cell lines and measure growth rates	1-12	60

Subtask 2: A genome-wide shRNA screen for growth of REV3L-defective cells	12-24	20

Specific Aim 3: Determine radiation and chemosensitivity of a pol ζ-defective prostate cancer model in mice		
Major Task 1: Establish xenograft model and determine the response to ionizing radiation and cisplatin-based treatment will be quantified.	Months	MDA
Subtask 1: Obtain mice and establish xenografts	6-18	0
Subtask 2: Test drug and radiation resistance of xenografts	9-20	0

What was accomplished under these goals?

Major Activities: The major activities described below under "Key Outcomes" and also in the publications provided with the Appendix.

Specific Objectives: Identifying inactivation of *REV3L* in primary prostate cancers, and the reasons enabling tolerance of pol ζ inactivation were the primary objects of research.

A major objective of **Specific Aim 1** was to test the functionality and consequence of cancer-associated mutations in REV3L, the catalytic subunit of DNA polymerase ζ. We focused on measurement of sensitivity to the chemotherapeutic agent cisplatin, which produces lesions including DNA interstrand crosslinks. Cell lines were generated in REV3L KO cell lines REV3L 4(-/Cre)TAg MEF and REV3L 3(+/Cre)TAg MEF. Cancer associated mutations in REV3L included the planned P2744S and R2523C and the more recently reported R187W mutation found in a cohort study of 40 Spanish families with colorectal cancer. The parental cell lines are REV3L 4(-/Cre)TAg MEF and REV3L 3(+/Cre)TAg MEFs. Successfully transfected clones are isolated by fluorescence activated flow-sorting. They are then confirmed by checking plasmid integration using PCR and immunoblotting.

We constructed the following complemented MEF cell lines, several independent clones of each:

pCDH-FH-TR4-2 ASM REV3L KO
pCDH-FH-TR4-2 R2523C REV3L KO
pCDH-FH-TR4-2 P2744S REV3L KO
pCDH-FH-TR4-2 R187W REV3L KO
pCDH-FH-TR4-2 4A REV3L KO

Methodology:
To test sensitivity to chemical DNA damaging agents, the immortalized MEFs were plated into white 96-well plates (immortalized MEFs– 5,000 cells/well). The following day, various concentrations of cisplatin (Sigma) were added to the wells, and the cells were incubated for 48 hr. Then the cells were lysed, a reagent was added that emits light in the presence of ATP (ATPLite One Step, Perkin Elmer), and luminescence was measured using a plate reader (Biotek Synergy II).

The luminescence measurement was normalized to undamaged control. The ATP content measured by luminescence provides a measure of survival.

Significant Results:

Aim 1: Representative survival results were summarized in earlier reports. The cancer-associated R2523C, P2744S, and R187W mutations did not affect the functionality of the TR4-2 REV3L. However, it is significant that inactivation of the four REV7 binding sites in REV3L (in the 4A mutant) was sufficient to inactivate REV3L function (**Fig 1A**). To obtain a readout of genetic instability in REV3L-defective cells for future use in assessing REV3L defects, we assayed micronuclei, which are diagnostic of unresolved double strand breaks in chromosomes (**Fig 1D**).

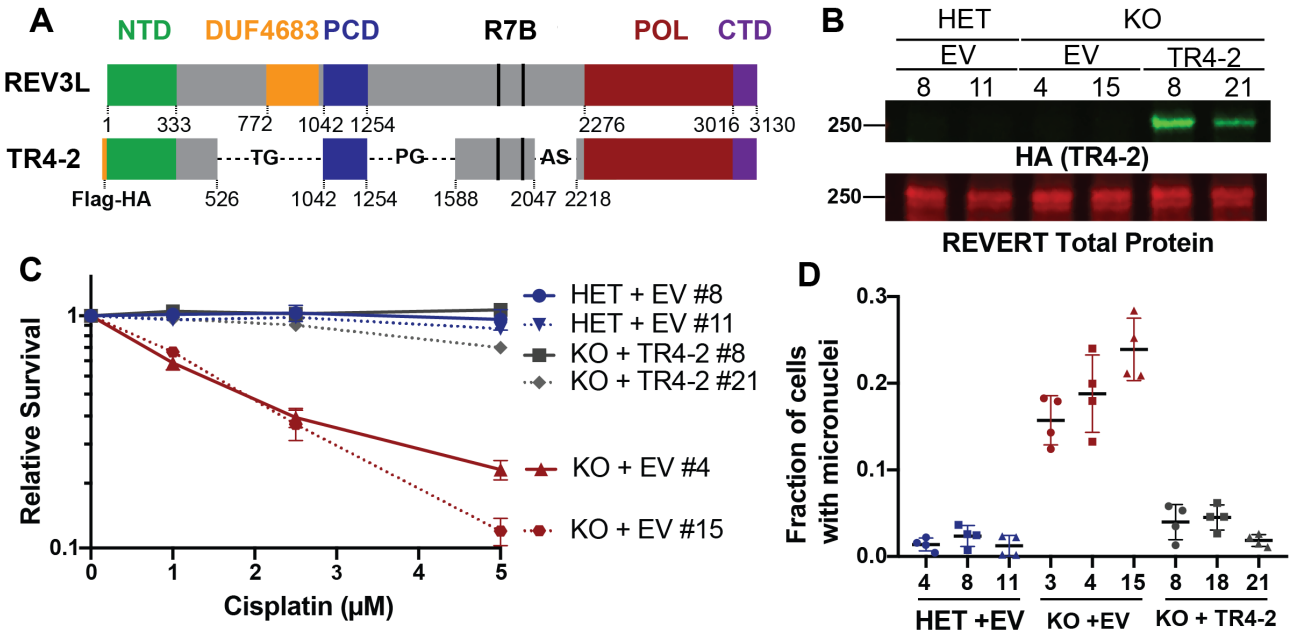


Figure 1: Shortened REV3L construct rescues phenotypes of pol ζ disruption
A) Schematic of full-length human REV3L and human REV3L construct, TR4-2. TR4-2 retains most conserved domains and binding sites of REV3L including the regions that coordinate interactions with the accessory subunits of pol ζ, the C-terminal domain (CTD, purple) the two REV7 binding sites (R7B, black), and a positively charged domain (PCD) of uncertain function. The B-family catalytic core is formed by folding of the N-terminal domain (NTD, green) and the polymerase domain (POL, red). B) Immunoblot with HA antibody showing stable expression of TR4-2 with an N-terminal Flag-HA tag in *Rev3l* KO MEF clones. C) Stable expression of TR4-2 in *Rev3l* KO MEF clones reverses hypersensitivity to cisplatin. MEFs were exposed to the indicated cisplatin concentrations for 48 h and relative survival was quantified with the ATPlite assay. D) Stable expression of TR4-2 decreases micronuclei formation in unchallenged *Rev3l* KO MEF clones.

Aim 2: A major objective of Specific Aim 2 is to construct REV3L knockout cell line(s) in prostate cancer cells. As described in last year's report, we did extensive screening of Crispr-Cas9 targeted clones in DU145 and PC3 prostate cancer cell cells. None of the many candidate clones had full REV3L deletion. The inviability of these cells following REV3L deletion was, in some ways

expected according to our current hypothesis, because neither cell line harbors the large deletion on chromosome 6q that encompasses REV3L and surrounding genes. We also cultured REV3L knockout cells from the human Jurkat cell line for a total 12 months to attempt to find clones that regained cell growth speed sufficient for CRISPR-Cas9 screening, but we did not find such cell lines.

These results show that there is very strong selection to retain REV3L function in primary cells and in most cancer cells. We have been careful in our genetic analysis, in contrast to some other reports in the field. For example, a recently published manuscript describes REV3L knockout cells from human HEK293T cells (Su et al 2019, J. Biol Chem. PMID: 30842261 10.1074/jbc.RA119.007925). We examined the deleted DNA sequences given in this paper. They do not prove that both alleles of REV3L are inactivated. One of the targeted mutations eliminates a splice site and part of an intron, leaving open the strong possibility that alternative splicing may permit *REV3L* function in that cell line.

Loss of polymerase ζ alters the transcriptome

In order to uncover the type of stress occurring in cells lacking pol ζ, we performed genome-wide mRNA sequencing on a controlled set of immortalized clones: *Rev3l* HET + empty vector, *Rev3l* KO + empty vector and *Rev3l* KO + TR4-2. To focus on major changes, we set a strict threshold (> |2| $\log_2$ fold change and FDR < 0.05). Expression analysis of 17,346 mapped transcripts revealed that 1117 transcripts were either upregulated or downregulated in the *Rev3l* KO + empty vector relative to the *Rev3l* HET + empty vector MEFs (Fig 2A). The majority (~68%) were upregulated (Fig 2A). These upregulated or downregulated genes displayed no statistically significant enrichment or depletion for DNA replication or canonical DNA damage sensing pathways. This is not completely unexpected given that p53 promotes much of the transcriptional response to DNA damage, while our MEFs have inactivated p53 due to large T-antigen immortalization. In these immortalized high passage cells, we are likely to observe stable transcriptional alterations, rather than an acute response.

Instead, the upregulated genes displayed an enrichment in immune system-related pathways, as revealed by gene ontology analysis (Fig 2B). Upstream regulator analysis was used to analyze all differentially expressed genes. This revealed that the alterations in the transcriptome are consistent with activation of positive regulators of the interferon response, including key transcription factors in this pathway, IRF3 and IRF7 (Fig 2C). Importantly the predicted activation of IRF3 and IRF7 was reversed by expression of the *Rev3l* TR4-2 cDNA, showing that these results stem from a function of *Rev3l.* The same trends were also observed by applying a substantially lower threshold ($\log_2$ FC > |0.5|) for differentially expressed genes. This increased the dataset to 2071 differentially expressed transcripts. Pathway analysis showed a negative correlation with predicted TRIM24 activation. TRIM24 suppresses interferon-stimulated gene expression [16], confirming that our data is consistent with expression of interferon stimulated genes.

To explore whether our data set is in fact consistent with an interferon response, we analyzed a curated set of 25 known interferon stimulated genes and observed increased expression in the *Rev3l* KO relative to the *Rev3l* HET MEFs, which was partially reduced by TR4-2 stable expression (Fig 2D). Together our results reveal that disruption of pol ζ promotes induction of interferon-stimulated genes.

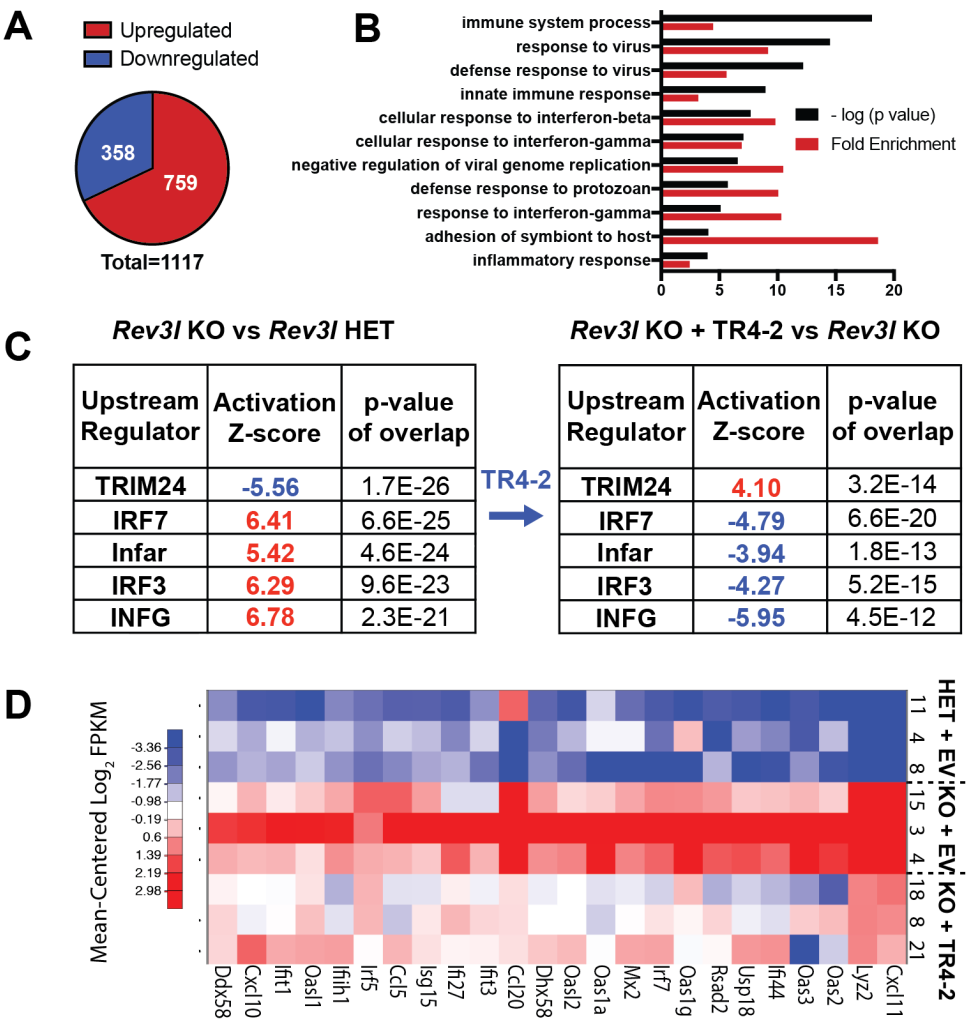


Figure 2: Cells deficient in DNA polymerase ζ have an altered transcriptome.

A) Differentially expressed genes in *Rev3l* KO + empty vector relative to *Rev3l* HET + empty vector using the threshold of a $\log_2$ fold change of > |2| and a false discovery rate of < 0.05. B) Top 10 GO (Gene Ontology) terms reveal enrichment of immune system related genes in upregulated genes in *Rev3l* KO MEFs. C) Upstream regulator analysis reveals the data set is consistent with predicted activation of positive regulators of an interferon response in *Rev3l* KO MEFs. D) Heatmap of a set of 25 known interferon-stimulated genes.

Disruption of *Rev3l* induces interferon-stimulated gene expression driven by the cGAS-STING axis

Given that our complemented cell lines were generated using lentivirus constructs and grown continually under selection, we moved to validate the results in the parental *Rev3l* KO and *Rev3l* HET MEF cell lines and one additional set of cell lines to limit extraneous variables. We confirmed an increase in mRNA expression of specific interferon stimulated genes in the *Rev3l* KO MEFs relative to the control cell lines, including key chemokines (Fig 3A).

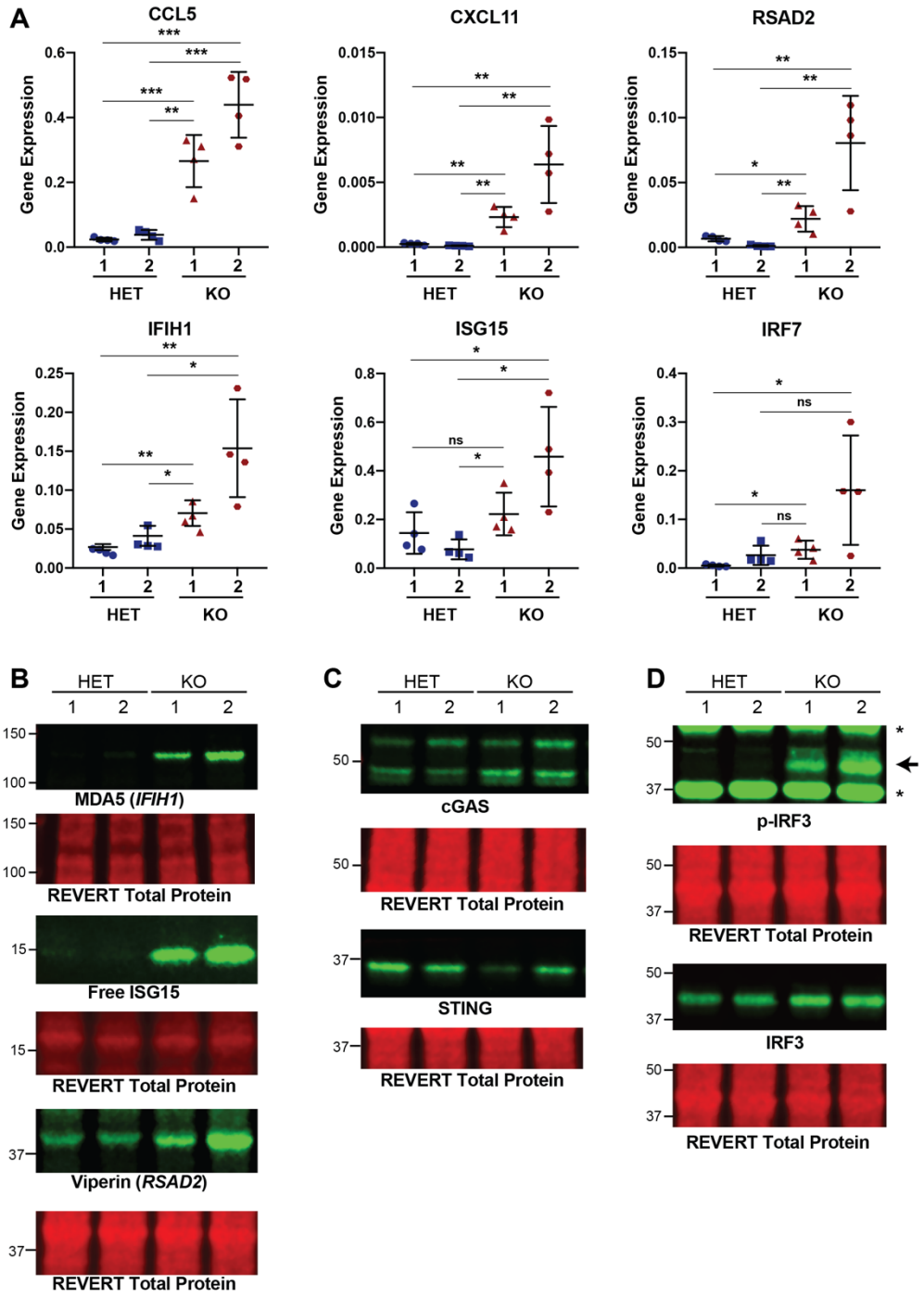


Figure 3: Disruption of *Rev3l* results in increased expression of interferon stimulated genes and proteins. A) Gene expression ($2^{-\Delta Ct}$) of selected interferon stimulated genes normalized to HPRT detected by qRT-PCR. Error bars represent standard deviation. Unpaired student t-test, * = $p < 0.05$, ** = $p < 0.01$ and *** = $p < 0.001$. B) Immunoblots showing increased protein levels of interferon stimulated genes, MDA5 and ISG15. C) Immunoblot showing presence of components of the innate immune system, cGAS and STING, in MEFs, with reduced STING in pol ζ knockout cells. D) Enhanced phosphorylation of S888 of IRF3 (corresponding to S396 in humans) in *Rev3l* KO MEFs.

In order to extend these findings to the protein level, we examined interferon stimulated gene products by immunoblotting. Corresponding to an increase in mRNA levels, we also observed an increase in protein levels of known interferon stimulated genes, MDA5 (encoded by the *IFIH1* gene), ISG15, and viperin (encoded by the *RSAD2* gene) (Fig 3B). Together these data indicate that an interferon branch of the innate immune system may be activated due to disruption of pol ζ function. Since it seems unlikely that pol ζ plays a direct role in transcriptional regulation, the next obvious question is how and why loss of pol ζ induces the expression of interferon stimulated genes.

The major consequence of pol ζ disruption in unchallenged mammalian cells is increased genomic instability as evidenced by multiple markers including γ-H2AX foci, chromosome fragmentation and aberrations, and micronuclei (Fig 1D). Therefore, it seems likely that this transcriptional response ultimately stems from the vast genomic damage induced by loss of pol ζ function. Consistent with this hypothesis, the innate immune system not only can recognize and mount an interferon response to foreign DNA, but also can respond to endogenous DNA that has escaped from the nucleus. In some instances, this response can halt cell growth providing organisms to shut down propagations of virally infected cells and cells with dangerously fragmented genomes.

Mammalian cells have a host of cytosolic nucleic acid sensors that patrol the cytosol for DNA. One of these, cGAS, is increasingly recognized to be of paramount importance in the induction of an interferon response to both exogenous and endogenous cytosolic DNA. When cGAS binds to double stranded DNA, it activates its enzymatic activity and results in the production of cGAMP, a cyclic dinucleotide. cGAMP binds to the STING receptor on the membrane of endoplasmic reticulum, resulting in activation of kinases including TBK1 which can in turn phosphorylate and activate IRF3, a central transcription factor in the interferon response.

Given that cGAS-STING axis has been implicated specifically in responding to endogenous DNA damage and has been correlated with micronuclei formation, we asked whether cGAS-STING promotes the induction of expression of interferon stimulated genes due to loss of function of pol ζ. Consistent with most MEFs having a functional innate immune system, both *Rev3l* KO and HET MEF cell lines expressed both cGAS and STING (Fig 3C). We noted a decrease in STING expression in *Rev3l* KO MEFs, which is consistent with a constitutive activation of the cGAS-STING pathway, as cGAS activation leads to a negative feedback loop resulting in STING degradation [19,20]. Importantly, we detected an increase of IRF3 phosphorylated at S888 (corresponding to S396 in humans) indicative of IRF3 activation in *Rev3l* KO MEFs (Fig 3D).

This led us to investigate if cGAS-STING drives expression of interferon-stimulated genes upon loss of pol ζ. Knockdown of either cGAS or STING significantly reduced the mRNA expression of selected interferon stimulated genes as well as the protein levels (Fig 4A-D). In addition, depletion of cGAS or STING in *Rev3l* KO MEFs markedly reduced S888 phosphorylation of IRF3 (Fig 4E). Together this indicates that disruption of pol ζ function promotes activation of an innate immune response driven by the cGAS-STING axis.

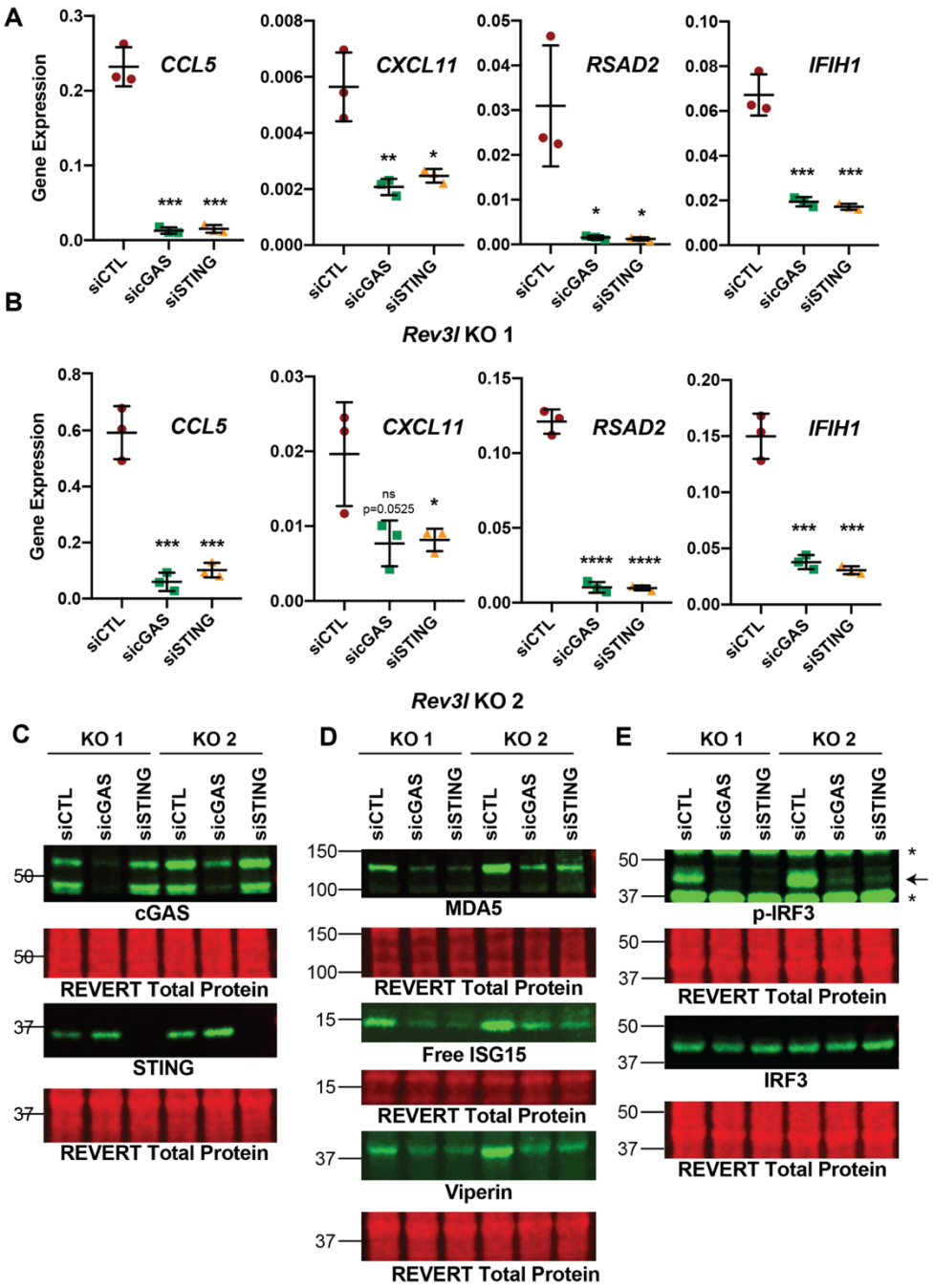


Figure 4: The cGAS-STING axis promotes expression of interferon stimulated genes due to loss of pol ζ function. A) Knockdown of cGAS or STING reduces mRNA expression of *CCL5, CXCL11, RSAD2* (which encodes Viperin protein), *IFIH1* (which encodes MDA5) as detected by qRT-PCR in *Rev3l* KO 1. Gene expression ($2^{-\Delta Ct}$) of selected interferon stimulated genes normalized to HPRT detected by qRT-PCR. Error bars represent standard deviation. Unpaired student t-test, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, and **** = $p < 0.0001$. B) Same as in A except with *Rev3l* KO 2. C) Efficient knockdown of cGAS or STING protein levels. D) MDA5, ISG15, and Viperin protein levels decrease with cGAS and STING knockdown. E) Phosphorylation of S888 in mouse (analogous to the human S396) of IRF3 in *Rev3l* KO MEFs decrease with knockdown of cGAS and STING.

An innate immune response cause by loss of pol ζ function

Pol ζ stands apart from the other translesion polymerases in that it is required for mammalian development and proliferation of primary cells. Now we can add that in addition to activating p53-dependent responses, disruption of pol ζ function invokes a prominent innate immune response promoted by the cGAS-STING pathway.

It is remarkable that disruption of an enzyme commonly thought of as a specialized transleion synthesis polymerase can lead to a constitutive innate immune response. Recently, cells with loss of function of key DNA repair enzymes, RNaseH2, BRCA2, and BLM have been shown to have an elevated cGAS-STING response that correlates with an increase in micronuclei that colocalize with cGAS. There are several sources of DNA damage that may give rise to a sustained response including cytosolic mitochondrial DNA and cytosolic DNA arisen from stalled and processed replication forks. DNA stress may continually arise from likely collapse of DNA replication forks in the absence of pol ζ, which could promote formation of micronuclei and also release small fragments of DNA. Further, some nuclear genes including pol ζ

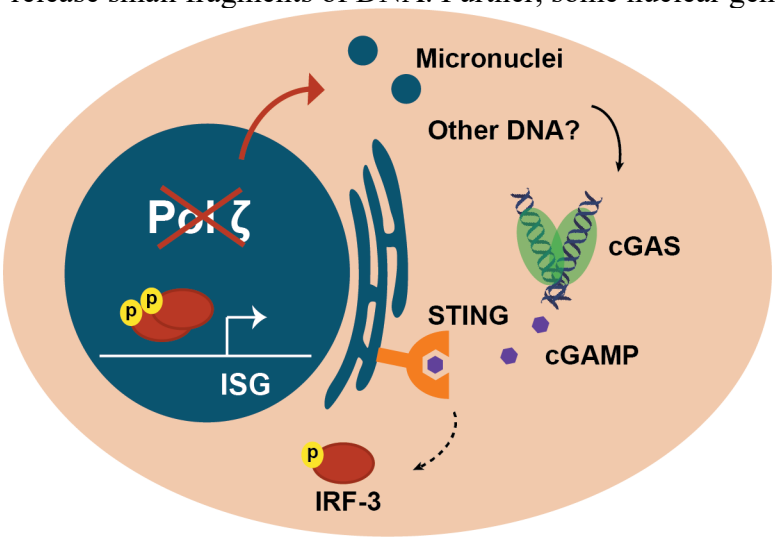


Figure 5: Model for disruption of pol ζ triggering an innate immune response. Loss of pol ζ induces genomic damage that results in accumulation of forms of cytosolic DNA, including micronuclei at a minimum. This results in DNA binding of cGAS and activation of STING which indirectly promotes phosphorylation and activation of IRF-3. This results in expression of interferon stimulated genes (ISG).

control mitochondrial DNA integrity, and there is evidence that mitochondrial function is compromised without pol ζ [29]. It remains to be seen whether micronuclei are the primary source of interferon signalling in cells lacking pol ζ, or whether they are more of an indicator of DNA degradation.

An interferon response can result in shutting down cell growth. Specifically, cGAS has been tied to promoting senescence in primary cells [30,31]. Our experiments were performed in T-antigen immortalized MEFs. In addition to blunting p53 activity, large T-antigen has been implicated in impairing an interferon response to nucleic acids [32]. Loss of pol ζ would likely induce an innate immune response of even greater magnitude in primary MEFs. Primary MEFs lacking pol ζ only make it approximately two cell divisions before cell growth completely halts, which is accompanied by an increase in senescent cells [11]. An essential function of pol ζ is also evident from the failure of *Rev3l*-defective embryos to develop, and the from the inability of *Rev3l*-defective primary keratinocytes to proliferate in a mouse model [2,10,11]. It is possible that cGAS-STING drives this severe growth arrest in primary cells due to loss of pol ζ function.

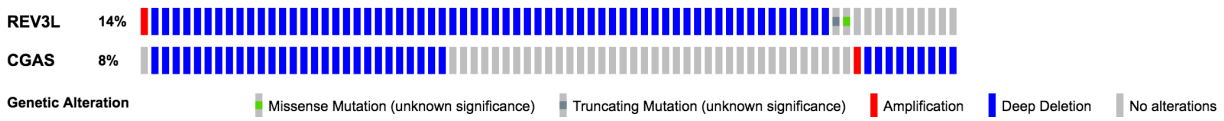
In addition to widening our understanding of the lengths cell go to protect themselves from the genomic damage induced by impairment of pol ζ, these results could have impact on translational approaches. A potential approach, suggested by experiments in laboratory settings, has been to disrupt pol ζ function to enhance chemotherapeutic effectiveness. For example, an inhibitor that impairs the interaction of pol ζ with the master regulator REV1 has been developed that sensitizes cancer cells and xenograft tumors to cisplatin treatment. Our work suggests that such inhibitors might also induce an interferon response. This approach would have multiple advantages for therapy by enhancing DNA damage sensitivity, limiting induced mutations, and potentially enhancing a cytotoxic immune response on targeted cells.

This is exciting and highly relevant to prostate cancer because our genetic analysis shows that *CGAS* is one of the co-deleted genes with *REV3L* in human primary prostate cancer, as it is located at human chromosome 6q13 in the deleted region (**Fig 7**). It is possible that *CGAS* deletion may allow *REV3L* defective cells to survive, and we are currently testing this candidate gene.

We are also testing *ATG5* as a candidate gene, deleted in a similar fraction of primary prostate adenocarcinoma. This is a collaboration with Dr. Bratton in our department.

If we obtain suitable human knockout cell lines for screening, we will use Crispr-Cas9 libraries as proposed.

Figure 7. Oncoprint generated from cBioportal for TCGA data of prostate adenocarcinoma (492 patients). About 13% of all patients have a deep deletion of *REV3L*, and about half of these also have a deep deletion of *CGAS*.



cBioPortal uses the GISTIC algorithm on copy number segments to determine if a gene has a deep deletion. We are finding that this algorithm overcalls deletions by looking at the underlying copy number segment data.

Another approach is to use copy number segment data and RNA-seq (**Figure 8**). This figure shows that copy number calls as deep (homozygous deletion) in TCGA are questionable for these data. Even the lowest copy number calls are not correlated with exceptionally low expression. Another view is the color-coded plot in **Figure 9**, covering the area of chromosome 6 harboring REV3L. The gradated spectrum of blue color shows that making a distinct call for homozygous deletion is not possible from these data. If one chooses only those patients (120) where coverage of the REV3L gene is < 0.2 (log2), it is interesting that those with some REV3L deletion show lower survival (**Figure 10**). We are investing this data by further bioinformatic analysis.

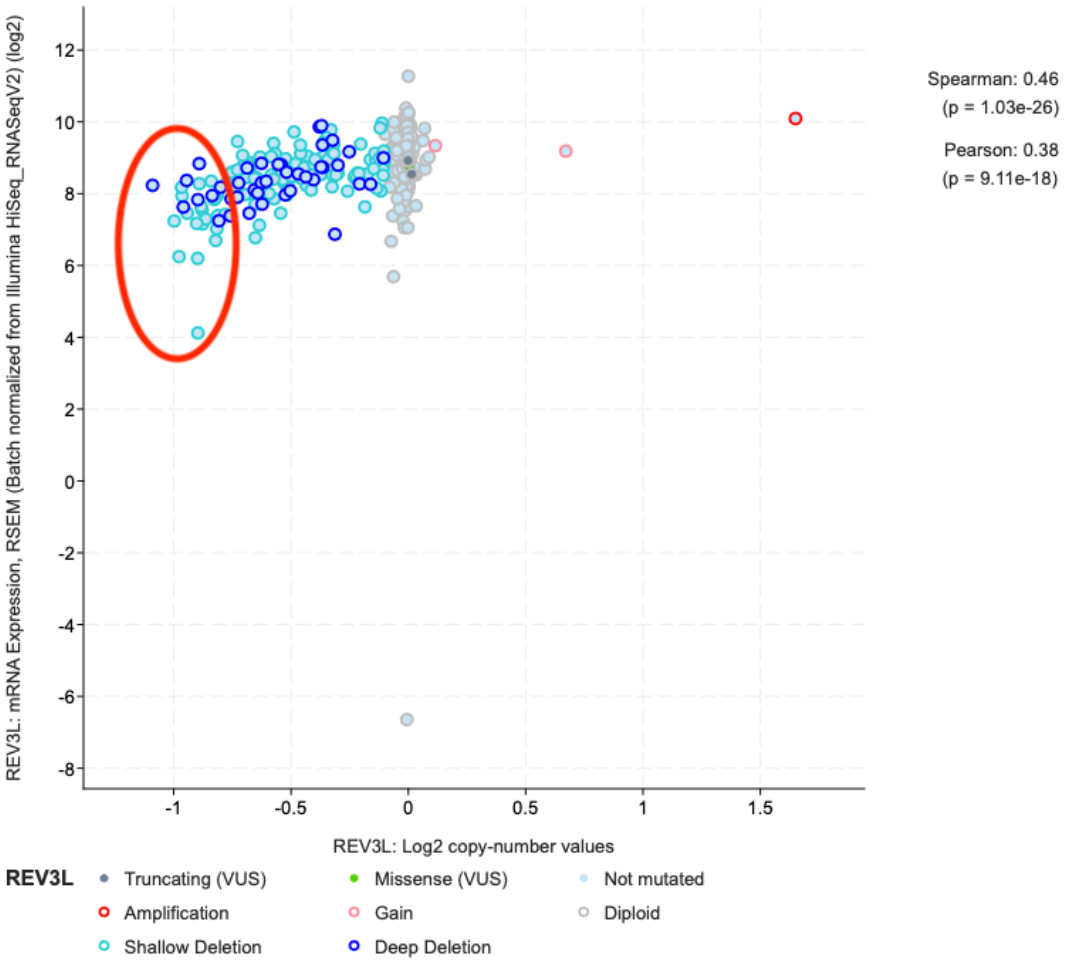


Figure 8. TCGA data for 499 Prostate Adenocarcinoma samples, plotting TCGA copy number values vs mRNA expression for *REV3L*. The points/samples within and around the red-circled area are those with multiple lines of evidence that there is a potential deep deletion. Ideally, one would like to identify samples in the extreme bottom left corner.

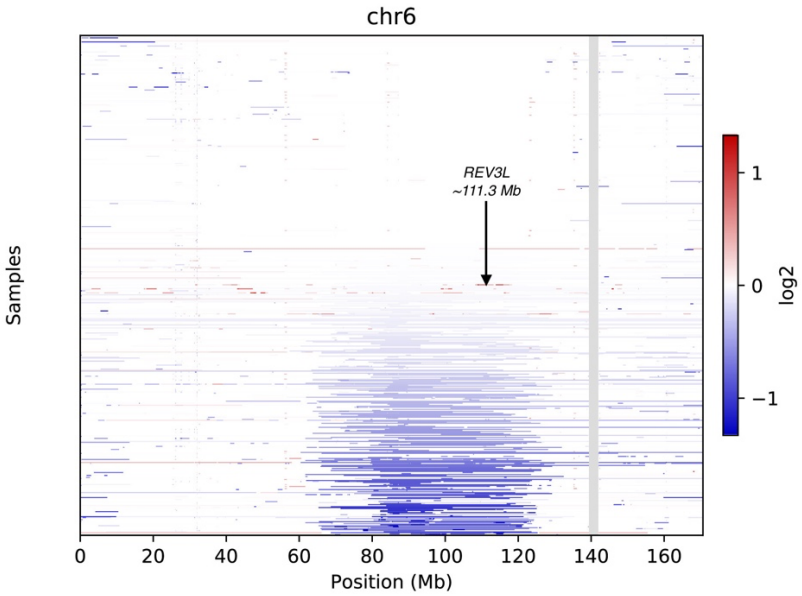


Figure 9. TCGA data for 499 Prostate Adenocarcinoma samples, plotting TCGA copy number values vs chromosome 6 location. *REV3L* is located at 111.3 to 111.4 Mb as indicated by the black arrow. The spectrum of blue color shows the difficulty in definitively assigning homozygous deletion calls from this type of data.

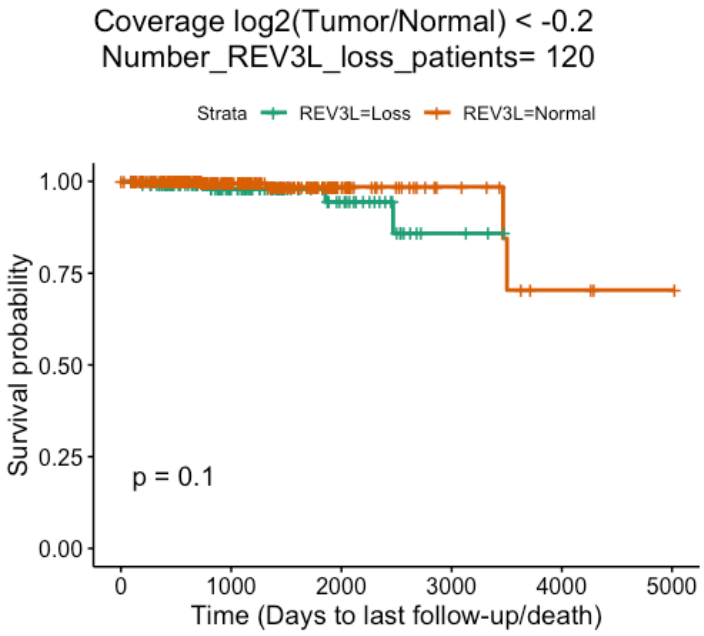
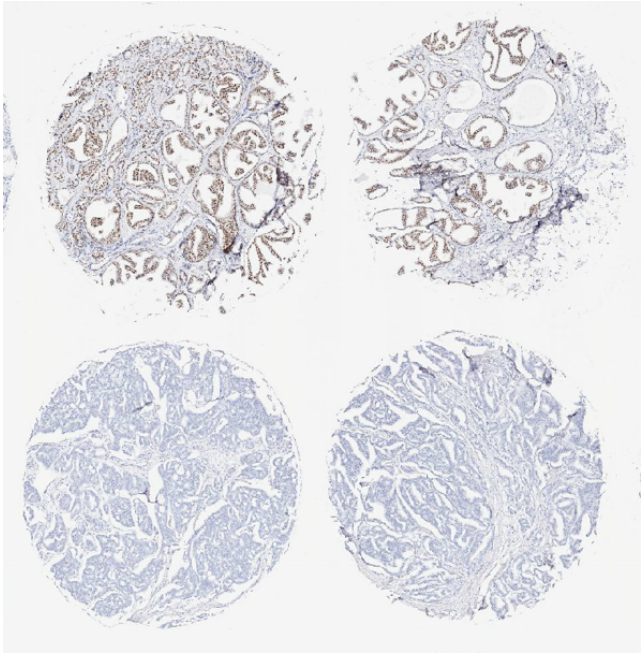


Figure 10. TCGA data for 120 Prostate Adenocarcinoma samples, where the coverage log 2 (Tumor/Normal) is less that -0.2. with p = 0.1, those samples with *REV3L* loss have a lower probability of survival than those with normal copy numbers of *REV3L*.

We wish to be able to detect cell containing the chromosome 6 deletion (including *REV3L*, *ATG5*, *CGAS* and *FOXO3A*) in an efficient manner. Therefore, we carried out immunohistochemical analysis of a prostate adenocarcinoma tissue microarray. An anti-FOXO3A antibody was used, and we found conditions that give a clear distinction between positive and negative samples (**Fig 11**). This is satisfying because it opens up the possibility of identifying *REV3L*-deleted human prostate cancer samples, which are expected to be vulnerable to chemotherapy or radiation.

Figure 11. Detection of FOXO3A-negative samples by staining a prostate adenocarcinoma tissue microarray (PR1921a) with anti-FOXO3A antibody (CST #12829). This antibody is diagnostic for the chromosome 6 deletion region and thus for REV3L deletion. The samples shown (duplicate tissue slices) are from malignant tumors of 73 year old individuals with prostate adenocarcinoma, Gleason grade 4. Formalin-fixed paraffin embedded tissue f was immunostained (brown) and counterstained with hematoxylin. The tumor from one patient (top two samples is FOXO3A positive), and from the other patient is FOXO3A negative. The FOXO3A status is predicted to reflect the REV3L status.



Significant Results or Key Outcomes: The research showed that identifying *REV3L* in tumors will require access to a large number of primary samples, and new methods for verifying deletions. The traditional method of using an antibody to REV3L for immunohistochemistry is unavailable. Further, only about 10% of prostate adenocarcinomas are predicted to harbor a deletion of the appropriate region of chromosome 6. We obtained 10 fresh-frozen biopsy slides from patients at MD Anderson, but none had relevant deletions as shown by staining with an antibody to the adjacent proxy marker FOXOA3. With the aid of our collaborator Dr. Dean Tang at the Roswell Park Cancer Institute, we have obtained additional sets of tissue microarrays (TMAs), which has provided a fresh approach. In the last year we stained the TMAs with FOXOA3 and have done preliminary scoring. The next step is to perform in situ hybridization of *REV3L* mRNA of the TMAs. Once the in situ hybridization for RNA is done, a serious study to identify suppressor genes could be done as proposed using mouse models. Until this background work is finished, a mouse model study as originally proposed In Specific Aim 3 is not practical or desirable.

During the course of the research, we made unexpected and important discoveries regarding the function of Rev3L in normal and cancer cells and reported them in four papers in top journals. All of these papers cite the support provided by the DOD grant. In the paper Martin SK, Tomida J,

Wood RD. *Cell Reports* (2021), we found that disruption of DNA polymerase ζ engages an innate immune response in mouse cells. The results are described further in the Appendix. In this paper we also set up the in situ hybridization technique for REV3L mRNA, and we are now poised to do similar experiments with the human transcript and TMAs. We reviewed the field of REV3L function in a survey and summary article for Nucleic Acids Research (Martin and Wood, 2019).

Other Achievements:

Adjacent to this exploration of REV3L function, we found that pol ζ operates in a role to help repair heterochromatin regions in normal cells. Ben Yamin et al, *EMBO Journal*, 2021. These regions are difficult to repair, and further help specify the essential role of REV3L.

As part of the synthetic lethality approach described in Aims 1 and Aim 2, we worked on a collaboration involving a related DNA polymerase. Relevant to this proposal, the results were able to answer whether there is a synthetic lethal relationship between subunits of pol ζ and pol θ. Synthetic lethality was detected with pol θ combined with suppression of the MAD2L2 (REV7) subunit, but not with the REV3L subunit (Feng et al,. *Nature Communications*, 2019). The synthetic lethality with MAD2L2 is related to the function of this protein as a component of the shieldin complex, where it helps control resection at double-strand breaks in DNA. The results show that pol θ inhibition is unlikely to be a first choice target for inhibition in REV3L-deficient prostate cancers.

What opportunities for training and professional development has the project provided?

During the overall project, two postdoctoral fellows and one graduate student have worked on the project, co-authored papers, and obtained positions afterwards, thus fostering their training and professional development. Further, two undergraduate students have been able to work with the key personnel on this project, and learn about mechanism of pol zeta action. They presented their work at the annual undergraduate symposium in the department.

How were the results disseminated to communities of interest?

The communities of interest are research scientists and physicians. Publications of our results has been the main mechanism of dissemination. We have also disseminated results by presenting them at department seminars, international meetings, and the to the prostate cancer SPRE group at MD Anderson.

What do you plan to do during the next reporting period to accomplish the goals?

Nothing to report (Final Report)

4. IMPACT:

Scientific feedback from our conference presentations on the work that is now published has been favorable. Peer review of all the papers was done before publishing the results in well regarded journals. The prospects for significant future impact are high, if we can identify the most relevant suppressor genes, and REV3L-deficient prostate cancers.

What was the impact on other disciplines?

This is hard to gauge at present, until we have more information on Rev3l deletion in prostate cancer.

What was the impact on technology transfer?

Technologies were transferred from scientist to scientist, including many trainees as described above. Thus, it seems the impact was high.

What was the impact on society beyond science and technology?

Education and training of undergraduates, and scientific education of medical professionals.

5. CHANGES/PROBLEMS:

Changes in approach and reason for change:

There were no major changes in approach and goals, but as described above we decided that the mouse tumor studies were premature and so we focused more on Aims and 2 of the project.

Actual or anticipated problems or delays and actions or plans to resolve them

Several factors combined to create delays during the course of this project. Following discussions with DOD we were able to obtain 2 no-cost extensions to the grant, allowing us to be significantly productive as shown by publications. Some delays were related to the ongoing COVID-19 pandemic (i) Because of the pandemic, our laboratories shut down on March 20 and reopened with shift work a few months later. All mouse work had to be stopped and could not be restarted in the interim (ii) We could not hire new positions, even grant funded, during this interim due to an institutional rule. US Government rules and uncertainties for obtaining a J1 visa for candidates are introducing further delays. We therefore requested a no-cost extension to conclude all studies, and this extension was approved. In addition, during 2021 we devoted months to moving of the laboratory (and our entire department) to a building in Houston.

Changes that had a significant impact on expenditures

These delays described above also spread out the rate of expenditure, but all funds from the grant have now been spent.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents:

No significant changes.

Significant changes in use or care of human subjects

No significant changes.

Significant changes in use or care of vertebrate animals

No significant changes.

Significant changes in use of biohazards and/or select agents

No significant changes.

6. PRODUCTS:

- **Publications, conference papers, and presentations**

Journal publications.

1. Martin SK, Tomida J, Wood RD. Disruption of DNA polymerase ζ engages an innate immune response. *Cell Reports*. 2021;34(8):108775. Epub 2021/02/25. doi: 10.1016/j.celrep.2021.108775. PubMed PMID: 33626348; PMCID: PMC7990024.

2. Ben Yamin B, Ahmed-Seghir S, Tomida J, Despras E, Pouvelle C, Yurchenko A, Goulas J, Corre R, Delacour Q, Droin N, Dessen P, Goidin D, Lange SS, Bhetawal S, Mitjavila-Garcia MT, Baldacci G, Nikolaev S, Cadoret JC, Wood RD, Kannouche PL. DNA polymerase zeta contributes to heterochromatin replication to prevent genome instability. *EMBO Journal*. 2021;n/a(n/a):e104543. doi: https://doi.org/10.15252/embj.2020104543.

3. Martin SK, Wood RD. DNA polymerase ζ in DNA replication and repair. *Nucleic Acids Res*. 2019;47(16):8348-61. doi: 10.1093/nar/gkz705. PubMed PMID: 31410467.

4. Feng W, Simpson D, Carvajal-Garcia J, Price BA, Kumar RJ, Mose LE, Wood RD, Rashid N, Purvis JE, Parker JS, Ramsden DA, Gupta GP. Genetic Determinants of Cellular Addiction to DNA Polymerase Theta. *Nature Commun*. 2019;10:4286. PubMed PMID: 31537809; PMCID: PMC6753077

Books or other non-periodical, one-time publications.

Nothing to Report.

Other publications, conference papers and presentations.

Leiden, The Netherlands, 5th DNA Polymerases Meeting, September 26, 2018, "DNA Polymerase zeta and cancer".

University of California, Berkeley, Department of Molecular and Cellular Biology, seminar presentation October 10, 2018, "DNA Polymerase, breaks and cancer".

New York University School of Medicine seminar presentation April 18, 2019, "A DNA polymerase for stress relief and cancer suppression".

Invited Speaker, University of Pittsburgh, Biological Sciences, Pittsburgh, PA, November 18, 2019, "DNA polymerases, breaks and cancer".

Invited Speaker, DNA Polymerases and Cancer, 4th DNA Repair/Replication Structures and Cancer Conference, Nassau, Bahamas, February 16, 2020, "DNA Polymerases and Cancer".

Invited Speaker, A DNA polymerase for stress relief and cancer suppression, Gordon Research Conference on DNA Damage, Mutation and Cancer, Ventura, CA, March 5, 2020, "A DNA Polymerase for Stress Relief and Cancer Suppression".

Speaker, Genome Instability Group (Houston), October 30, 2020 "DNA polymerases, breaks and cancer".

- **Website(s) or other Internet site(s)**

Nothing to report

- **Technologies or techniques**

As reported in the publications above

- **Inventions, patent applications, and/or licenses**

Nothing to report

- **Other Products**

Nothing to report

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Name:	Richard D. Wood (No Change)
Name:	Sarita Bhetawal
Project Role:	Senior Research Assistant
Nearest person month worked:	5
Contribution to Project:	Technical assistance with all experiments
Name:	Yuzhen Li
Project Role:	Postdoctoral Fellow
Nearest person month worked:	12
Contribution to Project:	Technical assistance with all experiments

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

NEW:
Title: *"Pol θ synthetic lethality in non-small cell lung cancer"*
Major Goals: Aim 1. In NSCLC cells with candidate DNA repair gene alterations, determine the synthetic lethality potential of POLQ disruption; Aim 2. Determine whether small molecule inhibitors inhibit proliferation of NSCLC cells with candidate DNA repair gene alterations; and Aim 3. Measure the level of POLQ protein expression in NSCLC cells
Status of Support: Active
Project Number: 2021-00059785-Y1
Name of PD/PI: Wood, Richard
Source of Support: Institutional Research Grant/Tobacco Pilot
Primary Place of Performance: The University of Texas MD Anderson Cancer Center, Houston, Texas
Project/Proposal Start and End Date: (MM/YYYY) (if available): 05/01/2021 – 8/31/2023
Total Award Amount (including Indirect Costs):
Person Months (Calendar/Academic/Summer) per budget period (.6 Calendar/Unpaid/year)

NEW:
Title: *"J. Ralph Meadows Chair in Carcinogenesis Research at Science Park – Research Division"*
Status of Support: Active
Project Number: 185800-80-102850-50
Name of PD/PI: Wood, Richard
Source of Support: Institutional Support
Primary Place of Performance: The University of Texas MD Anderson Cancer Center, Houston, Texas
Project/Proposal Start and End Date: (MM/YYYY) (if available): 08/01/2020 – Open
Person Months (Calendar/Academic/Summer) per budget period. (.42 Calendar/Year)
Total Award Amount (including Indirect Costs): Estimated annual income

What other organizations were involved as partners?

Roswell Park Comprehensive Cancer Center Elm and Carlton Streets, Buffalo, NY 14263 Collaboration with Dean Tang, Ph.D., Professor and Chair

8. SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS: ***N/A***
QUAD CHARTS: N/A

9. BIBLIOGRAPHY

Publications:

1. Martin SK, Tomida J, Wood RD. Disruption of DNA polymerase ζ engages an innate immune response. *Cell Reports*. 2021;34(8):108775. Epub 2021/02/25. doi: 10.1016/j.celrep. 2021.108775. PubMed PMID: 33626348; PMCID: PMC7990024.
2. Ben Yamin B, Ahmed-Seghir S, Tomida J, Despras E, Pouvelle C, Yurchenko A, Goulas J, Corre R, Delacour Q, Droin N, Dessen P, Goidin D, Lange SS, Bhetawal S, Mitjavila-Garcia MT, Baldacci G, Nikolaev S, Cadoret JC, Wood RD, Kannouche PL. DNA polymerase zeta contributes to heterochromatin replication to prevent genome instability. *EMBO Journal*. 2021;n/a(n/a):e104543. doi: https://doi.org/10.15252/embj.2020104543.
3. Martin SK, Wood RD. DNA polymerase ζ in DNA replication and repair. *Nucleic Acids Res*. 2019;47(16):8348-61. doi: 10.1093/nar/gkz705. PubMed PMID: 31410467.
4. Feng W, Simpson D, Carvajal-Garcia J, Price BA, Kumar RJ, Mose LE, Wood RD, Rashid N, Purvis JE, Parker JS, Ramsden DA, Gupta GP. Genetic Determinants of Cellular Addiction to DNA Polymerase Theta. *Nature Commun*. 2019;10:4286. PubMed PMID: 31537809; PMCID: PMC6753077.

Meetings/Abstracts:

Leiden, The Netherlands, 5th DNA Polymerases Meeting, September 26, 2018, "DNA Polymerase zeta and cancer".

University of California, Berkeley, Department of Molecular and Cellular Biology, seminar presentation October 10, 2018, "DNA Polymerase, breaks and cancer".

New York University School of Medicine seminar presentation April 18, 2019, "A DNA polymerase for stress relief and cancer suppression".

Invited Speaker, University of Pittsburgh, Biological Sciences, Pittsburgh, PA, November 18, 2019, "DNA polymerases, breaks and cancer".

Invited Speaker, DNA Polymerases and Cancer, 4th DNA Repair/Replication Structures and Cancer Conference, Nassau, Bahamas, February 16, 2020, "DNA Polymerases and Cancer".

Invited Speaker, A DNA polymerase for stress relief and cancer suppression, Gordon Research Conference on DNA Damage, Mutation and Cancer, Ventura, CA, March 5, 2020, "A DNA Polymerase for Stress Relief and Cancer Suppression".

Speaker, Genome Instability Group (Houston), October 30, 2020 "DNA polymerases, breaks and cancer".

List of Personnel receiving pay from the research effort:

Richard D. Wood, Professor and PI
Junya Tomida, Instructor
Yi Zhong, Senior Statistical Analyst
Sara Martin, Graduate Student
Sarita Bhetawal, Senior Research Assistant
Yuzhen Li, Postdoctoral Fellow

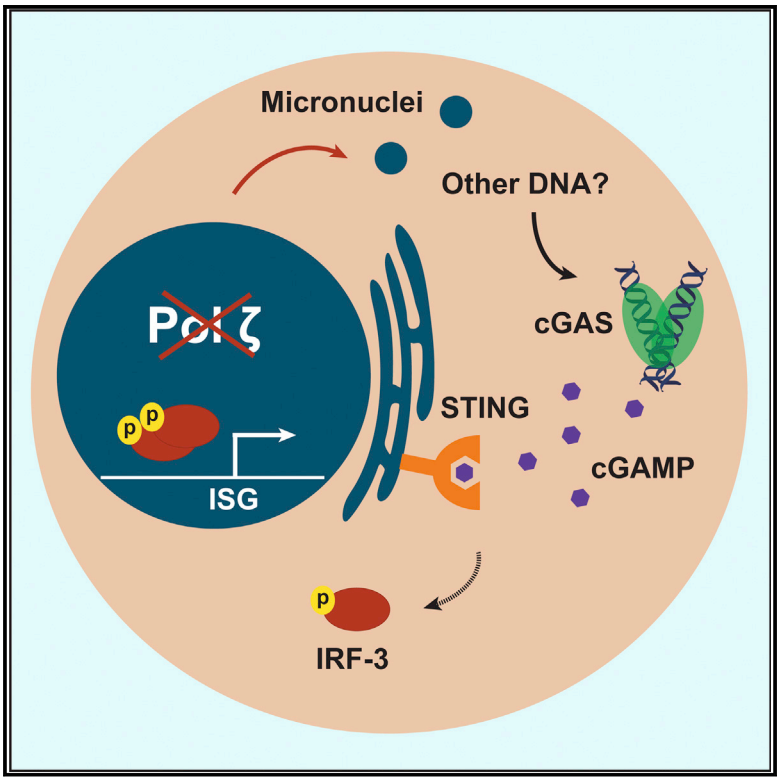
10. APPENDICES: (See publications attached)

Cell Reports

Report

Disruption of DNA polymerase ζ engages an innate immune response

Graphical Abstract



Authors

Sara K. Martin, Junya Tomida, Richard D. Wood

Correspondence

rwood@mdanderson.org

In brief

Martin et al. show that disruption of the catalytic subunit of DNA polymerase ζ, *Rev3l,* results in the accumulation of DNA damage and increased expression of interferon-stimulated genes. The cGAS-STING pathway, part of the innate immune system that responds to genomic stress, drives this expression signature in pol ζ-deficient cells.

Highlights

- Loss of pol ζ (*Rev3l*) leads to chromosome damage marked by increased micronuclei
- *Rev3l* loss increases expression of interferon-stimulated genes (ISGs) and proteins
- Expression of ISG chemokines is elevated in *Rev3l*-disrupted primary epithelial cells
- The cGAS-STING pathway drives this ISG expression signature in *Rev3l*-disrupted cells

Martin et al., 2021, Cell Reports *34*, 108775
February 23, 2021
https://doi.org/10.1016/j.celrep.2021.108775

CellPress

Cell Reports

CellPress
OPEN ACCESS

Report

Disruption of DNA polymerase ζ engages an innate immune response

Sara K. Martin,[1,2,3] Junya Tomida,[1,4] and Richard D. Wood[1,2,5,*]
[1]Department of Epigenetics & Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Smithville, TX 78507, USA
[2]The University of Texas MD Anderson Cancer Center, UT Health Graduate School of Biomedical Sciences, Houston, TX, USA
[3]Present address: Department of Biology, Tufts University, Medford, MA 02155, USA
[4]Present address: Department of Biological Sciences, University of North Carolina at Charlotte, Charlotte, NC 28223, USA
[5]Lead contact
*Correspondence: rwood@mdanderson.org
https://doi.org/10.1016/j.celrep.2021.108775

SUMMARY

In mammalian cells, specialized DNA polymerase ζ (pol ζ) contributes to genomic stability during normal DNA replication. Disruption of the catalytic subunit *Rev3l* is toxic and results in constitutive chromosome damage, including micronuclei. As manifestations of this genomic stress are unknown, we examined the transcriptome of pol ζ-defective cells by RNA sequencing (RNA-seq). Expression of 1,117 transcripts is altered by ≥4-fold in *Rev3l*-disrupted cells, with a pattern consistent with an induction of an innate immune response. Increased expression of interferon-stimulated genes at the mRNA and protein levels in pol ζ-defective cells is driven by the cyclic guanosine monophosphate-adenosine monophosphate synthase (cGAS)-signaling partner stimulator of interferon genes (STING) pathway. Expression of key interferon-stimulated chemokines is elevated in basal epithelial mouse skin cells with a disruption of *Rev3l*. These results indicate that the disruption of pol ζ may simultaneously increase sensitivity to genotoxins and potentially engage parts of the innate immune response, which could add an additional benefit to targeting pol ζ in cancer therapies.

INTRODUCTION

Mammalian genomes encode an array of translesion DNA polymerases, which provide a diverse tool kit to tolerate assorted genomic lesions (Lange et al., 2011). While most translesion DNA polymerases are required for cells to survive various exogenous genotoxic assaults, they are not essential for mammalian development or unchallenged cellular survival (Lange et al., 2011). An exception is DNA polymerase ζ (pol ζ). The pol ζ catalytic subunit is encoded by the *Rev3l* gene (Martin and Wood, 2019). Germline disruption of *Rev3l* results in embryonic lethality in mice (Bemark et al., 2000; Esposito et al., 2000; Wittschieben et al., 2000). The indispensable nature of pol ζ reflects genome protective functions that are inadequately understood. Disruption of *Rev3l* in B cells or keratinocytes in mice leads to acute genomic stress in the target tissues (Daly et al., 2012; Lange et al., 2013, 2018; Schenten et al., 2009; Wittschieben et al., 2010). Primary mouse embryonic fibroblasts (MEFs) rapidly accumulate chromosome breaks at the first metaphase following *Rev3l* disruption (Lange et al., 2012). In addition, chromosome rearrangements are observed in *Rev3l*-deficient MEFs in a p53 null background (Wittschieben et al., 2006). Loss of pol ζ in primary MEFs cripples cell proliferation, with the cells failing to replicate past roughly two rounds of cell division following *Rev3l* disruption (Lange et al., 2012). If cells can survive *Rev3l* disruption, this promotes spontaneous tumorigenesis in mouse models (Lange et al., 2013; Wittschieben et al., 2010). Given that the loss of pol ζ function results in chromosome breaks and structural abnormalities (Lange et al., 2012; Wittschieben et al., 2006), this could generate genetic variety to promote oncogenic growth, but only if the cells can overcome the disruption of pol ζ. This makes it of great interest to understand the cellular consequences of the loss of *Rev3l*.

The dramatic growth suppression in *Rev3l* knockout MEFs appears to be ameliorated by blunting DNA damage quality control responses, for example, by p53 deletion (Wittschieben et al., 2006) or large T-antigen expression (Lange et al., 2012) (which inhibits p53 and other targets). However, p53 deletion utterly fails to rescue the embryonic lethality of *Rev3l* disruption in mice (O-Wang et al., 2002; Van Sloun et al., 2002; Wittschieben et al., 2006). This implies that there are additional, p53-independent, growth-suppressive responses to the genomic strain caused by *Rev3l* disruption. We set out to clarify the consequences of the sustained genomic stress induced by the loss of pol ζ in mammalian cells. Starting with an unbiased transcriptome-wide approach, we discovered that the loss of pol ζ induces the constitutive expression of immune system-related genes, in particular, interferon-stimulated genes (ISGs). Furthermore, we found that the cytosolic nucleic acid sensor cyclic guanosine monophosphate–AMP synthase (cGAS) and its downstream signaling partner stimulator of interferon genes (STING) drive this response in absence of functional pol ζ. Given that the cGAS-STING axis can inhibit cell growth, this provides a further

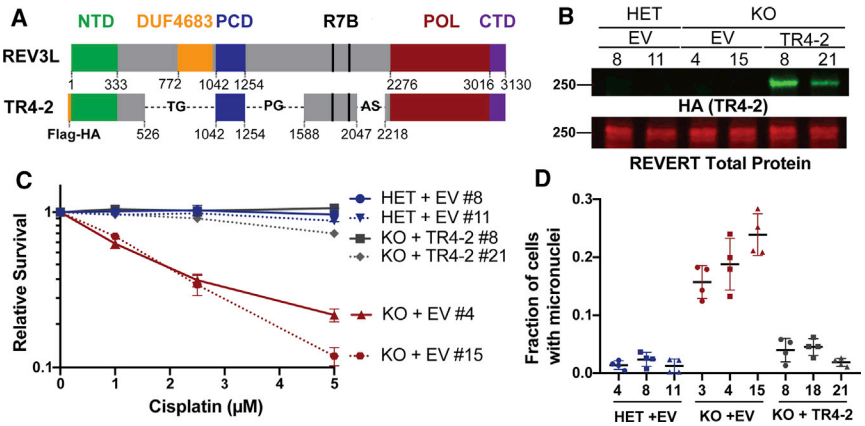


Figure 1. Shortened REV3L construct rescues phenotypes of pol ζ disruption

(A) Schematic of full-length human REV3L and human REV3L construct TR4-2. TR4-2 retains most conserved domains and binding sites of REV3L, including the regions that coordinate interactions with the accessory subunits of pol ζ, the C-terminal domain (CTD, purple), the 2 REV7 binding sites (R7B, black), and a positively charged domain (PCD) of uncertain function. The B-family catalytic core is formed by folding of the N-terminal domain (NTD, green) and the polymerase domain (POL, red) (Martin and Wood, 2019).

(B) Immunoblot with HA antibody showing stable expression of TR4-2 with an N-terminal FLAG-HA tag in *Rev3l* KO MEF clones.

(C) Stable expression of TR4-2 in *Rev3l* KO MEF clones reverses hypersensitivity to cisplatin. MEFs were exposed to the indicated cisplatin concentrations for 48 h, and relative survival was quantified with the ATPlite assay. The error bars represent the standard deviation of 3 replicates.

(D) Stable expression of TR4-2 decreases micronuclei formation in unchallenged *Rev3l* KO MEF clones. The error bars represent the standard deviation of four replicates.

See also Figure S1.

potential explanation for why the loss of pol ζ function is remarkably toxic to proliferating cells.

RESULTS AND DISCUSSION

A shortened REV3L construct rescues known phenotypes of pol ζ disruption

To dissect the consequences of the poorly resolved genome protection function of pol ζ, we set up a complementation system using large T-antigen immortalized MEFs (Lange et al., 2012), either with a pol ζ-proficient background, *Rev3l* heterozygous (HET), or a pol ζ deficient background, *Rev3l* knockout (KO). The KO cell lines (*Rev3l*$^{-/\Delta}$) were generated by adenovirus-Cre disruption of a single allele in *Rev3l*$^{-/lox}$ MEFs (Lange et al., 2012). HETs (*Rev3l*$^{+/\Delta}$), which retain pol ζ activity, provides a control for this process since they also had deletion of a single *Rev3l* allele using adenovirus-Cre with *Rev3l*$^{+/lox}$ MEFs as the parental cell line. We have not observed phenotypic differences between *Rev3l* HETs and WT cells or mice (Lange et al., 2012).

A biochemically active shortened human REV3L construct, TR4-2 (Lee et al., 2014) (Figure 1A), with an N-terminal FLAG-hemagglutinin (HA) tag, was stably expressed in *Rev3l* KO MEFs (Figures 1B and S1A). To test that our complementation system was in fact addressing both external and endogenous genome-protective functions of pol ζ, we tested both cellular sensitivity to cisplatin and the formation of micronuclei in unchallenged cells. Micronuclei are a marker of genomic instability and arise when a whole chromosome or fragment of a chromosome missegregates and forms its own discrete nuclear compartment. DNA breaks that result in chromosome fragmentation can lead to acentric chromosomes that form micronuclei (Fenech et al., 2011).

As expected, *Rev3l* KO + empty vector (EV) MEFs were hypersensitive to cisplatin relative to the *Rev3l* HET MEFs + EV (Figures 1C and S1B). Stable expression of TR4-2 in *Rev3l* KO MEFs reversed the hypersensitivity to cisplatin (Figures 1C and S1B). This indicates that in this context, TR4-2 can restore the function of REV3L in DNA damage tolerance. Consistent with large-scale genomic stress, ∼15%–23% of unchallenged *Rev3l* KO MEFs had at least 1 micronucleus, relative to *Rev3l* HET MEFs, which had a micronucleus frequency of ∼1%–3% (Figure 1D). Stable TR4-2 expression in *Rev3l* KO MEFs restored micronucleus frequency to near normal, showing that TR4-2 can restore some *Rev3l* genome-protective functions. We used this isogenic system as a tool to probe the unknown consequence of the loss of pol ζ.

Loss of polymerase ζ alters the transcriptome

To uncover the type of stress occurring in cells lacking pol ζ, we performed genome-wide mRNA sequencing on a controlled set of immortalized clones: *Rev3l* HET + EV, *Rev3l* KO + EV, and *Rev3l* KO + TR4-2. To focus on major changes, we set a strict threshold (>|2| $\log_2$ fold change and false discovery rate [FDR] < 0.05). Expression analysis of 17,346 mapped transcripts revealed that 1,117 transcripts were either upregulated or downregulated in the *Rev3l* KO + EV relative to the *Rev3l* HET + EV MEFs (Figure 2A). The majority (∼68%) were upregulated (Figure 2A). These upregulated or downregulated genes displayed no statistically significant enrichment or depletion for DNA replication or canonical DNA damage-sensing pathways. This is not completely unexpected, given that p53 promotes much of the transcriptional response to DNA damage, and our MEFs have inactivated p53 due to large T-antigen immortalization. In these immortalized high-passage cells, we are likely to observe stable transcriptional alterations, rather than an acute response.

Instead, the upregulated genes displayed an enrichment in immune system-related pathways, as revealed by Gene Ontology (GO) analysis (Figure 2B). Upstream regulator analysis was used to analyze all differentially expressed genes. This revealed that the alterations in the transcriptome are consistent with the activation of positive regulators of the interferon response, including key transcription factors in this pathway, IRF3 and

CellPress
OPEN ACCESS

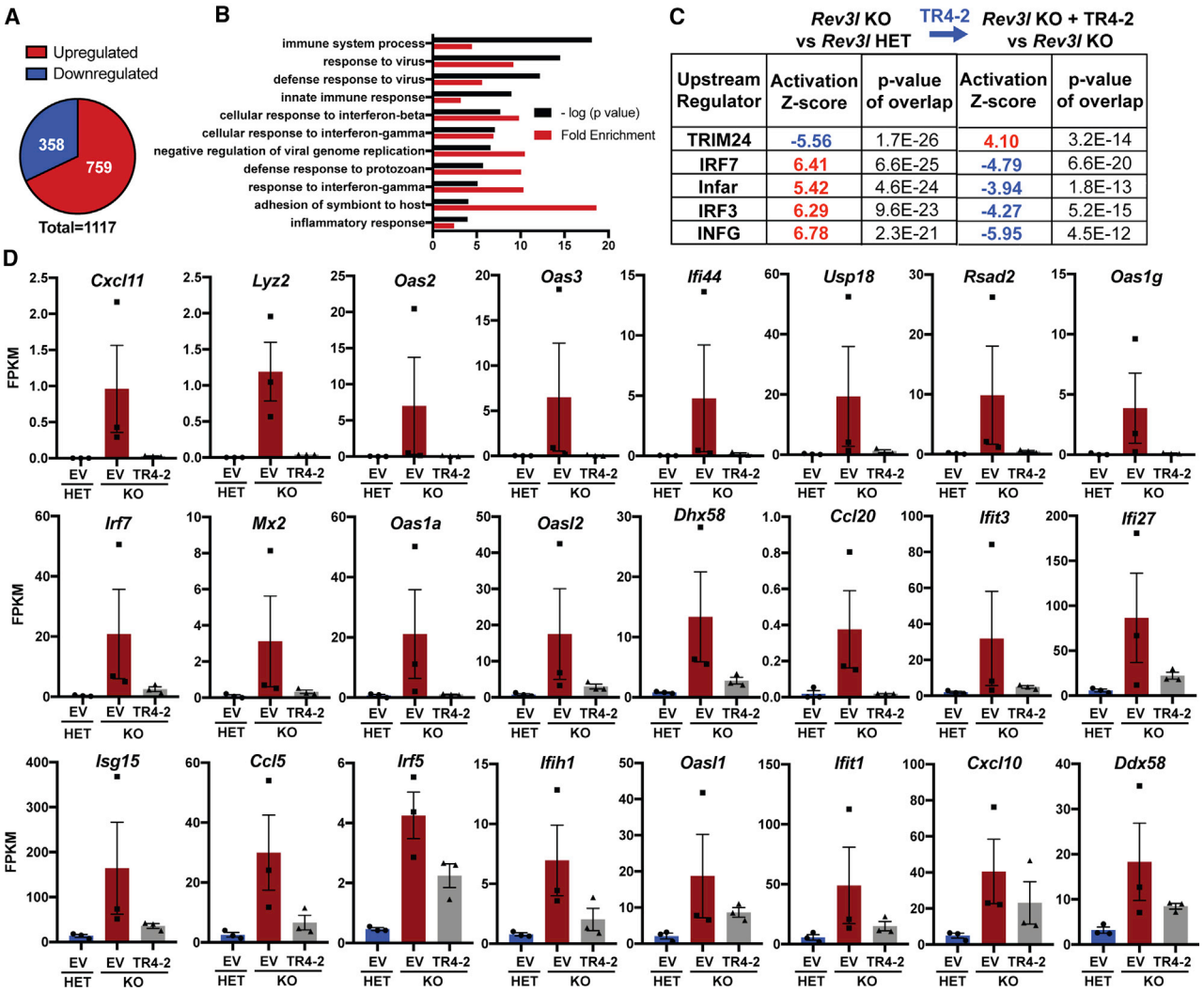


Figure 2. Cells deficient in DNA polymerase ζ have an altered transcriptome

(A) Differentially expressed genes in *Rev3l* KO + empty vector (EV) relative to *Rev3l* HET + EV using the threshold of a fold change >4 and a false discovery rate of <0.05.

(B) Top 10 GO (Gene Ontology) terms reveal enrichment of immune system-related genes in upregulated genes in *Rev3l* KO MEFs.

(C) Upstream regulator analysis reveals the dataset is consistent with the predicted activation of positive regulators of an interferon response in *Rev3l* KO MEFs.

(D) Expression of individual interferon-stimulated genes (ISGs) is elevated in *Rev3l* KO MEFs and is suppressed by TR4-2 expression. The fragments per kilobase of transcript per million mapped reads (FPKM) of 3 clones for each condition, HET + EV (ID nos. 4, 8, 11), KO + EV (ID nos. 3, 4, 15), and KO +TR4-2 (ID nos. 8, 18, 21) is graphed. The error bars represent SEMs.

See also Figure S2 and Table S1.

IRF7 (Figure 2C). Importantly, the predicted activation of all five of the identified upstream regulators was reversed by expression of the *Rev3l* TR4-2 cDNA (Figure 2C). This shows that the changes in differential gene expression stem from a function of *Rev3l*. The same trends were also observed by applying a substantially lower threshold ($\log_2$ FC > |0.5|) for differentially expressed genes. This increased the dataset to 2,071 differentially expressed transcripts. Pathway analysis showed a negative correlation with predicted TRIM24 activation. TRIM24 suppresses ISG expression (Tisserand et al., 2011), confirming that our data are consistent with the expression of ISGs.

To explore whether our dataset is in fact consistent with an interferon-like response, we analyzed a curated set of 24 known ISGs and observed increased expression in the *Rev3l* KO relative to the *Rev3l* HET MEFs, which was largely reduced by the stable expression of the *Rev3l* cDNA TR4-2 (Figure 2D). It is notable that while the expression of TR4-2 conferred nearly complete rescue of cisplatin sensitivity (Figure 1C) and micronuclei formation (Figure 1D), the expression of some individual ISGs was only partially reversed (Figures 2D and S2; Table S1). It is possible that the expression of some genes in the MEF clones may have been stabilized by epigenetic changes, not reversible by the reintroduction

of *Rev3l*. It is also possible that segments of REV3L not present in TR4-2 (e.g., the DUF4683 domain) (Figure 1A) may contribute to a specific gene expression function of REV3L but not to the DNA damage bypass function. These results reveal that the disruption of pol ζ promotes the induction of ISGs.

The cGAS-STING pathway drives the expression of ISG expression in *Rev3l*-disrupted cells

Given that our complemented cell lines were generated using lentivirus constructs and grown continually under selection, we moved to validate the results in the parental *Rev3l* KO and *Rev3l* HET MEF cell lines and 1 additional set of cell lines to limit extraneous variables. We confirmed an increase in the mRNA expression of specific ISGs in the *Rev3l* KO MEFs relative to the control cell lines, including key chemokines (Figure 3A). To extend these findings to the protein level, we examined ISG products by immunoblotting. Corresponding to an increase in mRNA levels, we also observed an increase in the protein levels of known ISGs MDA5 (encoded by the *Ifih1* gene), ISG15, and viperin (encoded by the *Rsad2* gene) (Figure 3B).

These results were obtained by analysis of isogenically matched pairs of *Rev3l*-proficient and *Rev3l*-deficient immortalized cell lines. An important question is whether the disruption of REV3L function drives the induction of ISGs in primary cells. This is experimentally challenging, because complete disruption in primary cells rapidly promotes senescence in primary MEFs (Lange et al., 2012). We instead took advantage of a mouse model system in which *Rev3l* is specifically disrupted only in keratin-5-expressing cells, principally the epithelial cells of the skin and hair follicles. These mice survive, but they have a constitutively lower cell density in the basal epithelium (Lange et al., 2013). The skin is extremely sensitive to physical insults, including UV irradiation and wounding (Lange et al., 2013, 2018). *Rev3l*-defective epithelial cells struggle with re-proliferation because of ongoing DNA replication stress, which accounts for the fragility of the epithelia (Lange et al., 2013, 2018).

We hypothesized that the expression of ISGs in the immortalized MEFs may depend on DNA damage incurred during replication. Thus, we examined primary cells arising from proliferation in the skin. In normal mice, stem cells in the hair follicles undergo division and migrate to continuously repopulate the epithelium. Using skin sections from mice previously reported (Lange et al., 2018), we tested the expression of the mRNA of CXCL10 and CCL5 by *in situ* hybridization (Figures 3C–3F and S3). We found that with 3 independent mice in *Rev3l*-defective cells, CXCL10 and CCL5 levels were significantly upregulated by 2.6-fold and 2.9-fold, respectively (Figures 3E and 3F).

These data indicate that an interferon-like branch of the innate immune system may be activated due to the disruption of pol ζ function. Since it seems unlikely that pol ζ plays a direct role in transcriptional regulation, the next obvious question is how and why loss of pol ζ induces the expression of ISGs.

The major consequence of pol ζ disruption in unchallenged mammalian cells is increased genomic instability as evidenced by multiple markers, including chromosome fragmentation and aberrations and micronuclei (Lange et al., 2012, 2013, 2016; Wittschieben et al., 2006) (Figure 1D). Therefore, it seems likely that this transcriptional response ultimately stems from the vast genomic damage induced by the loss of pol ζ function. Consistent with this hypothesis, the innate immune system can respond to endogenous DNA that has escaped from the nucleus, in addition to its more canonical role in recognizing and mounting an interferon response to foreign DNA. In some instances, this response can halt cell growth to shut down propagations of virally infected cells and cells with dangerously fragmented genomes.

Mammalian cells have a host of cytosolic nucleic acid sensors that patrol the cytosol for DNA. One of these, cGAS, is increasingly recognized to be of paramount importance in the induction of an interferon response to both exogenous and endogenous cytosolic DNA (Ablasser and Chen, 2019). Binding of double-stranded DNA to cGAS activates the enzymatic activity of cGAS and leads to the production of 2′,3′ cyclic GMP-AMP (cGAMP). cGAMP binds to the STING receptor on the membrane of the endoplasmic reticulum, resulting in the activation of kinases, including TBK1, which can in turn phosphorylate and activate interferon regulatory factor 3 (IRF3), a central transcription factor in the interferon response.

Given that the cGAS-STING axis has been implicated specifically in responding to endogenous DNA damage and has been correlated with micronuclei formation (a potential source for cGAS activation), we asked whether cGAS and STING promote the induction of expression of ISGs due to the loss of function of pol ζ. Consistent with most MEFs having a functional innate immune system, both *Rev3l* KO and HET MEF cell lines expressed both cGAS and STING (Figure 3G). In the *Rev3l* KO cells, the steady-state levels of STING (but not cGAS) are somewhat lower (decreased by ~3- to 4-fold for KO 1 and decreased by ~20%–60% in KO 2). The decreased STING protein level is consistent with a constitutive activation of the cGAS-STING pathway, as cGAS activation leads to a negative feedback loop resulting in STING degradation (Gonugunta et al., 2017; Prabakaran et al., 2018). Importantly, we detected an increase in IRF3 phosphorylated at S888 (corresponding to S396 in humans), indicative of IRF3 activation in *Rev3l* KO MEFs (Panne et al., 2007; Yoneyama et al., 2002) (Figure 3F).

This led us to investigate whether cGAS-STING drives the expression of ISGs in cells with disrupted pol ζ. Knockdown of either cGAS or STING significantly reduces the mRNA expression of selected ISGs and dramatically decreases corresponding protein levels (Figures 4A–4D). In addition, the depletion of cGAS or STING in *Rev3l* KO MEFs markedly reduced S888 phosphorylation of IRF3 (Figure 4E). This indicates that the disruption of pol ζ function promotes the upregulation of ISGs driven by the cGAS-STING pathway.

Pol ζ stands apart from the other translesion polymerases in that it is required for mammalian development and proliferation of primary cells. Now we can say that in addition to activating p53-dependent responses, disruption of pol ζ function promotes the activation of the cGAS-STING pathway.

It is remarkable that the disruption of an enzyme commonly thought of as a specialized translesion synthesis polymerase can lead to the constitutive activation of the cGAS-STING pathway. Recently, cells with a loss of function of key DNA repair enzymes RNaseH2, BRCA2, and BLM have been shown to have an elevated cGAS-STING response that correlates with

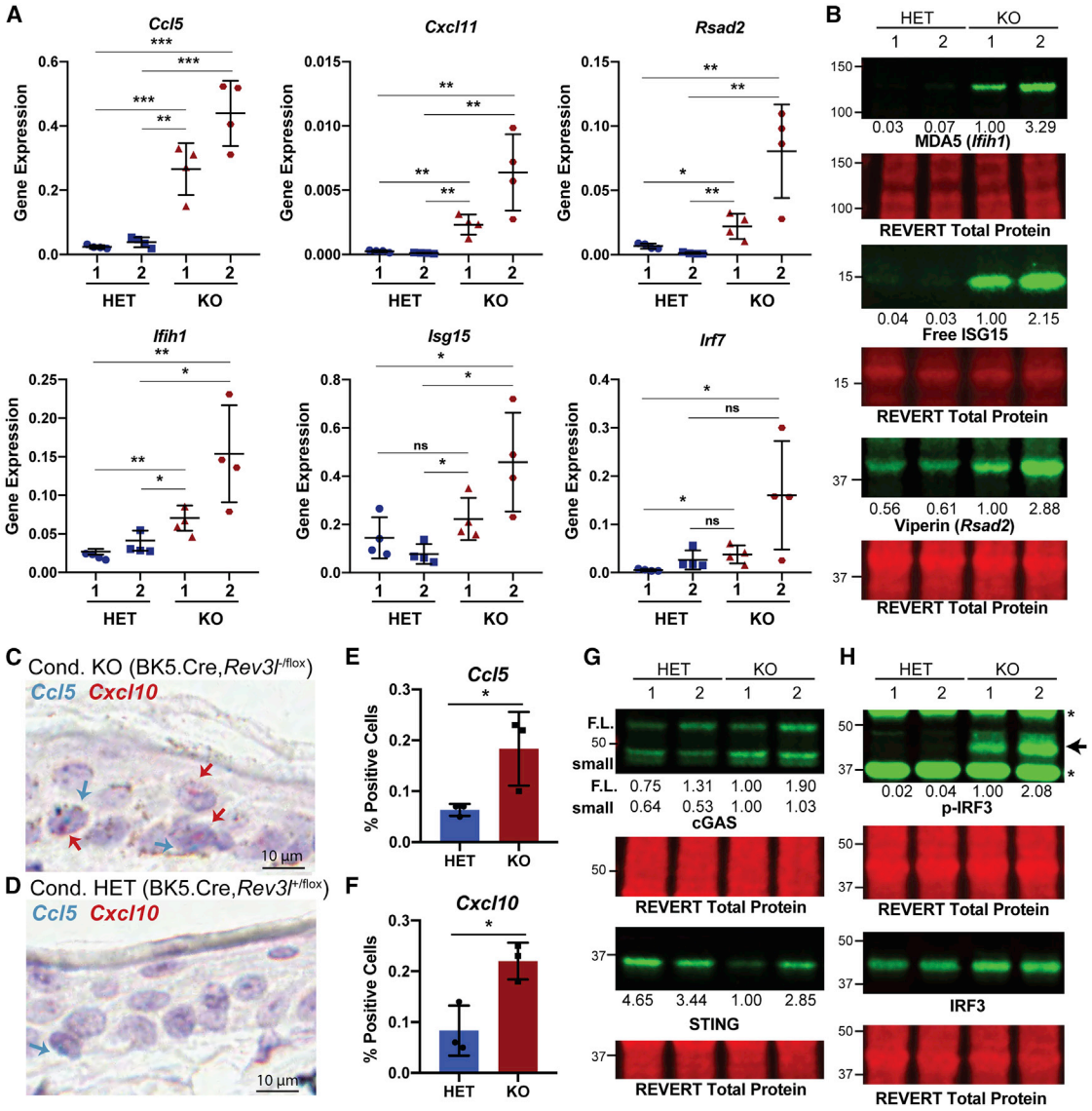


Figure 3. Disruption of *Rev3l* results in increased expression of ISGs and proteins

(A) Gene expression ($2^{-\Delta Ct}$) of selected ISGs normalized to hypoxanthine phosphoribosyltransferase (HPRT) detected by qRT-PCR. Error bars represent standard deviation. Unpaired 2-tailed Student's t test, *p < 0.05, **p < 0.01, and ***p < 0.001.

(B) Immunoblots showing increased protein levels of ISGs MDA5, ISG15, and viperin. The quantification of each protein relative to KO 1 is shown underneath the blot. The signals were first normalized to total protein.

(C) Representative image of CCL5 (blue) and CXCL10 (red) signals in the epithelia of conditional *Rev3l* KO mice (BK5.Cre,*Rev3l*$^{-/flox}$) positive cells are shown with blue and red arrows, respectively. A 10 μm scale bar is shown. The brown staining is epithelial melanin present in the KO model.

(D) Same as (C), except for conditional *Rev3l* HET mice (BK5.Cre,*Rev3l*$^{+/flox}$). A 10 μm scale bar is shown.

(E) Quantification for percentage of cells positive for CCL5 in the epithelia for 3 mice of each genotype.

(F) Same as (E), except for CXCL10.

(G) Immunoblot showing the presence of components of the innate immune system, cGAS and STING, in MEFs, with reduced STING in pol ζ KO cells. Quantification of the signal intensity of cGAS and STING relative to KO 1 is displayed underneath the blot. The signal intensity of each band was first normalized to total protein. For cGAS, we detect 2 bands: 1 at full-length molecular weight (F.L.) and 1 (small) that we confirm is another cGAS species in Figure 4C.

(H) Enhanced phosphorylation of S888 of IRF3 (corresponding to S396 in humans) in *Rev3l* KO MEFs. The relative quantification of p-IRF3 signal intensity normalized to IRF3 is displayed underneath the p-IRF3 blot. The signal intensities were first normalized to total protein signal.

See also Figure S3.

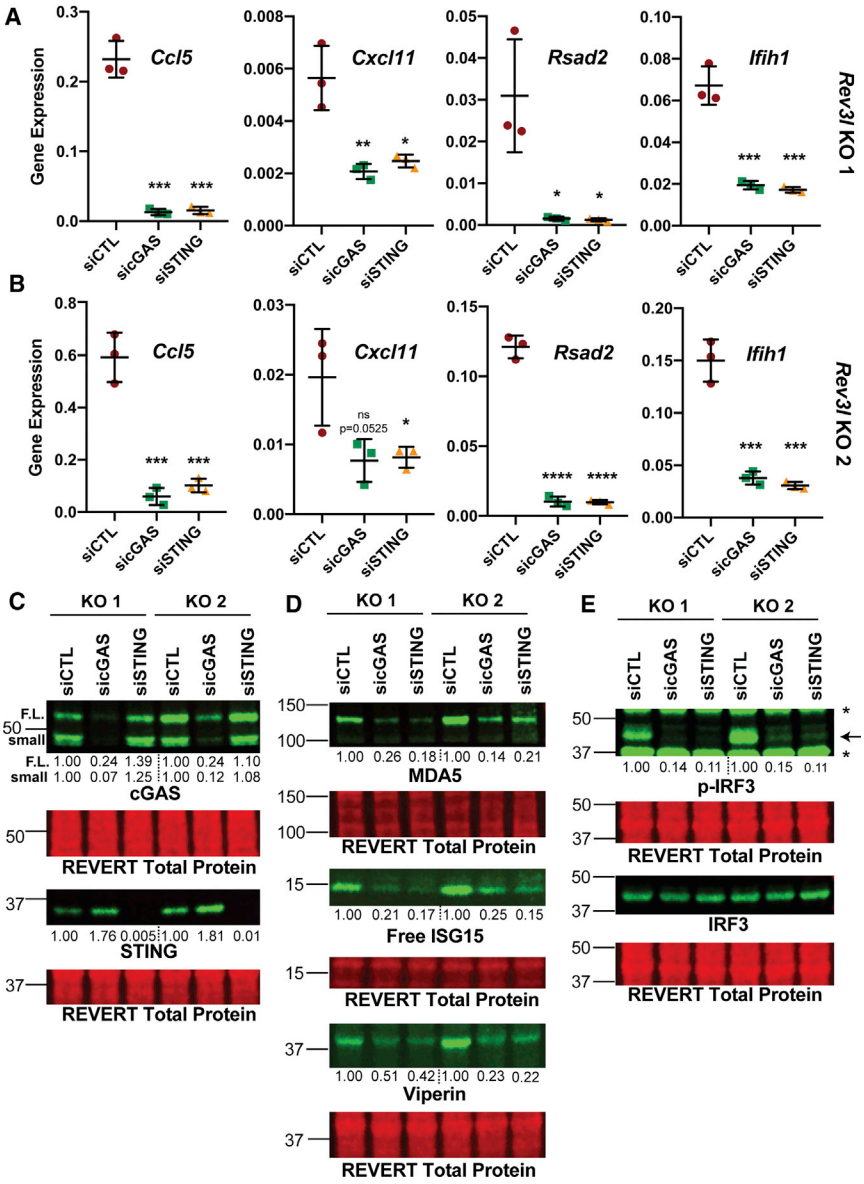


Figure 4. The cGAS-STING axis promotes expression of ISGs due to loss of pol ζ function

(A) Knockdown of cGAS or STING reduces mRNA expression of *CCL5*, *CXCL11*, *RSAD2* (which encodes viperin), and IFIH1 (which encodes MDA5) as detected by qRT-PCR in *Rev3l* KO 1. Gene expression ($2^{-\Delta Ct}$) of selected ISGs normalized to *HPRT* detected by qRT-PCR. Error bars represent standard deviation. Unpaired 2-tailed Student's t test, *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.

(B) Same as in (A), except with *Rev3l* KO 2.

(C) Efficient knockdown of cGAS or STING protein levels. The relative quantification of the full-length cGAS (F.L.), a cGAS smaller species (small), and STING are displayed underneath each blot. The signal intensities were first normalized to total protein.

(D) MDA5, ISG15, and viperin protein levels decrease with cGAS and STING knockdown. The quantification of each protein relative to negative control dicer-substrate short interfering RNA duplex (siCTL) is shown underneath the blot. The signal intensities were first normalized to total protein signal.

(E) Phosphorylation of S888 in mouse (analogous to the human S396) of IRF3 in *Rev3l* KO MEFs decreases with knockdown of cGAS and STING. The relative quantification of p-IRF3 signal intensity-normalized IRF3 is displayed underneath the p-IRF3 blot. The signal intensities were first normalized to total protein signal.

an increase in micronuclei (Gratia et al., 2019; Mackenzie et al., 2016, 2017; Reisländer et al., 2019). There are several sources of DNA damage that may give rise to a sustained response, including cytosolic mitochondrial DNA (West et al., 2015) and cytosolic DNA arising from stalled and processed replication forks (Coquel et al., 2018). DNA stress may continually arise from DNA replication fork collapse in the absence of pol ζ, which could promote the formation of micronuclei and also release small fragments of DNA. Furthermore, some nuclear genes, including pol ζ, control mitochondrial DNA integrity. There is evidence that mitochondrial function is compromised without pol ζ (Singh et al., 2015). It remains to be seen whether micronuclei are the primary instigator of cGAS activation in cells lacking pol ζ or whether they are more of an indicator of DNA degradation.

Our experiments were performed with MEFs immortalized by large T-antigen, which allowed cellular survival in the absence of pol ζ. In addition to blunting p53 activity and inactivating Rb, large T-antigen has been implicated in impairing an interferon response to nucleic acids in primary MEFs and suppressing cGAS-STING activation in human 293 cells through an unknown mechanism or mechanisms (Lau et al., 2015; Reus et al., 2020). Which functions of the large T-antigen enable cells to tolerate *Rev3l* deletion is an interesting but unexplored question. It could involve the poorly understood role of large T-antigen in inhibiting an interferon response or be tied to the blockage of p53 and Rb functions, or both.

An interferon response can result in shutting down cell growth. Specifically, cGAS has been tied to promoting senescence in primary cells (Glück et al., 2017; Yang et al., 2017). Primary MEFs lacking pol ζ only make it approximately two cell divisions before cell growth completely halts, which is accompanied by an increase in senescent cells (Lange et al., 2012). Further demonstrating the essential nature of pol ζ, *Rev3l*-defective primary keratinocytes are unable to proliferate in a mouse model, and germline disruption of *Rev3l* blocks embryo development (Lange et al., 2012, 2013; Martin and Wood, 2019). It is possible that

cGAS-STING drives this severe growth arrest in primary cells due to the loss of pol ζ function.

In addition to widening our understanding of the lengths that cells go to protect themselves from the genomic damage induced by the impairment of pol ζ, these results could have an impact on translational approaches. A potential approach, suggested by experiments in laboratory settings, has been to disrupt pol ζ function to enhance chemotherapeutic effectiveness (Berdis, 2008; Sail et al., 2017; Xu et al., 2013). An inhibitor that impairs the interaction of pol ζ with its regulator REV1 has been developed that sensitizes cancer cells and xenograft tumors to cisplatin treatment (Wojtaszek et al., 2019). This approach would have multiple advantages for therapy by enhancing DNA damage sensitivity and eliminating pol ζ-dependent point mutations during lesion bypass. Our work suggests that such inhibitors may also induce a cGAS-STING response due to the increased chromosomal instability. Promoting the activation of the cGAS-STING pathway as a way to engage the immune system is being widely explored as a new tool in cancer therapy (Hoong et al., 2020). Given our results, it would be exciting to investigate whether pol ζ disruption could also engage the innate immune system in a way that could be exploited for cancer treatment.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Generation of TR4-2 expressing Rev3l KO MEFs
  - RNA isolation
  - Genome-wide mRNA sequencing
  - Sequence mapping
  - Identifying differential expression
  - Analysis of differentially expressed genes
  - Gene expression analysis by quantitative PCR
  - Immunoblotting
  - Knockdown of cGAS and STING protein levels
  - Cisplatin sensitivity
  - Micronuclei frequency
  - Single-molecule RNA *in situ* hybridization
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.108775.

ACKNOWLEDGMENTS

We thank Sarita Bhetawal and Megan Lowery for assistance with cell culture and advice. We thank Dr. Yue Lu, Dr. Bin Liu, and Kevin Lin for invaluable assistance with the RNA expression data analysis, and Dr. Manu Sebastian, Jimi Lynn Young, and Carlos Perez for *in situ* hybridization. We are grateful to Winnie Cheng for assistance with the micronuclei experiments. We thank Dr. Wei Yang (NIH) for the gift of the TR4-2 cDNA. Funding was provided by National Institutes of Health grant no. CA193124, Department of Defense grant no. W81XWH-17-10239, and the J. Ralph Meadows Chair in Carcinogenesis Research to R.D.W. S.K.M. was supported by a CPRIT Research Training Grant award (RP170067). We thank the cores at The University of Texas MD Anderson Cancer Center in the Department of Epigenetics and Molecular Carcinogenesis that contributed to our data collection. RNA sequencing (Next Generation Sequencing Core) and quantitative PCR were supported by CPRIT grants RP120348 and RP170002. *In situ* hybridization performed by the Research Histology, Pathology, and Imaging Core was supported by P30 CA16672-39 DHHS/NCI Cancer Center Support Grant.

AUTHOR CONTRIBUTIONS

S.K.M. designed and performed the experiments, led the data interpretation, and drafted the manuscript. J.T. established the complemented cell line pairs and assisted with the manuscript. R.D.W. assisted with the experimental design, data interpretation, and writing.

DECLARATION OF INTERESTS

R.D.W. is a scientific advisor for Repare Therapeutics and a shareholder.

Received: March 12, 2020
Revised: December 22, 2020
Accepted: January 29, 2021
Published: February 23, 2021

REFERENCES

Ablasser, A., and Chen, Z.J. (2019). cGAS in action: expanding roles in immunity and inflammation. Science *363*, eaat8657.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Bemark, M., Khamlichi, A.A., Davies, S.L., and Neuberger, M.S. (2000). Disruption of mouse polymerase zeta (Rev3) leads to embryonic lethality and impairs blastocyst development in vitro. Curr. Biol. *10*, 1213–1216.

Berdis, A.J. (2008). DNA polymerases as therapeutic targets. Biochemistry *47*, 8253–8260.

Coquel, F., Silva, M.J., Técher, H., Zadorozhny, K., Sharma, S., Nieminuszczy, J., Mettling, C., Dardillac, E., Barthe, A., Schmitz, A.L., et al. (2018). SAMHD1 acts at stalled replication forks to prevent interferon induction. Nature *557*, 57–61.

Daly, J., Bebenek, K., Watt, D.L., Richter, K., Jiang, C., Zhao, M.L., Ray, M., McGregor, W.G., Kunkel, T.A., and Diaz, M. (2012). Altered Ig hypermutation pattern and frequency in complementary mouse models of DNA polymerase ζ activity. J. Immunol. *188*, 5528–5537.

Esposito, G., Godindagger, I., Klein, U., Yaspo, M.L., Cumano, A., and Rajewsky, K. (2000). Disruption of the Rev3l-encoded catalytic subunit of polymerase zeta in mice results in early embryonic lethality. Curr. Biol. *10*, 1221–1224.

Fenech, M., Kirsch-Volders, M., Natarajan, A.T., Surralles, J., Crott, J.W., Parry, J., Norppa, H., Eastmond, D.A., Tucker, J.D., and Thomas, P. (2011). Molecular mechanisms of micronucleus, nucleoplasmic bridge and nuclear bud formation in mammalian and human cells. Mutagenesis *26*, 125–132.

Glück, S., Guey, B., Gulen, M.F., Wolter, K., Kang, T.-W., Schmacke, N.A., Bridgeman, A., Rehwinkel, J., Zender, L., and Ablasser, A. (2017). Innate immune sensing of cytosolic chromatin fragments through cGAS promotes senescence. Nat. Cell Biol. *19*, 1061–1070.

Gonugunta, V.K., Sakai, T., Pokatayev, V., Yang, K., Wu, J., Dobbs, N., and Yan, N. (2017). Trafficking-mediated sting degradation requires sorting to acidified endolysosomes and can be targeted to enhance anti-tumor response. Cell Rep. *21*, 3234–3242.

Gratia, M., Rodero, M.P., Conrad, C., Bou Samra, E., Maurin, M., Rice, G.I., Duffy, D., Revy, P., Petit, F., Dale, R.C., et al. (2019). Bloom syndrome protein restrains innate immune sensing of micronuclei by cGAS. J. Exp. Med. *216*, 1199–1213.

Hoong, B.Y.D., Gan, Y.H., Liu, H., and Chen, E.S. (2020). cGAS-STING pathway in oncogenesis and cancer therapeutics. Oncotarget *11*, 2930–2955.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. *37*, 1–13.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Lange, S.S., Takata, K., and Wood, R.D. (2011). DNA polymerases and cancer. Nat. Rev. Cancer *11*, 96–110.

Lange, S.S., Wittschieben, J.P., and Wood, R.D. (2012). DNA polymerase ζ is required for proliferation of normal mammalian cells. Nucleic Acids Res. *40*, 4473–4482.

Lange, S.S., Bedford, E., Reh, S., Wittschieben, J.P., Carbajal, S., Kusewitt, D.F., DiGiovanni, J., and Wood, R.D. (2013). Dual role for mammalian DNA polymerase ζ in maintaining genome stability and proliferative responses. Proc. Natl. Acad. Sci. USA *110*, E687–E696.

Lange, S.S., Tomida, J., Boulware, K.S., Bhetawal, S., and Wood, R.D. (2016). The Polymerase Activity of Mammalian DNA Pol ζ Is Specifically Required for Cell and Embryonic Viability. PLoS Genet. *12*, e1005759.

Lange, S.S., Bhetawal, S., Reh, S., Powell, K.L., Kusewitt, D.F., and Wood, R.D. (2018). DNA polymerase ζ deficiency causes impaired wound healing and stress-induced skin pigmentation. Life Sci. Alliance *1*, e201800048.

Lau, L., Gray, E.E., Brunette, R.L., and Stetson, D.B. (2015). DNA tumor virus oncogenes antagonize the cGAS-STING DNA-sensing pathway. Science *350*, 568–571.

Lee, Y.S., Gregory, M.T., and Yang, W. (2014). Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. Proc. Natl. Acad. Sci. USA *111*, 2954–2959.

Mackenzie, K.J., Carroll, P., Lettice, L., Tarnauskaitė, Ž., Reddy, K., Dix, F., Revuelta, A., Abbondati, E., Rigby, R.E., Rabe, B., et al. (2016). Ribonuclease H2 mutations induce a cGAS/STING-dependent innate immune response. EMBO J. *35*, 831–844.

Mackenzie, K.J., Carroll, P., Martin, C.A., Murina, O., Fluteau, A., Simpson, D.J., Olova, N., Sutcliffe, H., Rainger, J.K., Leitch, A., et al. (2017). cGAS surveillance of micronuclei links genome instability to innate immunity. Nature *548*, 461–465.

Martin, S.K., and Wood, R.D. (2019). DNA polymerase ζ in DNA replication and repair. Nucleic Acids Res. *47*, 8348–8361.

Mizukami, H., Kim, J.D., Tabara, S., Lu, W., Kwon, C., Nakashima, M., and Fukamizu, A. (2019). KDM5D-mediated H3K4 demethylation is required for sexually dimorphic gene expression in mouse embryonic fibroblasts. J. Biochem. *165*, 335–342.

Mudge, J.M., and Harrow, J. (2015). Creating reference gene annotation for the mouse C57BL6/J genome assembly. Mamm. Genome *26*, 366–378.

O-Wang, J., Kajiwara, K., Kawamura, K., Kimura, M., Miyagishima, H., Koseki, H., and Tagawa, M. (2002). An essential role for REV3 in mammalian cell survival: absence of REV3 induces p53-independent embryonic death. Biochem. Biophys. Res. Commun. *293*, 1132–1137.

Panne, D., McWhirter, S.M., Maniatis, T., and Harrison, S.C. (2007). Interferon regulatory factor 3 is regulated by a dual phosphorylation-dependent switch. J. Biol. Chem. *282*, 22816–22822.

Prabakaran, T., Bodda, C., Krapp, C., Zhang, B.C., Christensen, M.H., Sun, C., Reinert, L., Cai, Y., Jensen, S.B., Skouboe, M.K., et al. (2018). Attenuation of cGAS-STING signaling is mediated by a p62/SQSTM1-dependent autophagy pathway activated by TBK1. EMBO J. *37*, e97858.

Reisländer, T., Lombardi, E.P., Groelly, F.J., Miar, A., Porru, M., Di Vito, S., Wright, B., Lockstone, H., Biroccio, A., Harris, A., et al. (2019). BRCA2 abrogation triggers innate immune responses potentiated by treatment with PARP inhibitors. Nat. Commun. *10*, 3143.

Reus, J.B., Trivino-Soto, G.S., Wu, L.I., Kokott, K., and Lim, E.S. (2020). SV40 Large T Antigen Is Not Responsible for the Loss of STING in 293T Cells but Can Inhibit cGAS-STING Interferon Induction. Viruses *12*, 137.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Sail, V., Rizzo, A.A., Chatterjee, N., Dash, R.C., Ozen, Z., Walker, G.C., Korzhnev, D.M., and Hadden, M.K. (2017). Identification of Small Molecule Translesion Synthesis Inhibitors That Target the Rev1-CT/RIR Protein-Protein Interaction. ACS Chem. Biol. *12*, 1903–1912.

Schenten, D., Kracker, S., Esposito, G., Franco, S., Klein, U., Murphy, M., Alt, F.W., and Rajewsky, K. (2009). Pol zeta ablation in B cells impairs the germinal center reaction, class switch recombination, DNA break repair, and genome stability. J. Exp. Med. *206*, 477–490.

Singh, B., Li, X., Owens, K.M., Vanniarajan, A., Liang, P., and Singh, K.K. (2015). Human REV3 DNA Polymerase ζ Localizes to Mitochondria and Protects the Mitochondrial Genome. PLoS ONE *10*, e0140409.

Tisserand, J., Khetchoumian, K., Thibault, C., Dembélé, D., Chambon, P., and Losson, R. (2011). Tripartite motif 24 (Trim24/Tif1α) tumor suppressor protein is a novel negative regulator of interferon (IFN)/signal transducers and activators of transcription (STAT) signaling pathway acting through retinoic acid receptor α (Rarα) inhibition. J. Biol. Chem. *286*, 33369–33379.

Tomida, J., Takata, K., Lange, S.S., Schibler, A.C., Yousefzadeh, M.J., Bhetawal, S., Dent, S.Y., and Wood, R.D. (2015). REV7 is essential for DNA damage tolerance via two REV3L binding sites in mammalian DNA polymerase ζ. Nucleic Acids Res. *43*, 1000–1011.

Van Sloun, P.P., Varlet, I., Sonneveld, E., Boei, J.J., Romeijn, R.J., Eeken, J.C., and De Wind, N. (2002). Involvement of mouse Rev3 in tolerance of endogenous and exogenous DNA damage. Mol. Cell. Biol. *22*, 2159–2169.

Wang, F., Flanagan, J., Su, N., Wang, L.C., Bui, S., Nielson, A., Wu, X., Vo, H.T., Ma, X.J., and Luo, Y. (2012). RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. J. Mol. Diagn. *14*, 22–29.

West, A.P., Khoury-Hanold, W., Staron, M., Tal, M.C., Pineda, C.M., Lang, S.M., Bestwick, M., Duguay, B.A., Raimundo, N., MacDuff, D.A., et al. (2015). Mitochondrial DNA stress primes the antiviral innate immune response. Nature *520*, 553–557.

Wittschieben, J., Shivji, M.K., Lalani, E., Jacobs, M.A., Marini, F., Gearhart, P.J., Rosewell, I., Stamp, G., and Wood, R.D. (2000). Disruption of the developmentally regulated *Rev3l* gene causes embryonic lethality. Curr. Biol. *10*, 1217–1220.

Wittschieben, J.P., Reshmi, S.C., Gollin, S.M., and Wood, R.D. (2006). Loss of DNA polymerase ζ causes chromosomal instability in mammalian cells. Cancer Res. *66*, 134–142.

Wittschieben, J.P., Patil, V., Glushets, V., Robinson, L.J., Kusewitt, D.F., and Wood, R.D. (2010). Loss of DNA polymerase zeta enhances spontaneous tumorigenesis. Cancer Res. *70*, 2770–2778.

Wojtaszek, J.L., Chatterjee, N., Najeeb, J., Ramos, A., Lee, M., Bian, K., Xue, J.Y., Fenton, B.A., Park, H., Li, D., et al. (2019). A Small Molecule Targeting Mutagenic Translesion Synthesis Improves Chemotherapy. Cell *178*, 152–159.e11.

Xu, X., Xie, K., Zhang, X.Q., Pridgen, E.M., Park, G.Y., Cui, D.S., Shi, J., Wu, J., Kantoff, P.W., Lippard, S.J., et al. (2013). Enhancing tumor cell response to chemotherapy through nanoparticle-mediated codelivery of siRNA and cisplatin prodrug. Proc. Natl. Acad. Sci. USA *110*, 18638–18643.

Yang, C.H., Wei, L., Pfeffer, S.R., Du, Z., Murti, A., Valentine, W.J., Zheng, Y., and Pfeffer, L.M. (2007). Identification of CXCL11 as a STAT3-dependent gene induced by IFN. J. Immunol. *178*, 986–992.

Yang, H., Wang, H., Ren, J., Chen, Q., and Chen, Z.J. (2017). cGAS is essential for cellular senescence. Proc. Natl. Acad. Sci. USA *114*, E4612–E4620.

Yoneyama, M., Suhara, W., and Fujita, T. (2002). Control of IRF-3 activation by phosphorylation. J. Interferon Cytokine Res. *22*, 73–76.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-HA-Tag (C29F4)	Cell Signaling Technologies	Cat#3724; RRID:AB_1549585
Rabbit anti-cGAS (Mouse specific) (D3O8O)	Cell Signaling Technologies	Cat#31659; RRID:AB_2799008
Rabbit anti-STING (D2P2F)	Cell Signaling Technologies	Cat#13647; RRID:AB_2732796
Rabbit anti-MDA-5 (D74E4)	Cell Signaling Technologies	Cat#5321; RRID:AB_10694490
Rabbit anti-Phospho-IRF-3 (Ser396) (4D4G)	Cell Signaling Technologies	Cat#4947; RRID:AB_823547
Rabbit anti-IRF-3 (D83B9)	Cell Signaling Technologies	Cat#4302; RRID:AB_1904036)
Rabbit anti-STAT1 (D1K9Y)	Cell Signaling Technologies	Cat#14994; RRID:AB_2737027
Rabbit anti-TBK1 (D1B4)	Cell Signaling Technologies	Cat#3504; RRID:AB_2255663
rabbit anti-Phospho-TBK1 (S172) (D52C2)	Cell Signaling Technologies	Cat#5483; RRID:AB_10693472
Mouse anti-Viperin	Millipore	Cat#MABF106 RRID:AB_11203644
Goat anti-Rabbit 800CW	LI-COR	Cat#926-32211; RRID:AB_621843
Goat anti-Mouse 800CW	LI-COR	Cat#827-08364; RRID:AB_10793856
Critical commercial assays		
REVERT Total Protein Stain Kit	LI-COR	Cat#926-11010
ATPlite Luminescence Assay System	Perkin Elmer	Cat#6016941
MycoAlert Mycoplasma Detection Kit	Lonza Bioscience	Cat#LT07-218
Deposited data		
Raw and analyzed RNA sequencing data	This paper	GEO: GSE163313
Full immunoblots for the cropped immunoblots in Figures 1, 3, 4, and S1.	This paper	Mendeley Data: https://dx.doi.org/10.17632/5348dspthz.1
Experimental models: cell lines		
Rev3l HET 1 MEFs: Rev3L(+/Δ) mT/mG (+/−) MEFs from mouse embryo 3, TAg immortalized clone 6 (or 3(+)6)	Lange et al., 2012, 2016	N/A
Rev3l HET 2 MEFs: same as above except clone 10 (or 3(+)10)	Lange et al., 2012, 2016	N/A
Rev3l KO 1 MEFs: Rev3L(-/Δ) mT/mG (+/−) MEFs from mouse embryo 4, TAg immortalized clone 5 (or 4(-)5)	Lange et al., 2012, 2016	N/A
Rev3l KO 2 MEFs: same as above except clone 10 (or 3(+)10)	Lange et al., 2012, 2016	N/A
Rev3l HET + EV #Clone: HET 1 + pCDH-EF1-Flag-HA-Empty, clones: 4, 8 and 11	This study	N/A
Rev3l KO + EV #Clone: KO 1 + pCDH-EF1-Flag-HA-Empty, clones: 3, 4, and 15	This study	N/A
Rev3l KO + TR4-2 #Clone: KO 1 + pCDH-EF1-Flag-HA-TR4-2, clones: 8, 18, and 21	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
"siCTL": Negative Control DsiRNA Duplex	Integrated DNA Technologies	Cat#51-01-14-03
"sicGAS": cGAS (MB21D1) DsiRNA Duplex	Integrated DNA Technologies	Design ID: mm.Ri.Mb21d1.13.1
"siSTING": STING (TMEM173) DsiRNA Duplex	Integrated DNA Technologies	Design ID: mmRi.Tmem173.13.2
For primers used for qPCR in this study, see Table S2.	See Table S2	N/A
Recombinant DNA		
pLEXm-His_8-MBP-TR4-2	Lee et al., 2014	N/A
pCDH-EF1-Flag-HA-TR4-2	This study	N/A
pCDH-EF1-Flag-HA-Empty	Tomida et al., 2015	N/A
Software and algorithms		
TopHat (version 2.0.10)	Kim et al., 2013	N/A
HTSeq package (version 0.6.0)	Anders et al., 2015	N/A
R/Bioconductor package edgeR (version 3.8.6)	Robinson et al., 2010	N/A
Prism (version 8)	GraphPad	N/A

RESOURCE AVAILABILITY

Lead contact

Additional information and requests for resources and reagents can be directed to and fulfilled by Lead Contact, Dr. Richard Wood (rwood@mdanderson.org).

Materials availability

This study did not generate unique reagents.

Data and code availability

The datasets produced in this study have been uploaded to Gene Expression Omnibus. Accession number for the RNA sequencing data is GEO: GSE163313. The full immunoblot images in Figures 1, 3, 4, and S1 have been uploaded to Mendeley Data: https://dx.doi.org/10.17632/5348dspthz.1

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The immortalized *Rev3l* knockout (KO) and heterozygous (HET) mouse embryonic fibroblast (MEF) parental cell lines in this study were as described (Lange et al., 2012). In brief, the *Rev3l* KO MEFs were generated by large T-antigen immortalization of primary MEFs isolated from mouse embryos with one null *Rev3l* and one floxed allele, followed by *ex vivo* Ad-Cre deletion of the floxed allele, and clonal selection to ensure a homogeneous population with the genotype $Rev3l^{-/\Delta}$. The *Rev3l* HET MEFs were isolated in the same manner expect starting MEFs from embryos with one wild-type and one floxed allele, resulting in MEFs with the genotype $Rev3l^{+/\Delta}$. The floxed allele was generated by placing loxP sites flanking two conserved exons, corresponding to residues 2776-2860, that contain the three conserved catalytic aspartate residues of REV3L (Wittschieben et al., 2010). The null allele replaced these two exons with a lacZ-neo^R cassette (Wittschieben et al., 2000). In this study two sets of clones were analyzed. Cells were grown in DMEM (Sigma #5796), 10% fetal bovine serum (FBS) and 1 X penicillin/streptomycin (GIBCO #15140122). Cell lines were genotyped for *Rev3l* allele status as described (Lange et al., 2012), and karyotyping confirmed the presence of mouse chromosomes only. The RNA sequencing data show that the $Rev3l^{+/\Delta}$ MEF cell line expresses characteristic Y chromosome genes (*Kdm5d, Uty, Eif2s3y, Ddx3y*) indicating that it was derived from a male embryo (Mizukami et al., 2019). The $Rev3l^{-/\Delta}$ MEF does not express these genes but expresses *Xist,* indicating derivation from a female embryo. All cell lines were negative for mycoplasma infection as shown by regular checks with a MycoAlert Mycoplasma Detection Kit (Lonza Bioscience).

METHOD DETAILS

Generation of TR4-2 expressing Rev3l KO MEFs

The TR4-2 construct, a gift from Dr. Wei Yang (Lee et al., 2014), was cloned into the pCDH backbone with an N-terminal Flag-HA tag as the same previously described for the full length *Rev3l* cDNA (Tomida et al., 2015). The pCDH-Flag-HA-TR4-2 or pCDH-Flag-HA empty vector was stably inserted into the MEF cell lines using lentiviral infection as previously described (Lange et al., 2016). Three single clones were isolated for analysis and continually grown in 1 μg/mL of puromycin to ensure stable integration of the construct.

RNA isolation

RNA was isolated from 1.5 × 10^6 (Figure 2, RNA seq) or 2.5 × 10^5 cells (Figure 3, gene expression) or 1 × 10^6 cells (Figure 4, gene expression), using the RNeasy Kit (QIAGEN # 74104) following the manufacturer's instructions including the on-column DNase I digestion (QIAGEN #79254).

Genome-wide mRNA sequencing

mRNA libraries were prepared using the Illumina TruSeq Stranded mRNA kit following manufacturer's instructions and 75 base paired end sequencing was run on the Illumina HiSeq 3000. Three biological replicates were prepared for each condition. 41-46 million pairs of reads were generated per sample. Each pair of reads represents a cDNA fragment from the library.

Sequence mapping

The reads were mapped to the mouse genome mm10 by TopHat (version 2.0.10) (Kim et al., 2013). By reads, the overall mapping rate is 96.0%–97.1%. 93.9%–95.3% fragments have both ends mapped to the mouse genome.

Identifying differential expression

GENCODE Release M19 (Mudge and Harrow, 2015) using htseq-count from HTSeq package (version 0.6.0) (Anders et al., 2015) was used to determine the number of fragments in identified genes. A normalized estimate of gene expression was calculated as fragments per kilobase of exon model per million reads mapped (FPKM). If genes had less than 10 fragments in all the samples, they were excluded from differential expression analysis. R/Bioconductor package edgeR (version 3.8.6) (Robinson et al., 2010) was used to statistically assess differentially expressed genes. Genes with FDR (false discovery rate) $\leq$ 0.05, fold change $\geq$ |0.5| and length > 200 bp were called as differentially expressed. More stringent cut-offs for fold change were used for various analyses as described below.

Analysis of differentially expressed genes

For gene ontology analysis, genes upregulated more than 4-fold with an FDR < 0.05 were entered into DAVID 6.8 on 02/22/20 and the top 10 GOTERMS_BP_Direct were plotted based on -log(p value) (Huang et al., 2009a, 2009b). For upstream regulator analysis, differential expressed genes both upregulated and downregulated more than 4-fold with an FDR < 0.05. were entered into IPA (QIAGEN Inc., https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-ipa/). Given that upregulated genes are overrepresented in the Rev3l KO differentially expressed genes, this results in an expected bias for upstream regulators that also predominately result in the upregulation of targeted genes which we see in our dataset. Since, bias-corrected z-scores are not reported for upstream regulators with a |bias term| > 0.5, we've reported the uncorrected z-score here. Given that the majority of differentially expressed genes are upregulated in our dataset, bias is to be expected for transcription regulators that primarily induces expression of genes.

Gene expression analysis by quantitative PCR

High Capacity cDNA Reverse Transcription (Applied Biosciences #4368814) was used to prepare cDNA from 1000 ng of total RNA from each sample. qPCR was run iTaq Universal SYBR Green Supermix (Biorad #1725121) on the Applied Biosystems 7900HT Fast Real-Time PCR System. Gene expression of calculated using the $2^{-\Delta Ct}$ method normalizing to the *Hprt* gene. See STAR methods table for the primers used for mouse *Hprt* and target genes (Mackenzie et al., 2016; West et al., 2015; Yang et al., 2007).

Immunoblotting

Cells were lysed (3 million cells /100 μL) in lysis buffer (Tris-HCl: 50 mM, pH 7.5, NaCl 250 mM, EDTA 1 mM, Triton X-100 0.1%, 1 X Protease/Phosphatase Inhibitor Cocktail CST #5872) for 30 min on ice with mixing every 10 min. Debris was pelleted by centrifugation at 15,000 x g for 20 min at 4°C. Protein amounts were quantified using Biorad Protein Assay (Biorad #500-0006) and a bovine serum albumin standard curve (Biorad #500-0007) following manufacturer's instructions. The samples were denatured using 4 x loading buffer (LI-COR #928-40004). 25 μg of protein / well and Precision Plus Protein All Blue Standards (Biorad #161-0373) were run on 4%–20% polyacrylamide gels (Biorad #4561096) in 1 x Tris/Glycine/SDS buffer (Biorad #161-0772). Protein was transferred to Immobilon-FL PVDF Membrane (Millipore #IPFL00010) in 1 x Tris/Glycine buffer (Biorad #161-0772), 20% methanol. After transfer, membranes were dried. Total protein was measured using REVERT total protein stain kit (LI-COR #926-11010) following the manufacturer's instructions. Membranes were blocked for 1 h in 0.5X Odyssey Blocking Buffer (OBB, LI-COR # 927-50000) in Tris Buffered

Saline (TBS) and then incubated in primary antibody overnight at 4°C. The primary antibodies were used at the following dilutions in 0.5XOBB/TBS/0.2% Tween-20. Rabbit anti-HA-Tag (C29F4) (CST #3724, 1:1000), Rabbit anti-cGAS (Mouse specific) (D3O8O) (CST #31659, 1:1000), rabbit anti-STING (D2P2F) (CST # 13647, 1:1000), rabbit anti-MDA-5 (D74E4) (CST # 5321, 1:1000), rabbit anti-ISG15 (CST # 2743, 1:500), rabbit anti-Phospho-IRF-3 (Ser396) (4D4G) (CST #4947, 1:1000), rabbit anti-IRF-3 (D83B9) (CST #4302, 1:1000), mouse anti-viperin (Millipore #MAB106, 1:250), rabbit anti-STAT1 (D1K9Y) (CST: #14994, 1:1,000), rabbit anti-TBK1 (D1B4) (CST: #3504, 1:1,000), and rabbit anti-Phospho-TBK1 (S172) (D52C2) (CST: #5483, 1:1,000).

After primary incubation membranes were washed three times in TBS/0.1% Tween-20, and incubated for 1 h in secondary antibody either goat anti-Rabbit 800CW (LI-COR #926-32211) or goat anti-mouse 800CW (Li-Cor #827-08364) diluted 1:20,000 in 0.5X OBB /TBS /0.2% Tween-20/0.01% SDS. After primary incubation, membranes were washed three times in TBS/0.1% Tween-20, rinsed with TBS, then dried and imaged on the LI-COR Odyssey FC.

Knockdown of cGAS and STING protein levels

600,000 cells were seeded into 10 cm plates. The next day, cells were transfected with 1 nM of the appropriated siRNA using Lipofectamine RNAiMAX (ThermoFisher #13778150) following the manufacturer's protocols. The following dicer-substrate short interfering RNAs were used: siCTL (IDT: Negative Control DsiRNA # 51-01-14-03), sicGAS (IDT: DsiRNA Duplex mm.Ri.Mb21d1.13.1), and siSTING (IDT: DsiRNA Duplex mmRi.Tmem173.13.2). After 48 h, cells were harvested in paired pellets for RNA and protein analysis and flash frozen in liquid nitrogen.

Cisplatin sensitivity

10,000 cells were seeded in triplicate into 96 well plates. The following day cells were treated with the appropriate concentration of cisplatin or vehicle control. After 48 h, the relative survival of the cells was calculated by using the ATPlite assay (Perkin Elmer #606016943) following the manufacturer's instructions.

Micronuclei frequency

20,000 cells per chamber were seeded on four-chamber slides. After 48 h, cells were fixed in 100% methanol and slides were stained with DAPI. The slides were mounted and imaged. Micronuclei were counted as discrete units distinct from the nucleus.

Single-molecule RNA *in situ* hybridization

Skin tissues were from previously described 7-10 wk old mice of genotypes *BK5.Cre*; *Rev3l*$^{+/lox}$ and *BK5.Cre*; *Rev3l*$^{-/lox}$. In these mice, the floxed allele of *Rev3l* is deleted specifically in epithelial cells where the keratin 5 promoter (BK5) is active (Lange et al., 2018). Sections of formalin fixed, paraffin-embedded archival blocks were stained with hematoxylin and eosin. RNA *in situ* hybridization experiments were performed using RNAscope® (Wang et al., 2012). Paired double-Z oligonucleotide probes were designed against target RNA using custom software. The following probes were used Mm-Ccl5, cat no. 469601, NM_013653.3, 11 pairs, nt 4-527 and Mm-Cxcl10-C2, cat no. 408921-C2, NM_021274.2, 16 pairs, nt 11-1012. The RNAscope® 2.5 HD Duplex Reagent Kit (Advanced Cell Diagnostics, Newark, CA) was used according to the manufacturer's instructions. Each sample was quality controlled for RNA integrity with probes specific to the housekeeping genes *PPIB and POLR2A.* Negative control background staining was evaluated using a probe specific to the *Bacillus subtilis dapB* gene. Brightfield images were acquired using an Aperio ScanScope CS microscope using a 40x objective. Fields of interfollicular basal epithelial cells with defined nuclei were scored. Positive cells usually contained only one hybridization focus. Images from three individual mice of each genotype were visualized with QuPath and scored (15 fields per mouse, total > 800 epithelial cells/mouse).

QUANTIFICATION AND STATISTICAL ANALYSIS

Unpaired two-tailed Student's t tests were applied to qPCR, micronuclei frequency, and *in situ* hybridization results using Prism 8. The significance threshold was set at $\alpha < 0.05$. On all graphs the mean is graphed and the error bars represent $\pm$ SD. For the graph of gene expression (FPKM) from RNA sequencing data (Figure 2D), the error bars represent SEM.

Cell Reports, Volume *34*

Supplemental information

Disruption of DNA polymerase ζ engages an innate immune response

Sara K. Martin, Junya Tomida, and Richard D. Wood

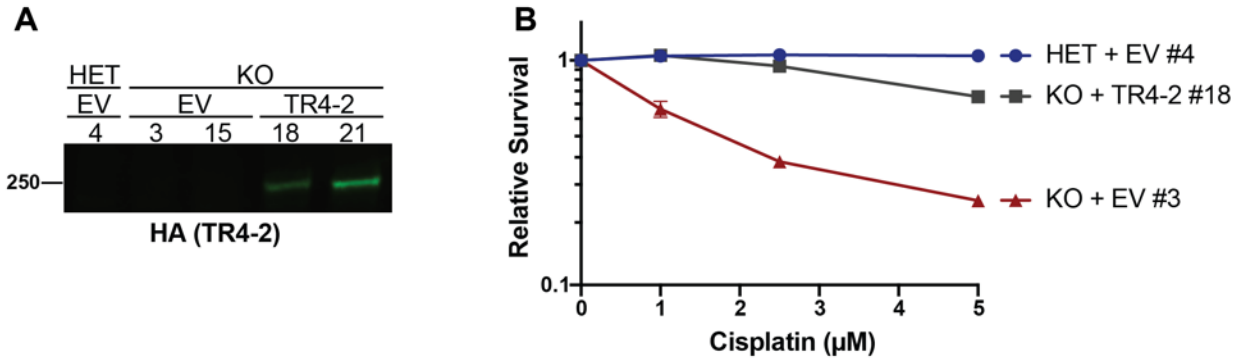


Figure S1. Shortened REV3L construct rescues phenotypes of pol ζ disruption in an additional set of clones, Related to Figure 1

(A) Stable expression of TR4-2 with an N-terminal Flag-HA tag in *Rev3l* KO clones as detected by HA immunoblot. For REVERT total protein loading control see Fig 1 and S1 in Mendeley Data: doi: 10.17632/5348dspthz.1

(B) Stable expression of TR4-2 in *Rev3l* KO MEF clones reverses hypersensitivity to cisplatin. MEFs were exposed to the indicated cisplatin concentrations for 48 hr and relative survival was quantified with the ATPlite assay.

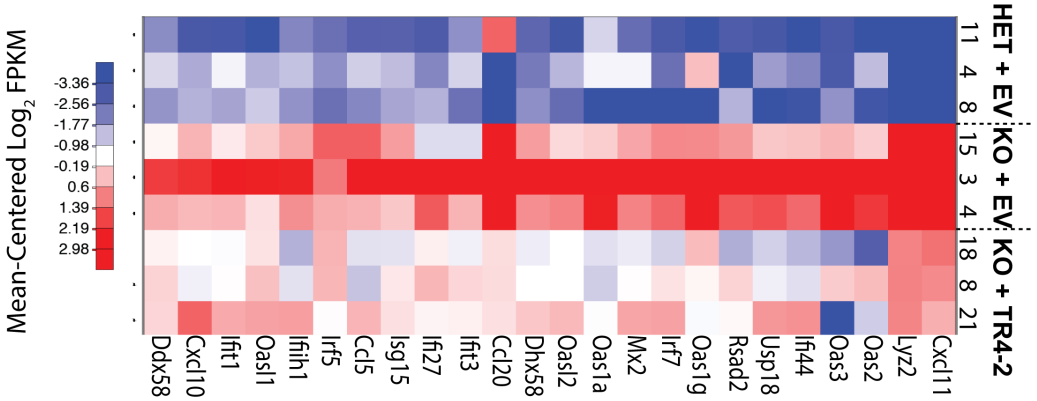


Figure S2. Heatmap of expression of 24 interferon-stimulated genes, Related to Figure 2
The heatmap was generated with the same RNA-seq data, genes and samples as in Fig 2D and Table S1.

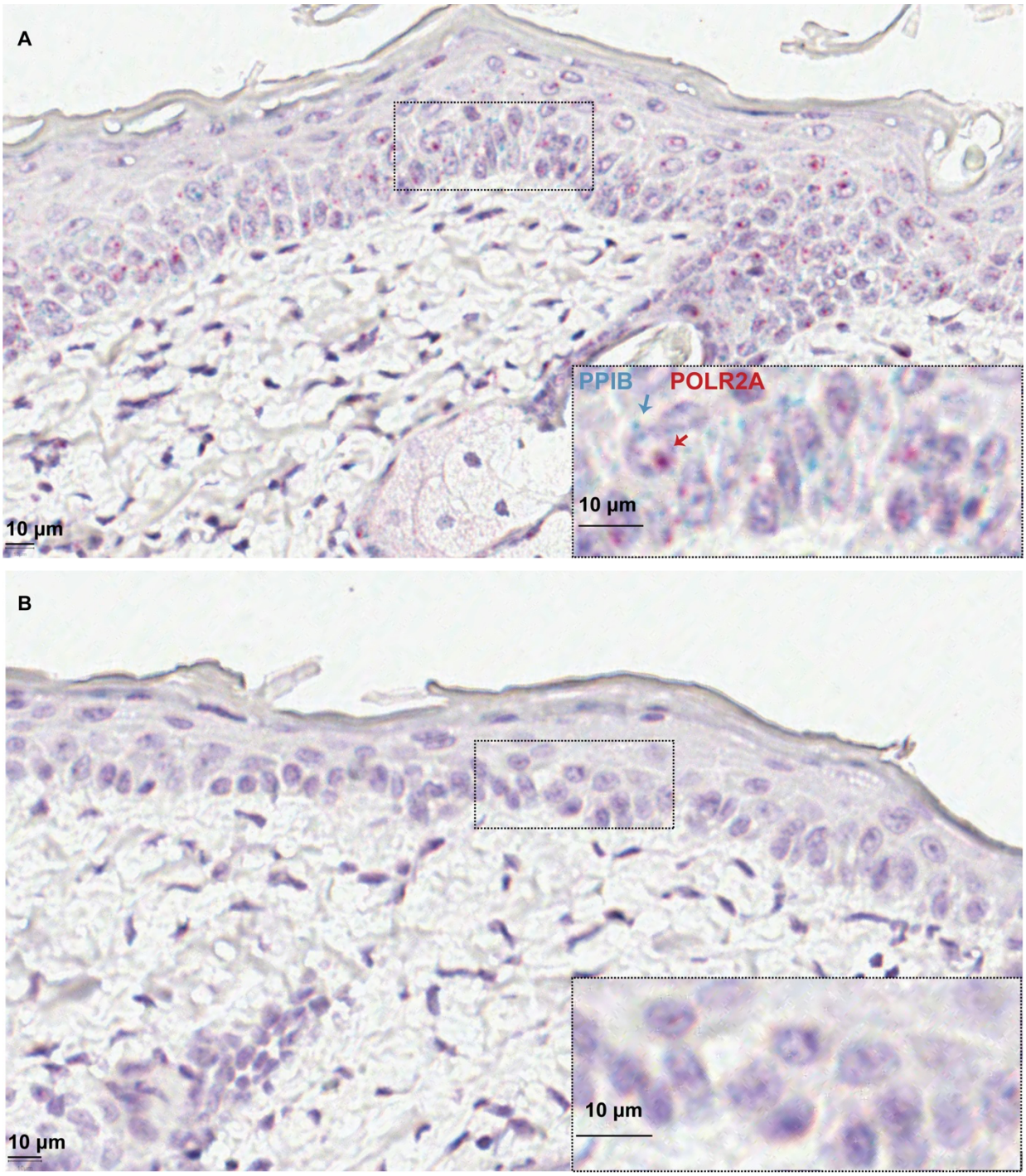


Figure S3. Positive and Negative controls for single-molecule RNA in situ hybridization, Related to Figure 3.
(A) Positive Control. Skin section of a keratin 5 conditional heterozygous *Rev3l* mouse (BK.5 Cre *Rev3l*$^{(+/lox)}$) was analyzed with RNA in situ hybridization probes specific for the housekeeping genes *Ppib* (blue dots) *and Polr2a* (red dots). Arrows point to one example of probe signals.
(B) Negative control. Single-molecule RNA in situ hybridization in skin of keratin 5 conditional heterozygous *Rev3l* mouse (BK.5 Cre *Rev3l*$^{(+/lox)}$) using a probe specific for the *Bacillus subtillus dapB* gene. As expected, no signal is detected.

	Rev3l KO + EV vs *Rev3l* HET + EV		*Rev3l* KO + TR4-2 vs *Rev3l* KO + EV	
Gene Name	Log2 Fold Change	FDR	Log2 Fold Change	FDR
Cxcl11	10.08783	1.43E-08	-5.0312	0.000362
Lyz2	8.870484	1.31E-09	-4.772444	1.91E-05
Oas2	8.267013	3.67E-05	-6.567388	0.002339
Oas3	7.959514	3.31E-06	-6.542195	0.000362
Ifi44	7.479926	1.44E-06	-4.825919	0.003736
Usp18	6.572054	7.54E-07	-4.059615	0.005592
Rsad2	6.196355	3.40E-05	-4.282616	0.01059
Oas1g	6.161099	0.000707	-5.145596	0.011973
Irf7	5.839463	5.85E-11	-3.041723	0.001875
Mx2	4.957664	0.000188	-3.258566	0.041769
Oas1a	4.870041	0.000783	-4.308707	0.009657
Oasl2	4.500346	9.79E-08	-2.551515	0.010097
Dhx58	4.068542	2.92E-09	-2.264432	0.004736
Ccl20	4.171942	0.004204	-4.143873	0.013561
Ifit3	3.995032	9.81E-05	-2.691874	0.032516
Ifi27	3.896872	1.04E-07	-1.961613	0.035
Isg15	3.587676	2.05E-07	-2.203871	0.007627
Ccl5	3.546222	3.94E-06	-2.176328	0.021028
Irf5	3.182247	5.70E-05	-0.922754	0.53922
Ifih1	3.155781	1.23E-06	-1.783102	0.032395
Oasl1	3.143506	0.000161	-1.113838	0.441074
Ifit1	3.131239	6.84E-05	-1.706951	0.113497
Cxcl10	3.042178	8.26E-06	-0.810127	0.544102
Ddx58	2.516276	2.12E-07	-1.110754	0.114366

Table S1. Log2 Fold Change of 24 selected interferon stimulated genes, Related to Figure 2

Species	Target	Direction	Sequence (5' to 3')	Source
Mouse	*Hprt*	Forward	CTGGTGAAAAGGACCTCTCG	Mackenzie et al., 2016
Mouse	*Hprt*	Reverse	CAAGGGCATATCCAACAACA	Mackenzie et al., 2016
Mouse	*Ccl5*	Forward	ACGTCAAGGAGTATTTCTACAC	Mackenzie et al., 2016
Mouse	*Ccl5*	Reverse	GATGTATTCTTGAACCCACT	Mackenzie et al., 2016
Mouse	*Cxcl11*	Forward	AGGAAGGTCACAGCCATAGC	Yang et al., 2007
Mouse	*Cxcl11*	Reverse	CGATCTCTGCCATTTTGACG	Yang et al., 2007
Mouse	*Rsad2*	Forward	ATAGTGAGCAATGGCAGCCT	West et al., 2015
Mouse	*Rsad2*	Reverse	AACCTGCTCATCGAAGCTGT	West et al., 2015
Mouse	*Ifih1*	Forward	CGGAAGTTGGAGTCAAAGC	West et al., 2015
Mouse	*Ifih1*	Reverse	TTTGTTCAGTCTGAGTCATGG	West et al., 2015
Mouse	*Isg15*	Forward	CTAGAGCTAGAGCCTGCAG	West et al., 2015
Mouse	*Isg15*	Reverse	AGTTAGTCACGGACACCAG	West et al., 2015
Mouse	*Irf7*	Forward	CAATTCAGGGGATCCAGTTG	West et al., 2015
Mouse	*Irf7*	Reverse	AGCATTGCTGAGGCTCACTT	West et al., 2015

Table S2. Primers used for qPCR in this study, Related to STAR Methods.

Article

SOURCE DATA | TRANSPARENT PROCESS | OPEN ACCESS

THE EMBO JOURNAL

DNA polymerase zeta contributes to heterochromatin replication to prevent genome instability

Barbara Ben Yamin[1,†], Sana Ahmed-Seghir[1,†,§], Junya Tomida[2,†,¶], Emmanuelle Despras[1,††], Caroline Pouvelle[1], Andrey Yurchenko[3], Jordane Goulas[1], Raphael Corre[1], Quentin Delacour[1], Nathalie Droin[4], Philippe Dessen[5], Didier Goidin[6], Sabine S Lange[2,‡‡], Sarita Bhetawal[2], Maria Teresa Mitjavila-Garcia[7], Giuseppe Baldacci[8], Sergey Nikolaev[3], Jean Charles Cadoret[8,‡], Richard D Wood[2,‡] & Patricia L Kannouche[1,*,‡]

Abstract

The DNA polymerase zeta (Polζ) plays a critical role in bypassing DNA damage. REV3L, the catalytic subunit of Polζ, is also essential in mouse embryonic development and cell proliferation for reasons that remain incompletely understood. In this study, we reveal that REV3L protein interacts with heterochromatin components including repressive histone marks and localizes in pericentromeric regions through direct interaction with HP1 dimer. We demonstrate that Polζ/REV3L ensures progression of replication forks through difficult-to-replicate pericentromeric heterochromatin, thereby preventing spontaneous chromosome break formation. We also find that *Rev3l*-deficient cells are compromised in the repair of heterochromatin-associated double-stranded breaks, eliciting deletions in late-replicating regions. Lack of REV3L leads to further consequences that may be ascribed to heterochromatin replication and repair-associated functions of Polζ, with a disruption of the temporal replication program at specific loci. This is correlated with changes in epigenetic landscape and transcriptional control of developmentally regulated genes. These results reveal a new function of Polζ in preventing chromosome instability during replication of heterochromatic regions.

Keywords DNA replication; heterochromatin; replication timing; REV3L; TLS polymerase

Subject Categories Chromatin, Transcription & Genomics; DNA Replication, Recombination & Repair

DOI 10.15252/embj.2020104543 | Received 22 January 2020 | Revised 20 August 2021 | Accepted 28 August 2021

The EMBO Journal (2021) e104543

Introduction

It is well established that DNA polymerase delta (Polδ) and epsilon (Polε) catalyze the high-fidelity duplication of the genome (reviewed in Burgers & Kunkel, 2017). However, an emerging concept is that translesion synthesis (TLS) DNA polymerases, known for their error-prone lesion-bypass properties, can also facilitate synthesis of non-damaged DNA. TLS polymerases may be employed when replicative DNA polymerases are not able to pass through stalled sites of replication at structured DNA sequences and/or hard-to-replicate genomic regions (reviewed in Tsao & Eckert, 2018). For example, DNA polymerase eta (Polη) and zeta (Polζ) are required to maintain common fragile site stability (CFS) in human cells (Bergoglio *et al*, 2013; Bhat *et al*, 2013; Despras *et al*, 2016). It is not

1 CNRS-UMR9019, Equipe labellisée Ligue Contre le Cancer, Gustave Roussy, Paris-Saclay Université, Villejuif, France
2 Department of Epigenetics and Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center and The University of Texas MD Anderson Cancer Center UT Health Graduate School of Biomedical Sciences, Houston, TX, USA
3 INSERM U981, Gustave Roussy, Université Paris Saclay, Villejuif, France
4 Genomics Facility, Gustave Roussy, Villejuif, France
5 Bioinformatics Core Facility, Gustave Roussy, Villejuif, France
6 Life Sciences and Diagnostics Group, Agilent Technologies France, Les Ulis, France
7 INSERM U1197, Paul Brousse Hospital, Paris-Saclay Université, Villejuif, France
8 Institut Jacques Monod, UMR7592, CNRS and University of Paris, Paris, France
*Corresponding author (Lead contact). Tel: +33 142 114 030; E-mail: patricia.kannouche@gustaveroussy.fr
†These authors contributed equally to this work as first authors
‡These authors contributed equally to this work as senior authors
§Present address: Department of Radiation Oncology, Memorial Sloan-Kettering Cancer Center, New York, NY, USA
¶Present address: Department of Biological Sciences, University of North Carolina at Charlotte, Charlotte, NC, USA
††Present address: Sorbonne Université, INSERM UMRS 938, Équipe Instabilité des Microsatellites et Cancer, Équipe Labellisée par la Ligue Nationale Contre le Cancer et SIRIC CURAMUS, Centre de recherche Saint Antoine, Paris, France
‡‡Present address: Texas Commission on Environmental Quality, Austin, TX, USA

known whether specialized DNA polymerases are recurrently recruited to DNA more broadly during S-phase to assist replicative DNA polymerases for replicating unconventional DNA structures preventing thus genome instability at the cost of increased point mutations. Intriguingly, growing evidence suggests that replication timing influences genomic mutation rates with an increasing gradient of single-nucleotide substitutions that correlate with late-replicating regions (Stamatoyannopoulos *et al*, 2009; Koren *et al*, 2012; Polak *et al*, 2015), but the causative underlying mechanisms remain elusive.

Polζ is a TLS polymerase complex in eukaryotes, consisting of four subunits: Rev3, the catalytic subunit, Rev7 that enhances Rev3 activity, and two subunits shared with the replicative polymerase Polδ (Baranovskiy *et al*, 2012; Johnson *et al*, 2012; Makarova *et al*, 2012; Lee *et al*, 2014). In contrast to other TLS enzymes that belong to the Y-family of polymerases, Polζ belongs to the B-family that includes the highly accurate replicative DNA Polδ and Polε (Gan *et al*, 2008). Polζ lacks an intrinsic 3′–5′ exonuclease activity, making this TLS polymerase error-prone with a spontaneous mutation rate that is 10- to 100-fold greater than that of replicative DNA polymerases in yeast (McCulloch & Kunkel, 2008; Stone *et al*, 2009). Polζ is a key player in translesion DNA synthesis by elongating primer termini that are positioned opposite base damage and non-instructional lesions (Johnson *et al*, 2000). This DNA polymerase has been extensively characterized in the budding yeast *S. cerevisiae*, showing that spontaneous as well as damage-induced mutagenesis is largely dependent on Polζ (Makarova & Burgers, 2015).

Counterparts of the yeast *REV* genes have been identified in other eukaryotes. Mouse and human Rev3-like (*Rev3l*) orthologs have a large extra segment which is not conserved in yeast *REV3* and are thus about twice the size of the 173-kDa yeast Rev3 (350 and 353 kDa, respectively). Despite the established participation of *Rev3l* in important cellular processes (Martin & Wood, 2019), the role of REV3L protein is incompletely understood and studies have been hampered by the inability to detect this large protein in cells. Polζ is unique among TLS polymerases in mammalian cells, because inactivation of *Rev3l* gene leads to embryonic lethality in mice (Esposito *et al*, 2000; Wittschieben *et al*, 2000; Van Sloun *et al*, 2002). $Rev3l^{-/-}$ mouse embryonic stem cells are not viable, and primary mouse embryonic fibroblasts (MEFs) obtained from conditional *Rev3l* knockouts show genome instability and growth defects without external damage to DNA (Wittschieben *et al*, 2006; Lange *et al*, 2012). All these data underscore an essential role of Polζ in mammalian cells.

In this study, we investigated the biological function of Polζ during S-phase under normal growth conditions. We show that Polζ/REV3L is able to interact with heterochromatin through direct interaction with HP1 dimer and ensures progression of replication forks through the hard-to-replicate pericentromeric heterochromatin, thus preventing spontaneous DNA and chromosome break formation. We also find that lack of REV3L compromises heterochromatin-associated DSB repair and shows increased number of genomic deletions in late-replicating regions. Lack of REV3L leads to further consequences that may be ascribed to heterochromatin replication-associated functions of Polζ, with a disruption of the temporal replication program at specific loci associated with changes in epigenetic and transcriptional landscape. These results show that Polζ helps maintain genome stability by contributing to heterochromatin replication.

Results

S-phase progression is altered in $Rev3l^{-/-}$ MEFs

Polζ is critical for proliferation of normal primary cells, suggesting that Polζ participates in unchallenged DNA replication (Lange *et al*, 2012). To gain cellular and molecular insights into its function during DNA replication, we first investigated the dynamics of S-phase progression in the absence of REV3L, the catalytic subunit of Polζ. Unchallenged $Rev3l^{-/-}$ and $Rev3l^{+/+}$ mouse embryonic fibroblasts (MEFs) (Lange *et al*, 2012) were pulse-labeled with the thymidine analog 5′-bromo-2′-deoxyuridine (BrdU) for 15 min, released into fresh medium, and harvested at the indicated time points. FACS analysis showed that cell cycle progression was altered in $Rev3l^{-/-}$ MEFs (Fig 1A). Indeed, $Rev3l^{-/-}$ cells exhibited a significant higher percentage in late S/G2-phase at 4, 6, and 8 h as compared to control cells. This is mirrored by a depletion of BrdU-positive cells in G1/early S at the same time periods (Fig 1A, see red arrows in upper panel and histograms in lower panel). Moreover, analysis of BrdU-negative cells showed that $Rev3l^{-/-}$ cells tend to accumulate in G2-phase as compared to $Rev3l^{+/+}$ (Fig 1A, see green arrows in upper panel). Similar observations were obtained using REV3L-depleted human cells (Fig EV1). These results indicate that cell cycle progression in $Rev3l^{-/-}$ cells is impaired in late S-phase and G2/M-phase and prompted us to investigate the patterns of spatio-temporal replication in *Rev3l* null MEFs.

S-phase can be divided into different stages by immunofluorescence observation of characteristic thymidine analog incorporation patterns corresponding to early, mid-, or late S-phase (Dimitrova & Berezney, 2002; Guenatri *et al*, 2004). To analyze the spatial organization of DNA replication in $Rev3l^{-/-}$ MEFs, cells were pulsed-labeled with EdU and the replication patterns typically observed in mouse cells were identified and quantified (Guenatri *et al*, 2004). Representative replication patterns for each category are represented: early S-phase pattern with a high density of small foci distributed throughout the nucleus (stage I), early–mid-S-phase pattern with foci which become bigger and less abundant than stage I (stage II), mid–late S-phase pattern with a typical ring-shaped labeling around pericentromeric heterochromatin domains (stage III), and late S-phase pattern with a few big dots located mainly at the nuclear periphery and in heterochromatin area (stage IV) (Fig 1B, upper panel). We found that in the absence of REV3L, the proportion of cells with an early replication spatial pattern (stage I) decreased whereas the proportion of cells in mid- and late S-phase increased (Fig 1B, middle and lower panels). A similar effect was found after depletion of REV3L in human cells (Fig 1C). These results suggest that lack of REV3L could affect the temporal control of DNA replication.

Loss of REV3L delays replication timing in specific genomic loci

We thus examined the replication timing in $Rev3l^{+/+}$ and $Rev3l^{-/-}$ cells after a few (p5 and p7 in $Rev3l^{+/+}$ and $Rev3l^{-/-}$ cells, respectively) or serial passages (p60) in cell culture (Fig EV2A). To perform genome-wide profiling of replication timing (RT), cells were pulse-labeled with BrdU and sorted into early and late S-phase fractions by flow cytometry (see methods (Fernandez-Vidal *et al*, 2014)). Newly synthesized DNA of each fraction was BrdU-immunoprecipitated and

specifically labeled (Cy3 for early fraction and Cy5 for late fraction) before co-hybridization on microarrays. The replication timing of genomic domains was obtained by measuring the log2-ratio of early versus late fractions and analyzed using the START-R program (Haddadj *et al*, 2020). We first compared the replication timing obtained at different passages for each cell line. For *Rev3l*$^{+/+}$ MEF, there was a

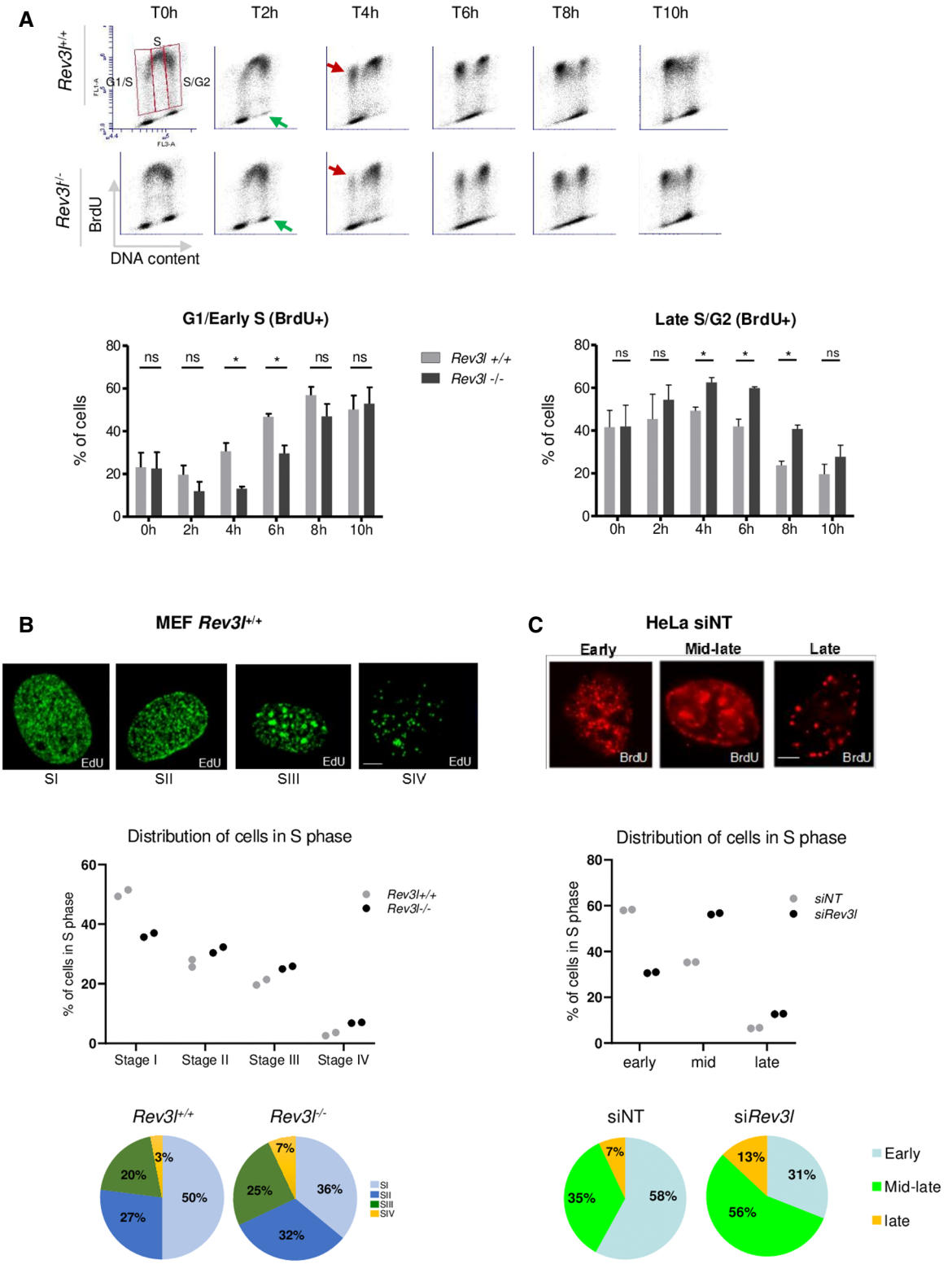


Figure 1.

Figure 1. S-phase progression is impaired in *Rev3l*$^{-/-}$ MEFs.

A *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs were pulse-labeled with BrdU prior to harvesting and analyzed by flow cytometry at different time points. The analysis was focused on S-phase divided into three parts: G1/early S, middle S, and late S/G2 BrdU^{+} cells. Red arrow represents BrdU^{+} cells that re-entry in S-phase, and green arrow shows BrdU^{-} cells in G2-phase. Histograms represent the percentage of cells in G1/early and late S/G2-phase after BrdU pulse. Error bars represent standard error for three independent experiments. ns: not significant, *$P < 0.05$ by Student's *t*-test.

B *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs were pulse-labeled with EdU for 15 min, permeabilized, fixed, and stained for EdU incorporation (green). S-phase sub-stages from I to IV were evaluated by visual inspection of the cycling population (> 300 EdU^{+} cells, top panel). Scale bar = 5 μm. Dot plots and pie charts show the relative proportion (percentage of total S) of each sub-stage from I to IV (middle and bottom panel, respectively). Each dot represents the mean of two technical replicates.

C Seventy-two hours after transfection with non–targeting (NT) or *Rev3l* siRNA, HeLa cells were pulse-labeled with BrdU for 15 min, permeabilized, fixed, and stained for BrdU (red). As in (B), S-phase sub-stages were evaluated by visual inspection of the cycling population (> 300 BrdU^{+} cells, top panel). Scale bar = 5 μm. Dot plots and pie charts show the relative proportion (percentage of total S) of early, middle–late, and late S-phase (middle and bottom panel, respectively). Each dot represents the mean of two technical replicates.

1.9% difference between p5 and p60. For *Rev3l*$^{-/-}$ cells, we observed only 0.2% difference between p7 and p60, suggesting that the replication timing remains stable during cell culture passages in both cell lines, and can be considered as two independent replicates.

We next compared the replication timing between *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEF at early passage. The START-R analysis identified a set of genomic compartments that changed replication timing in response to *Rev3l* inactivation (Fig 2A). About 5.7% of the whole genome was affected; 19.2% (in bp) of these regions were advanced in timing, and 80.8% of regions were delayed (Fig EV2B). A major effect of *Rev3l* loss was boundary shifts as exemplified in Fig 2A, corresponding to a delay in regions that lie between early- and late-replicating domains called temporal transition regions (TTR) by Gilbert and colleagues (Hiratani *et al*, 2008). We observed that 67% of disturbed domains fall in TTR (corresponding to 83% of delayed regions).

We then explored the correlation between genomic regions that changed replication timing in *Rev3l*$^{-/-}$ cells (called disturbed regions) and the coverage for GC and LINE-1 contents (Fig 2B) and also for H3K27ac, H3K4me3, and H3K27me3 epigenetic marks (Fig 2C). For each parameter, we observed a molecular signature intermediate between that of early and late domains, strengthening our hypothesis that these disturbed regions correspond to TTR. In *Rev3l*$^{-/-}$ cells, 18.7% of TTR were disturbed. These results reveal that *Rev3l* inactivation in MEFs induces changes in the temporal replication program, especially in specific genomic regions located in TTR and suggest that REV3L/Polζ might contribute to replicate these specific loci.

To confirm these observations, we examined the replication timing in human cells depleted for REV3L. For that, HeLa cells were transfected with non-targeting siRNA (siNT) or siRNA against *Rev3l*. Sixty hours later, cells were pulse-labeled with BrdU for 1.5 h and processed as for MEFs. Analysis by the START-R program revealed 116 disturbed regions in REV3L-depleted cells, corresponding to approximatively 5.4% of the whole genome with more than 70% of delayed regions (Fig EV2C and D). We found that depletion of REV3L in human cells affected mostly genomic domains that replicate late in S-phase (Fig EV2E). Thus, the specific replicating domains disturbed after REV3L depletion are not exactly the same in MEFs and HeLa genomes (TTR versus Late, respectively).

Rev3l inactivation results in epigenetic changes with down-regulation of numerous developmentally regulated genes

Links between transcription and replication timing have been well documented (Hiratani *et al*, 2008; Sima *et al*, 2019). Genome-wide analyses have identified a strong correlation between early replication and high transcriptional potential (MacAlpine *et al*, 2004; Farkash-Amar *et al*, 2008; Hiratani *et al*, 2008). We reasoned that changes in replication timing in the absence of *Rev3l* may impact the transcriptional program. We therefore performed microarray-based transcriptome profiling from *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs (Fig 3A). Genes altered by at least threefold between conditions were considered (FDR < 0.05). Analysis of microarray data indicated a total of 317 genes with mRNA expression altered by ≥ 3-fold in *Rev3l*-deficient cells as compared to control cells, with 112 genes up-regulated and 205 genes that were down-regulated. Genes displaying a down-regulation in *Rev3l*$^{-/-}$ MEFs had significant enrichment for Gene Ontology terms related to organ and system development (Fig 3B). Prominent repressed genes are in the *HoxB*

Figure 2. Loss of REV3L disrupts the replication timing in specific genomic loci.

A Cells were pulse-labeled with BrdU for 1.5 h and sorted by flow cytometry in two fractions, S1 and S2, corresponding to early and late S-phase fractions, respectively. Neo-synthesized DNA was immunoprecipitated with BrdU antibodies. Early and late neo-synthesized DNAs were labeled with Cy3 and Cy5, respectively, and hybridized on microarrays. After processing analysis with the START-R software, replication-timing profiles were obtained from two replicates (see Fig EV2A). Shown are the zoomed microarray profiles of the timing of replication on chromosome 1 (136.6–143.5 Mb), chromosome 5 (127.3–134.7 Mb), and chromosome 9 (46.1–51.3 Mb) from *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs overlaid. Blue lines represent replication timing from *Rev3l*$^{+/+}$ MEFs, and red lines represent replication timing from *Rev3l*$^{-/-}$ MEFs. Genomic regions displaying significant difference between *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs by START-R are indicated by a pink line ($P < 0.01$).

B Analysis of GC and Line-1 content in early, mid, late, and TTR compared with disturbed replicating regions found in *Rev3l*$^{-/-}$ MEFs. Bar in boxplot represents the median, and red points represent the mean. The limit of the boxes corresponds to the 0.25–0.75 quartiles with whiskers extending to the maximum value of 1.5 times the interquartile range. Graphs show data from two biological experiments.

C Analysis of the active marks H3K27ac and H3K4me3 and facultative heterochromatin mark H3K27me3 content in early, mid, late, and TTR compared with disturbed replicating regions found in *Rev3l*$^{-/-}$ MEFs. Bar in boxplot represents the median, and red points represent the mean. The limit of the boxes corresponds to the 0.25–0.75 quartiles with whiskers extending to the maximum value of 1.5 times the interquartile range. Graphs show data from two biological experiments.

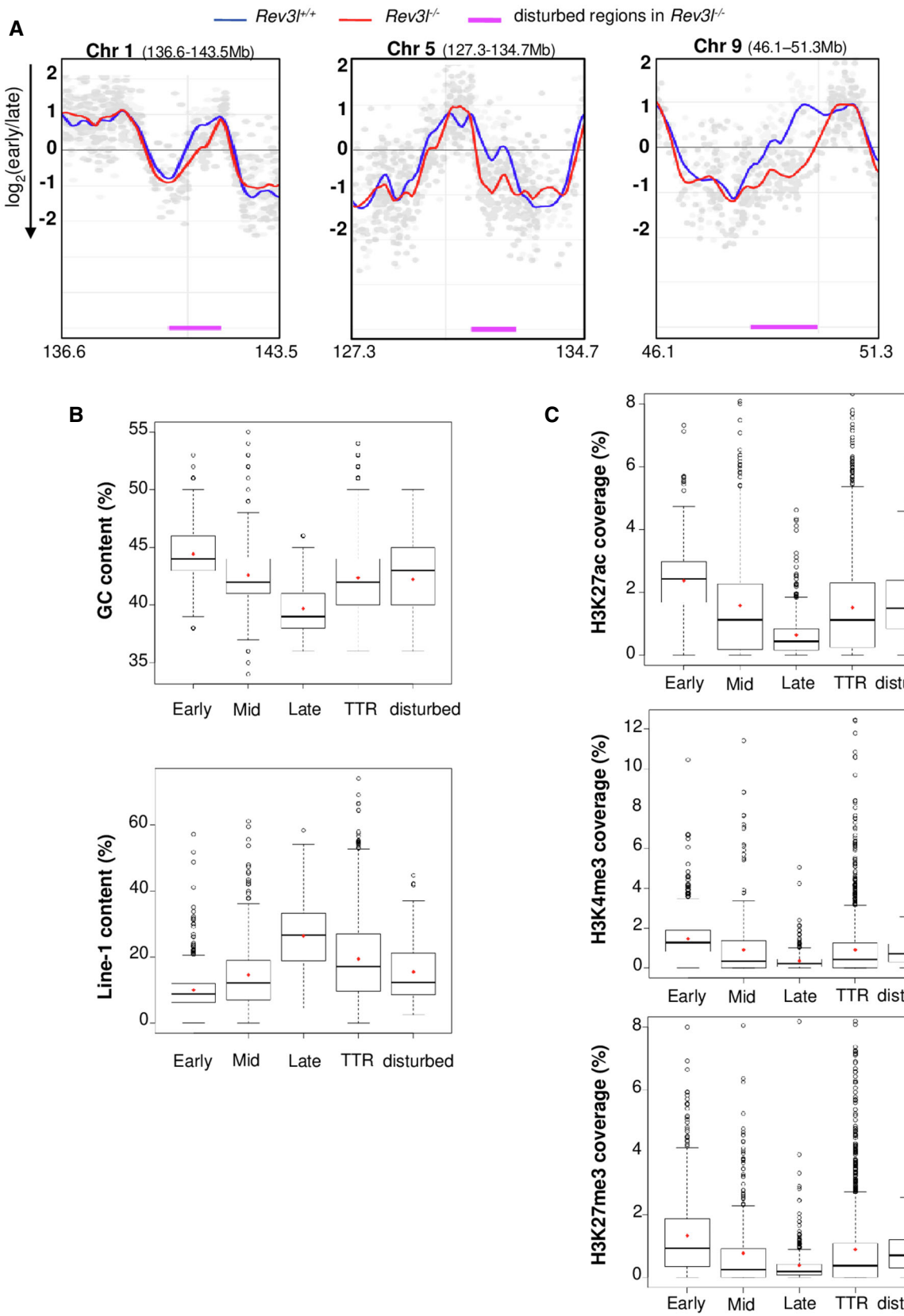

A
Rev3l+/+
Rev3l-/-
disturbed regions in Rev3l-/-
Chr 1 (136.6-143.5Mb)
Chr 5 (127.3-134.7Mb)
Chr 9 (46.1–51.3Mb)
log2(early/late)
136.6
143.5
127.3
134.7
46.1
51.3
B
GC content (%)
Line-1 content (%)
C
H3K27ac coverage (%)
H3K4me3 coverage (%)
H3K27me3 coverage (%)
Early
Mid
Late
TTR
disturbed

Figure 2.

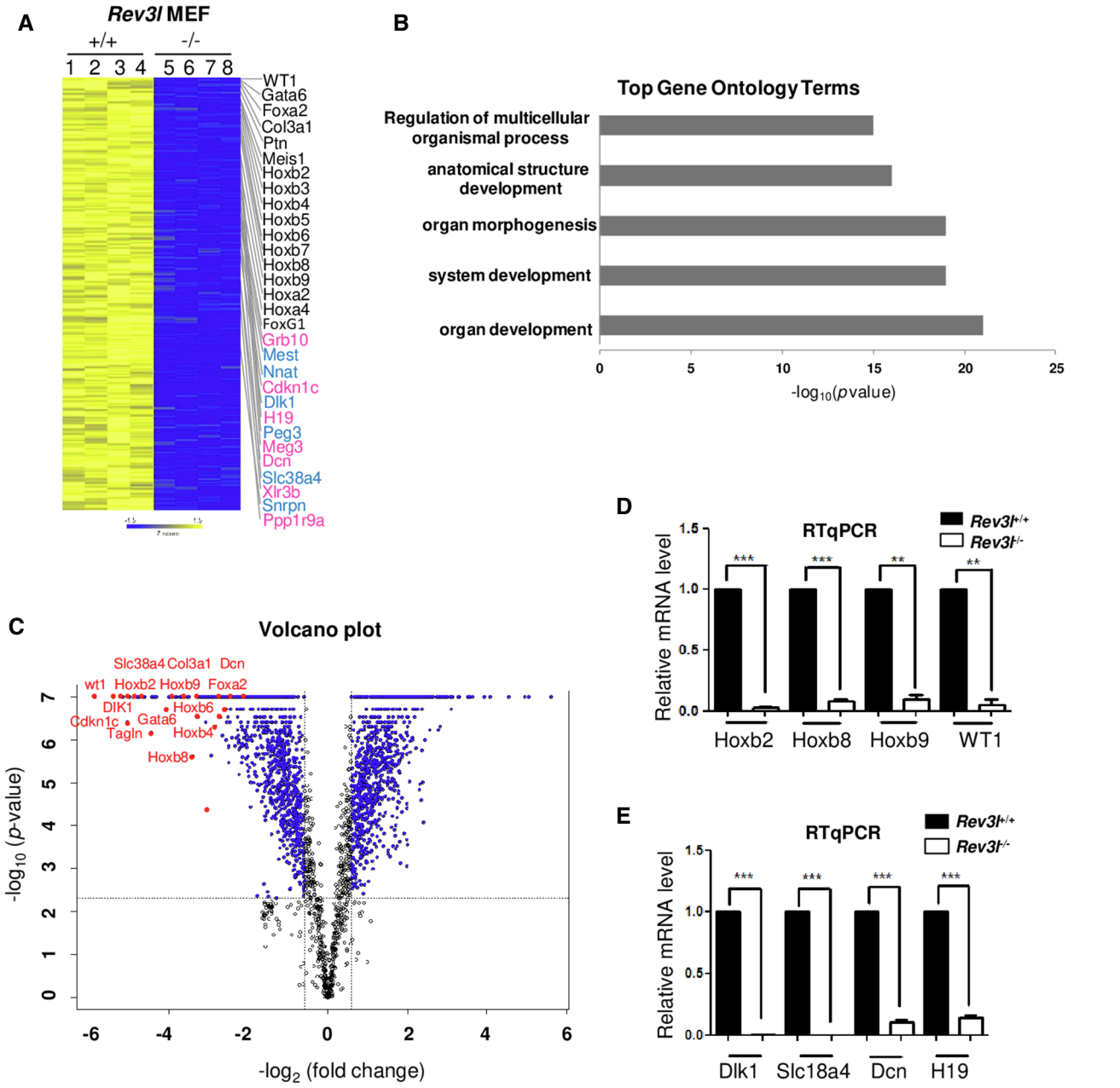


Figure 3. Inactivation of *Rev3l* impairs expression of numerous developmentally and imprinted genes.

A Heat map showing log2 fold change in differentially expressed genes in *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs from 2 independent biological experiments (Rev3l$^{+/+}$: samples 1 and 2, Rev3l$^{-/-}$: samples 5 and 6) with 2 technical replicates (samples 3,4,7,8: technical repeats from samples 1,2,5,6, respectively). Several developmentally regulated genes and imprinting genes (paternally in blue and maternally in pink) down-regulated in *Rev3l*$^{-/-}$ MEFs are indicated on the right. Yellow and blue indicate high and low mRNA expression levels, respectively.

B Top five biological process gene ontology (GO) terms of genes found down-regulated in *Rev3l*$^{-/-}$ MEFs transcriptome analysis.

C Volcano plot shows in red genes involved in development and imprinting with high fold changes ≥ 3 (FDR < 0.05).

D Relative mRNA levels of four genes involved in development (*Hoxb2*, *Hoxb8*, *Hoxb9*, and *WT1*) were validated through qRT–PCR from *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEF samples. The data were normalized to the amount of GAPDH mRNA. Error bars indicate standard error of the mean for three independent experiments. ***$P < 0.001$ and **$P < 0.005$ by Student's *t*-test.

E Relative mRNA levels of genes involved in imprinting (paternal expressed genes: Dlk1 and Slc18a4 and maternal expressed genes: Dcn1 and H19) were validated through qRT–PCR from *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEF samples. The data were normalized to the amount of *GAPDH* mRNA. Error bars indicate standard error of the mean for three independent experiments. ***$P < 0.001$ by Student's *t*-test.

Source data are available online for this figure.

gene cluster, including *Hoxb2* to *Hoxb9* as well as two *HoxA* genes (Fig 3A and C). *Rev3l* inactivation also resulted in the decrease in expression of approximately 14 genes from imprinted loci (corresponding to 6.8% of down-regulated genes,). These included paternally expressed genes *Dlk1*, *Snrpn*, *Peg3*, and *Mest* (Fig 3A, blue) and maternally expressed genes *H19*, *Dcn*, *Meg3*, and *Cdkn1c* (Fig 3A, pink). *WT1* is a well-characterized developmental gene that is mutated in Wilms' tumor and was identified as being imprinted (Schwienbacher *et al*, 2000). This gene was significantly down-regulated by at least 50-fold in $Rev3l^{-/-}$ MEFs as compared to control cells. qRT–PCR from independent cell cultures of $Rev3l^{-/-}$ and $Rev3l^{+/+}$ MEFs confirmed the down-regulation of multiple *HoxB* and imprinted genes (Fig 3D and E).

We next asked whether genes showing altered mRNA expression in the absence of REV3L are located in genomic regions that displayed a RT delay. For this, we integrated the data on gene expression and genome-wide profiling of replication timing. We found that of the 24 genes (corresponding to 7.6% of deregulated genes) located in these specific genomic domains (Table 1), all showed less expression in REV3L-defective cells. These observations are consistent with the fact that late-replicating genes are often silenced. However, the majority of deregulated genes do not fall within disturbed regions of replication timing, indicating that down-regulation of genes might be a secondary consequence of REV3L inactivation. Nevertheless, these data suggest that loss of *Rev3l* is correlated to dysregulation of genes involved in growth and development.

In addition to gene expression, various genomic features are linked to replication timing. These include histone modification, DNA methylation, DNA repeat sequences, ordered chromatin structure, and nuclear compartmentalization (Aladjem, 2007; Hiratani *et al*, 2008). Changes in replication timing as well as dysregulation of gene expression in response to *Rev3l* inactivation prompted us to assess whether these modifications were associated with variations in global levels of histone modifications. Total histones from $Rev3l^{+/+}$ and $Rev3l^{-/-}$ MEFs were screened by immunoblotting with a panel of histone modification-specific antibodies. *Rev3l* inactivation resulted in a fourfold up-regulation of histone H3K27 tri-methylation (Fig 4A) and a more than two-fold increase in H3K9me3 and H3K4me3. While H3K27me3 and H3K9me3 are repressive marks, H3K4me3 is associated with active genes.

We then examined repressive histone modifications (H3K27me3 and H3K9me3) at gene loci down-regulated in $Rev3l^{-/-}$ MEFs (including Hoxb2, Hoxb8, WT1). By ChIP-qPCR, we found that H3K27me3 and H3K9me3 levels were higher in $Rev3l^{-/-}$ than in $Rev3l^{+/+}$ MEFs (Fig 4B). In contrast, we observed a trend toward

Table 1. Loci displaying both delay in RT and misregulation in $Rev3l^{-/-}$ MEFs.

Domain	Chr	Size (bp)	Replication timing domain	Gene	Misregulation (fold change[a])
1	1	1,042,476	TTR	Pax3	−3.57
2	2	2,573,606	TTR	Plcb4	−4.17
3	6	6 628 128	TTR	Frmd4b	−3.13
4	8	671,566	TTR	Hand2	−7.69
5	8	641,165	TTR	Lpl	−27.03
6	10	567,222	TTR	Ptprb	−5.56
6	10	567,222	TTR	Tspan8	−7.14
7	11	920,184	TTR	Meis1	−5.56
8	12	487,988	TTR	Foxg1	−20.41
9	12	548,987	TTR	Dlk1	−6.25
9	12	548,987	TTR	Meg3	−5.26
9	12	548,987	TTR	Mirg	−6.67
10	12	1,372,498	TTR	Mycn	−5.88
11	13	1,179,925	TTR	Sfrp4	−6.25
12	14	671,779	TTR	Wnt5a	−4.55
13	15	1,054,278	TTR	Fbxl7	−8.33
14	15	1,145,389	TTR	Has2	−4.55
15	17	808,766	TTR	Rftn1	−12.35
16	18	402,349	TTR	Gata6	−12.82
17	18	539,392	TTR	Pcdhb20	−3.45
17	18	539,392	TTR	Pcdhb21	−5.56
18	18	1,232,897	TTR	Ppp2r2b	−8.33
19	19	509,162	TTR	Aldh1a1	−11.33
19	19	509,162	TTR	Aldh1a7	−3.57

[a]FDR < 0.05.

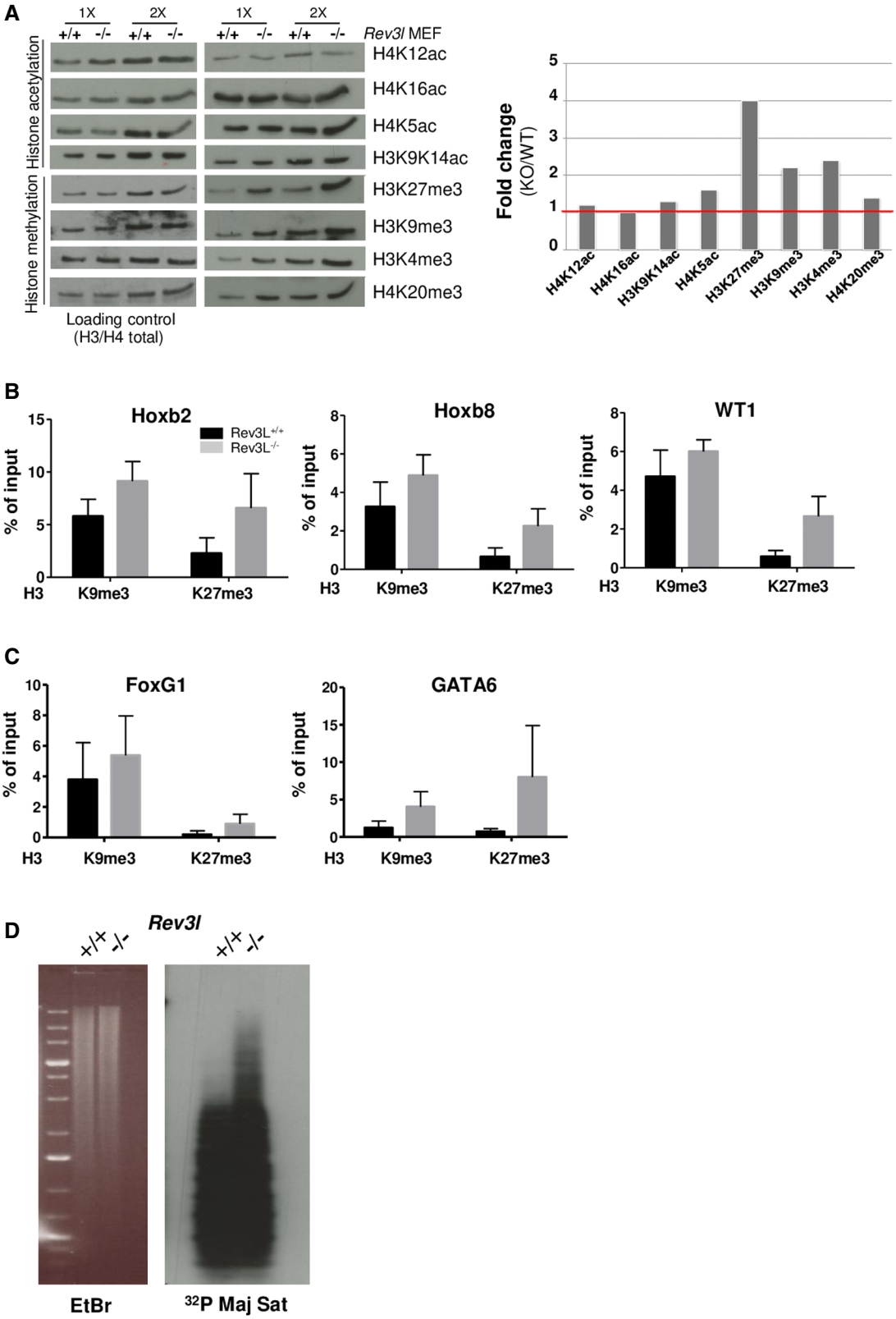


Figure 4.

Figure 4. Loss of REV3L results in epigenetic changes.

A Total histones were acid extracted from *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs. Samples were analyzed by Western blot using indicated antibodies specific to histone marks (right panel). An anti-pan H4 or H3 was used as loading control (left panel). Histograms show WT/KO *Rev3l* fold change relative to the immunoblot shown on the top panel (using 1× sample intensity).

B Histone mark levels were examined at selected genes repressed in *Rev3l*$^{-/-}$ MEFs (*Hoxb2*, *Hoxb8*, *WT1*). Histograms represent enrichment of H3K9me3 and H3K27me3 at indicated loci assessed by ChIP-qPCR in *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs. Error bars indicate standard error of the mean for three independent experiments.

C H3K9me3 and H3K27me3 mark levels were evaluated at selected genes that localized in genomic loci displaying replication timing delay and down-regulated in *Rev3l*$^{-/-}$ MEFs (*FoxG1*, *GATA6*). Histograms represent enrichment of H3K9me3 and H3K27me3 assessed by ChIP-qPCR at indicated loci in *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs. Error bars indicate standard error of the mean for three independent experiments.

D Southern blot analysis of genomic DNA extracted from *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs and digested with the CpG methylation-sensitive enzyme MaeII (5′ -ACGT- 3′). The membrane was hybridized with radiolabeled probes specific to major satellites. The presented data are representative of two repeats (see Source Data file for this figure).

Source data are available online for this figure.

reduced active marks H3K9ac and H3K27ac in *Rev3l*$^{-/-}$ cells (Appendix Fig S1). We also observed enrichment of repressive histone marks in gene loci that showed a delay in replication timing in *Rev3l*$^{-/-}$ cells such as *FoxG1* and *GATA6* genes (Fig 4C), supporting the idea that the down-regulation of gene expression observed in *Rev3l*$^{-/-}$ cells might be caused by epigenetic silencing.

Constitutive heterochromatin exhibits a chromatin landscape marked by high levels of H3K9me3, DNA methylation, and histone hypoacetylation. The pericentromere is a heterochromatic domain that provides a structural scaffold for centromere formation and plays a crucial role in genome stability (Allshire & Karpen, 2008). In mouse cells, pericentromeric heterochromatin consists of ~10^5 major satellite DNA repeats that are methylated, decorated by H3K9me3, and enriched in the heterochromatin protein 1 alpha (HP1α) (Maison *et al*, 2002). Since we observed an up-regulation of H3K9me3 levels in *Rev3l*-deficient cells, we next investigated the DNA methylation pattern at major satellite repeats. Genomic DNA was extracted from *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs, digested with methylation-sensitive restriction enzymes MaeII, and DNA blots hybridized with major satellite probes. In *Rev3l*-deficient cells, a fraction of DNA satellite was resistant to enzymatic digestion, indicating an increase in DNA methylation in the absence of REV3L (Fig 4D). Therefore, our data suggest that REV3L can play a role in the duplication of DNA repeat-rich pericentromeric regions known to be replicated in mid-to-late S-phase in mouse cells (Guenatri *et al*, 2004). Loss of REV3L results in modification of epigenetic marks in such heterochromatic regions.

Direct interaction of REV3L with HP1 dimer allows REV3L recruitment at pericentromeric regions

We reasoned that if REV3L is implicated in heterochromatin replication, it may interact with heterochromatin components. As no commercially available antibodies can detect endogenous REV3L, we generated a MEF cell line with an alteration at the genomic *Rev3l* locus so that endogenous REV3L protein harbors a 3X-Flag tag (Flag-REV3L) (Fig EV3A). We first verified that we were able to detect endogenous Flag-REV3L protein in this cell line (Fig EV3B and C). We then used co-immunoprecipitation to examine interaction of REV3L with heterochromatin proteins. We found that both HP1α and ATRX, a chromatin remodeling protein known to localize in heterochromatin (Eustermann *et al*, 2011), co-immunoprecipitated with REV3L. As positive control, we confirmed that REV7, the regulatory subunit of Polζ, was also co-immunoprecipitated (Fig 5A).

To strengthen these observations, we performed a proximity ligation assay (PLA) that detects protein associations within 40 nm. As expected, we observed specific PLA signals between REV7 and Flag-REV3L in the nucleus. We also detected PLA signals between HP1α and Flag-REV3L (Figs 5B and EV3D). We then tested whether REV3L was spatially close to particular modified histone marks. We found REV3L enriched in proximity to H3K9me3, H3K27me3, H3K4me3, H4K20me3, but not to H3K9ac (Figs 5C and EV3E). Together, these results show that REV3L localizes in heterochromatin regions and interacts with heterochromatin proteins and repressive histone marks.

To better characterize the molecular mechanisms linking Polζ/REV3L to heterochromatin, we examined the subcellular localization of REV3L. In mouse cells, pericentromeric heterochromatin (PHC) domains are easily identified by their intense DNA staining with the dye DAPI. They are clearly discernible during interphase when major satellites from different chromosomes associate with form clusters (so called chromocenters) that co-localize with HP1α (Maison *et al*, 2002). To explore REV3L subcellular localization, we generated a panel of eGFP-tagged REV3L truncated proteins (Fig 5D, upper panel). All constructs were detected mainly in the nucleus and localized into chromocenters, visualized by dense DAPI staining (Fig 5D, lower panel). Interestingly, the smaller REV3L construct (REV3L$^{761–1,029}$) consisting of 268 residues was sufficient to target most, if not all the fusion protein into the nucleus, suggesting that this polypeptide fragment contains a putative nuclear localization signal (NLS); the amino acid sequence KSRKRRKMSKKLPP at position 960–973 is a good candidate (Nguyen & Lavenier, 2009). More importantly, this small construct is also sufficient to target fusion protein into chromocenters. We then confirmed that location in PHC by visualizing the colocalization of REV3L$^{761–1,029}$ with HP1α and H3K9me3 by immunofluorescence (Fig 5E). In the mouse, the association of HP1α at pericentromeric heterochromatin is severely compromised in cells lacking the H3 Lys9 methyltransferases, *Suv39h1* and *Suv39h2* (Peters *et al*, 2001). We therefore investigated REV3L localization in *Suv39h* double-null (dn) cells. HP1α was no longer enriched at PHC in the *Suv39h* dn cells as previously reported (Maison *et al*, 2002). Likewise, REV3L$^{761–1,029}$ displayed a diffuse staining throughout the nucleus (Fig EV4A). We also observed that REV3L$^{761–1,029}$ staining was disrupted when cells were treated with trichostatin A (TSA), a histone deacetylase inhibitor which reversibly increases the acetylation level of histone H3K9 at pericentromeric regions. In TSA-treated cells, HP1α dissociates from heterochromatic domains (Taddei *et al*, 2001), and REV3L$^{761–1,029}$

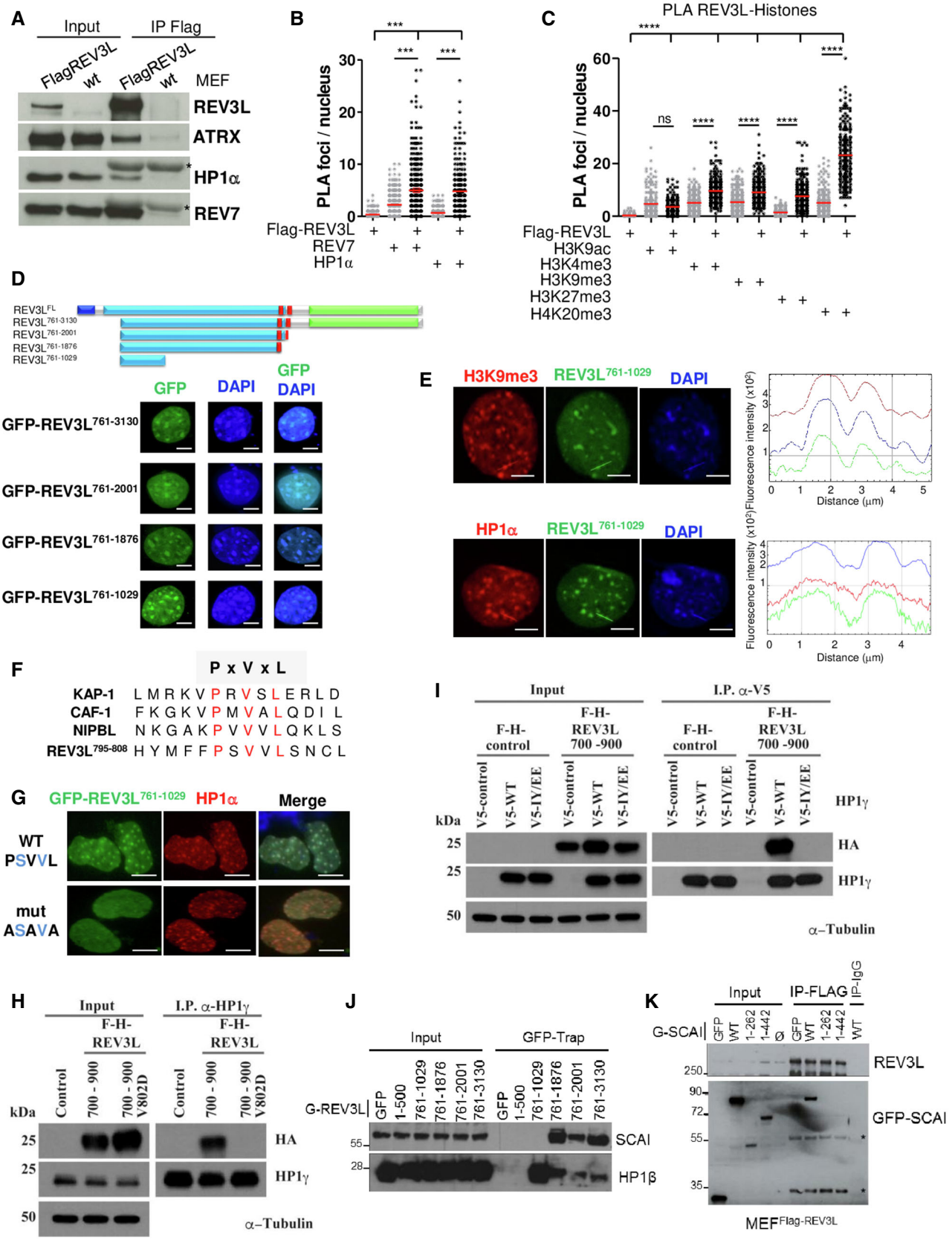


Figure 5.

Figure 5. REV3L localizes in heterochromatin through a direct interaction with HP1 dimer.

A Asynchronous MEFs expressing Flag-tagged REV3L from the endogenous locus were lysed, and REV3L was immunoprecipitated using anti-Flag (M2) antibodies. Co-immunoprecipitated proteins were analyzed by immunoblotting using the indicated antibodies. *: IgG light chain. This experiment was repeated 2–4 times.

B Asynchronous MEFs expressing Flag-tagged REV3L were subjected to *in situ* proximity ligation assay (PLA) to test the interactions REV3L-REV7 and REV3L-HP1α. Nuclear foci were quantified (more than 150 nuclei for each condition were counted). Reactions omitting one of the primary antibodies were used as negative controls. Horizontal bars show the mean. Mann–Whitney test, ns: not significant, ***$P < 0.001$. Experiments were repeated three times.

C Asynchronous MEFs expressing Flag-tagged REV3L were subjected to PLA to test the interactions REV3L-H4K20me3, REV3L-H3K9me3, REV3L-H3K27me3, REV3L-H3K4me3, and REV3L-H3K9ac. Nuclear foci were quantified as in (B). Horizontal bars show the mean. Mann–Whitney test, ns: not significant, ****$P < 0.0001$. Three independent experiments were performed.

D Schematic representation of human REV3L and truncated constructs. Conserved domains between yeast and human REV3L are in black blue (the N-terminal domain, 1–333 aa) and in green (the catalytic domain, 2,276–3,130 aa), the REV7 interacting domains are in red (1,880–2,001 aa). A large region in royal blue not found in yeast protein is almost exclusively encoded by exon 14. All the truncated constructs lack the NTD domain and were fused to GFP. MEFs cells were transiently transfected with various GFP-REV3L constructs and fixed with 4% formaldehyde. The distribution of the GFP-REV3L mutants was detected by autofluorescence, and nuclei were visualized using DAPI staining. Scale bar = 5 μm.

E MEFs cells were transiently transfected with GFP-REV3L$^{761–1,029}$ and fixed with 4% formaldehyde. The distribution of GFP- REV3L$^{761–1,029}$ was detected by autofluorescence (green), chromocenters were visualized by H3K9me3 immunostaining (red, top panel) or HP1α (red, bottom panel), and DNA was counterstained with DAPI. Line scans represent the colocalization of proteins within each image (right panels). Scale bar = 5 μm.

F Sequence alignment of proteins containing the PxVxL motif important for an interaction with HP1, with canonical residues shown in red.

G MEFs cells were transfected with GFP-REV3L$^{761–1,029}$ WT PSVVL or ASAVA mutant and fixed with 4% formaldehyde. The distribution of GFP-REV3L constructs was detected by autofluorescence (green), chromocenters were visualized by HP1α immunostaining (red), and DNA was counterstained with DAPI. Scale bar = 10 μm.

H Human 293T cells were transfected with either FH-REV3L$^{700–900}$, mutant V802D F-H-REV3L$^{700–900}$, or empty vector. Forty-eight hours after transfection, cell lysates were made and used for immunoprecipitation with HP1 antibody. Western blot was processed, and membranes were immunoblotted with the indicated antibodies.

I 293T cells were co-transfected with F-H-REV3L$^{700–900}$ or empty vector and V5-HP1γ, mutant IY/EE V5-HP1γ or empty vector. Forty-eight hours after transfection, cell lysates were made and used for immunoprecipitation with α-V5 antibody. After electrophoresis, samples were immunoblotted with anti-HA, anti-HP1γ, or anti-α-Tubulin as indicated.

J 293 cells were transfected with various GFP-REV3L constructs or empty vector (GFP). Twenty-four hours after transfection, cell lysates were made and GFP-REV3L was affinity-purified on GFP-Trap beads. After electrophoresis, samples were analyzed by immunoblotting with antibodies against SCAI or HP1β as indicated.

K MEF$^{Flag-REV3L}$ were transfected with various GFP-SCAI constructs, empty vector (GFP), or not transfected (Ø). Twenty-four hours after transfection, cell lysates were made and Flag-REV3L was affinity-purified on M2 beads. After electrophoresis, samples were analyzed by immunoblotting with antibodies against GFP or Flag (M2). *: non-specific bands.

Source data are available online for this figure.

followed this distribution (Fig EV4B). The localization to chromocenters was re-established rapidly after drug removal. Collectively, these results strongly suggest that HP1α is required for targeting REV3L to pericentromeric heterochromatin.

We therefore assessed whether REV3L directly interacts with HP1α. The chromoshadow domain of HP1α interacts with proteins containing a pentapeptide motif, PxVxL (Murzina *et al*, 1999; Thiru *et al*, 2004). Such a motif is present at position 800–804 in the 268-amino acid polypeptide (REV3L$^{761–1,029}$) that can target pericentromeric heterochromatin (Fig 5F). The PxVxL sequence and the putative NLS in this region are both completely conserved in vertebrate REV3L [(Fig EV4C) and Appendix Fig S1 in Lange *et al*, 2016)]. We tested whether this motif was critical for REV3L localization into chromocenters by introducing point mutations in the PSVVL motif. Localization of a transfected mutant GFP fusion protein (e.g., ASAVA) was strongly impaired (Fig 5G). Two alternative candidate PxVxL motifs are present near the C-terminus of REV3L. The expression of a REV3L construct containing these two motifs (GFP-REV3L$^{2,732–3,130}$) in mouse cells showed a diffuse staining in both nucleus and cytoplasm (Fig EV4D), indicating that these motifs alone cannot target REV3L at PHC. Taken together, these data reveal that the PxVxL motif at position 800–804 is important for REV3L localization at pericentromeric heterochromatin.

To confirm a direct interaction between REV3L and HP1, we performed a GST pull-down assay by using fusion polypeptides purified from *E. coli*. GST-XPA was used as negative control. While GST-REV3L$^{700–900}$ constructs bound to His-HP1α and His-HP1γ, the V802D mutant REV3L$^{700–900}$ did not (Fig EV4E), demonstrating that REV3L directly interacts with HP1 via its PxVxL motif at position 800–804. We confirmed these results by transfecting 293T cells with plasmids expressing F-H-REV3L$^{700–900}$, V802D F-H-REV3L$^{700–900}$, or empty vector. WT REV3L$^{700–900}$ but not the mutant co-immunoprecipitated using HP1γ antibody (Fig 5H).

Homodimerization of HP1 through its chromoshadow domain is critical for HP1 binding of H3K9me3 chromatin and chromatin condensation (Hiragami-Hamada *et al*, 2016). To gain insight into the molecular interaction between REV3L and HP1, we asked whether HP1 dimerization is mandatory for REV3L interaction with HP1. 293T cells were co-transfected with F-H-REV3L$^{700–900}$ or empty vector and V5-HP1γ, IY165-168EE V5-HP1γ mutant (unable to dimerize), or empty vector. We found that dimerization of HP1 is required for its interaction with REV3L (Fig 5I). We confirmed this result using F-H-tagged full-length REV3L (Fig EV4F). These results suggest that REV3L binds HP1 when HP1 forms condensed H3K9me3-modified chromatin.

We then evaluated at the endogenous level whether REV3L interacts with HP1 only at chromocenters, or also in heterochromatic regions dispersed in the genome. For that, we used images acquired for REV3L-HP1 PLA quantification (Figs 5B and EV3D) and determined the localization of each signal in the nucleus (Fig EV4G). Given that replication of pericentromeric regions occurs mainly at the surface of the chromocenters where PCNA is located (Quivy *et al*, 20042004), we distinguished foci "In/around" of chromocenters with foci "Out" of chromocenters. We used DAPI density to recognize these structures. We determined the localization of REV3L-HP1 PLA signals in 115 nuclei. We observed that 41/115 cells (36%) have ≥ 2 REV3L-HP1 PLA foci in/around chromocenters, suggesting that REV3L interacts with HP1 in pericentromeric

regions in one third of cells which probably undergo DNA replication. The mean of REV3L-HP1 PLA foci in and out of chromocenters/nucleus was 1.25 and 4.4, respectively (Fig EV4G), consistent with the idea that REV3L interacts with HP1 in pericentromeric heterochromatin, but also in other heterochromatic regions localized throughout the genome. To further explore specific interactions of REV3L with heterochromatin, we examined the HP1-binding protein SCAI (suppressor of cancer cell invasion; C9orf126) (Nozawa *et al*, 2010). SCAI functions in connection with 53BP1 to mediate ATM-dependent DSB signaling in heterochromatin (Hansen *et al*, 2016). A mass spectrometry-based approach identified REV3L as a high-confidence SCAI interactor (Appendix Fig S1 in Isobe *et al*, 2017). We validated the REV3L-SCAI interaction by co-immunoprecipitation with GFP-tagged REV3L constructs (Fig 5J). These experiments showed that SCAI-REV3L interaction is mediated by a region in REV3L between amino acids 1,030–1,876. HP1β was used as positive control for its interaction with residues 800–804 in REV3L. Further, full-length SCAI, but not $SCAI^{1–262}$ or $SCAI^{1–442}$, can be co-immunoprecipitated with endogenous REV3L (Fig 5K), indicating an interaction of REV3L with the C-terminal part of SCAI.

REV3L facilitates replication through pericentromeric heterochromatin, preventing DNA double-stranded breaks

These observations prompted us to investigate whether Polζ/REV3L may facilitate DNA replication at pericentromeric heterochromatin. DNA combing coupled with PNA probe hybridization was performed to specifically analyze pericentromeric fibers independently from bulk DNA fibers by using pericentromeric Satellite III (SatIII) repeats as probe (5′-TGGAA). We found that depletion of REV3L in HeLa cells resulted in shorter nascent DNA tracks and an increased fork ratio at SatIII regions, indicating that REV3L loss impairs fork progression at pericentromeric regions (Fig 6A). In contrast, REV3L depletion does not have a major effect on genome replication dynamics globally. However, a slight but significant increase in fork ratio was observed at globally after REV3L depletion, suggesting that Polζ/REV3L might assist replication of genomic regions in addition to pericentromeric heterochromatin.

Given that REV3L loss leads to genomic instability and chromosome breaks (Wittschieben *et al*, 2000), we asked whether DNA double-stranded breaks (DSBs) occur in heterochromatin regions. We first used PLA to examine whether DSB markers 53BP1 and γH2AX were spatially close to the heterochromatin-associated factor HP1 in $Rev3l^{-/-}$ cells. We observed that specific PLA signals between 53BP1 and HP1α (Fig 6B) and γH2AX and HP1β (Fig 6C) were higher in $Rev3l^{-/-}$ cells than in $Rev3l^{+/+}$ cells in unchallenged conditions. We confirmed these observations in human cells after depleting REV3L in RPE and HeLa cells (Fig 6D and E, respectively). These results suggest that DNA replication forks in heterochromatic regions might be prone to collapse in the absence of Polζ/REV3L.

Thus, we explored the consequences of REV3L depletion on PHC stability. DNA fluorescence *in situ* hybridization was performed on metaphase chromosome spreads using major satellite DNA as probe. For spontaneous breaks (in the absence of aphidicolin), a *Rev3l* defect elevates the frequency of breaks ~20-fold in pericentromeric regions (Fig 6F, upper panel), and about 10-fold in non-pericentromeric regions (Fig 6G). The consequence of *Rev3l* disruption is thus highly biased toward pericentromeric breaks, indicating an especially important function of Pol ζ in PHC. A variety of abnormalities were observed in PHC, including breaks/gaps, loss, duplication, and rearrangement (Fig 6F, lower panel). Exposure to aphidicolin (causing fork slowing) increases breaks in all regions. In PHC, the increase is *Rev3l*-dependent (Fig 6F), indicating that pol ζ is also particularly important in PHC for limiting the frequency of aphidicolin-stimulated breaks. In non-pericentromeric regions, *Rev3l* status did not affect the aphidicolin-stimulated frequency, showing that other mechanisms are effective globally in preventing the additional breaks initiated by aphidicolin.

Figure 6. REV3L facilitates replication in pericentromeric heterochromatin, preventing DNA double-stranded breaks.

A Representative fibers of newly synthesized DNA labeled with IdU (red) for 30 min and CldU (green) for 30 min in HeLa cells transfected with non-targeting siRNA (siNT) or siRNA against *Rev3l*. SatIII probe is visualized in blue. Scale bar: 10 μm = 20 kb. Relative *Rev3l* mRNA level normalized to GAPDH mRNA level is shown (top panel). Distribution of fork speeds (kb/min) and fork ratios (IdU/CldU track length) are shown in dot plots for SatIII and global DNA (bottom panel). The number of fibers analyzed is indicated in (n). Bars represent the median ± interquartile range (Mann–Whitney test. ns: not significant, *$P < 0.05$; **$P < 0.01$, ***$P < 0.001$ and ****$P < 0.0001$). The presented data are representative of three biological repeats.

B Asynchronous $Rev3l^{+/+}$ and $Rev3l^{-/-}$ MEFs were fixed with PFA then subjected to *in situ* proximity ligation assay (PLA) using 53BP1 and HP1α antibodies; then, PLA foci were counted in both cell lines (more than 150 nuclei for each condition were counted). *P*-values were calculated by Mann–Whitney test (****$P < 0.0001$). Red lines indicate the mean values. Error bars: SEM. Controls with a single antibody are also shown. Experiments were repeated three times. Representative images are shown. Scale bar = 5 μm.

C Asynchronous $Rev3l^{+/+}$ and $Rev3l^{-/-}$ MEFs were subjected to PLA as by using γH2AX and HP1β antibodies and processed as in (B). *P*-values were calculated by Mann–Whitney test (*$P < 0.05$). Red lines indicate the mean values. Error bars: SEM. Controls with a single antibody are also shown. Three independent experiments were performed.

D RPE cells were transfected with non-targeting siRNA (siNT) or siRNA against *Rev3l;* then, 72 h later cells were subjected to PLA by using 53BP1 and HP1α antibodies and processed as in (B). *P*-values were calculated by Mann–Whitney test (****$P < 0.0001$). Red lines indicate the mean values. Error bars: SEM. Three independent experiments were performed.

E HeLa cells were transfected with non-targeting siRNA (siNT) or siRNA against *Rev3l;* then, 72 h later cells were subjected to PLA by using 53BP1 and HP1α antibodies and processed as in (B). *P*-values were calculated by Mann–Whitney test (***$P < 0.001$). Red lines indicate the mean values. Error bars: SEM. Three independent experiments were performed.

F,G $Rev3l^{+/+}$ and $Rev3l^{-/-}$ MEFs were incubated with or without 0.23 μM aphidicolin for 24 h before metaphase spreading. FISH was performed using major satellite probe to quantify breaks in pericentromeric regions (F). Representative images of chromosomes showing abnormalities (see arrows) in pericentromeric regions from $Rev3l^{-/-}$ MEFs. Chromosomes were labeled with DAPI, and breaks in non-pericentromeric regions were quantified (G). Error bars indicate standard error of the mean for three independent experiments. Mann–Whitney test (ns: not significant, *$P < 0.05$; ***$P < 0.001$).

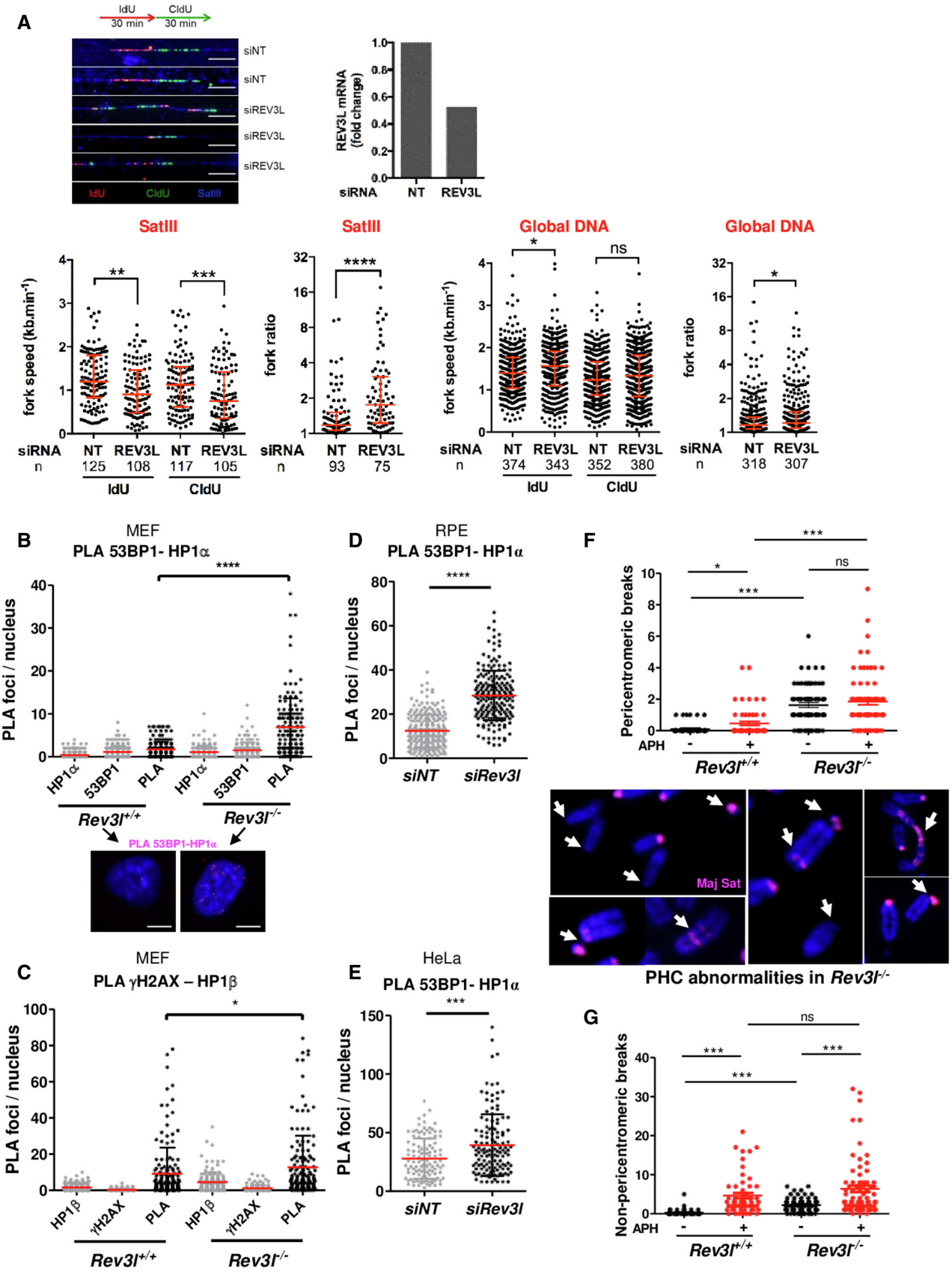

A
IdU 30 min
CldU 30 min
siNT
siNT
siREV3L
siREV3L
siREV3L
IdU
CldU
SatIII
REV3L mRNA (fold change)
siRNA NT REV3L
SatIII
SatIII
Global DNA
Global DNA
fork speed (kb.min-1)
fork ratio
siRNA NT REV3L NT REV3L
n 125 108 117 105
IdU
CldU
siRNA NT REV3L
n 93 75
siRNA NT REV3L NT REV3L
n 374 343 352 380
siRNA NT REV3L
n 318 307
B
MEF
PLA 53BP1- HP1α
PLA foci / nucleus
HP1α 53BP1 PLA HP1α 53BP1 PLA
Rev3l+/+
Rev3l-/-
PLA 53BP1-HP1α
C
MEF
PLA γH2AX – HP1β
PLA foci / nucleus
HP1β γH2AX PLA HP1β γH2AX PLA
Rev3l+/+
Rev3l-/-
D
RPE
PLA 53BP1- HP1α
PLA foci / nucleus
siNT
siRev3l
E
HeLa
PLA 53BP1- HP1α
PLA foci / nucleus
siNT
siRev3l
F
Pericentromeric breaks
APH
Rev3l+/+
Rev3l-/-
Maj Sat
PHC abnormalities in Rev3l-/-
G
Non-pericentromeric breaks
APH
Rev3l+/+
Rev3l-/-

Figure 6.

Rev3l-deficient cells show increased number of genomic deletions

To gain deeper insight into the role of Polζ/REV3L in preventing DSBs and genomic instability at the genome-wide level in mammalian cells, we characterized structural variations that arise upon loss of *Rev3l*. Whole-genome sequencing from subclones of *Rev3l*-deficient MEFs was analyzed (Fig 7A). We identified in total 102 structural variants, including 66 deletions, 13 inversions, 10 duplications, and 13 complex events from 4 samples (Fig 7B), using both GRIDSS2 and Manta algorithms (see Materials and Methods). The mean deletion size was 2,092 bp and ranged from 55 to 9,115 bp. Importantly, we found 61 deletions in $Rev3l^{-/-}$ samples and only 5 in $Rev3l^{+/+}$ samples (95%CI = 7.6–13.4, *P*-value < 2.2e-16, Poisson test) and the number of deletions was consistent between replicates (Fig 7C). Note that none deletion was shared between samples, suggesting that they occurred during independent cell cultures. Moreover, the size of the deletions that accumulate in $Rev3l^{-/-}$ cells was within a rather narrow range, with 21 deletions (34%) smaller than 1,000 bp.

We then correlated deletions with replication timing observed in $Rev3l^{-/-}$ cells and found that deletions were preferentially localized in late-replicating regions as well as in domains for which replication timing was disturbed in the absence of REV3L (Fig 7D). Given that heterochromatic regions (facultative and constitutive) replicate in mid- and late S-phase, we assume that deletions found in $Rev3l^{-/-}$ mouse cells might be initiated by replication-associated DSBs that arise when replication forks stall and collapse in difficult-to-replicate heterochromatic regions.

Loss of REV3L compromises heterochromatin-associated DSB repair

The results above imply that spontaneous DSBs originate in *Rev3l*-deficient cells, at least partly, as a consequence of inefficient replication fork progression through heterochromatin. However, it has been proposed that Polζ/REV3L can also facilitate DSB repair by homologous recombination (Sharma *et al*, 2012). Therefore, the high level of DSBs in the absence of Polζ could be due to a combination of increased replication-associated DSBs and inefficient repair of those breaks.

To investigate the effect of Polζ/REV3L in regulating repair of heterochromatin-associated DSBs, we challenged MEFs with UV radiation. This should lead to enhanced replication-associated DSBs in $Rev3l^{-/-}$ cells (Sonoda *et al*, 2003). DSB signaling and repair were monitored in several ways. Phosphorylation of the Kruppel-associated box (KRAB)-associated co-repressor, KAP1 by ATM in response to DSBs, occurs during DNA repair in heterochromatin (Goodarzi *et al*, 2008). There was a marked increase in S824-phosphorylated KAP1 from 2 J/m^2 in $Rev3l^{-/-}$ MEFs (Fig 8A and B) which remained elevated 24 h after UV irradiation. In contrast, induction of pKAP1 required higher doses in $Rev3l^{+/+}$ MEFs and $Rev3l^{-/-}$ cells complemented with REV3L, and largely disappeared after 24 h. This suggests that heterochromatin-associated DSBs accumulate in UV-irradiated cells when Polζ is missing. Moreover, $Rev3l^{-/-}$ cells had increased activation of phosphorylated (S4-S8) RPA2 (an indicator of resection at DSBs), and a persistent level of pCHK1 (Fig 8A and B). Similar results were obtained in human cells after down-regulating REV3L in HeLa and RPE cells (Fig EV5).

We further examined the dynamics of DSB formation and repair by single-cell immunofluorescence. Following irradiation with 4 J/m^2, the levels of 53BP1 foci in EdU$^+$ cells slightly increased 6 h after UV exposure in *Rev3l*-proficient cells and returned to a basal level 24 h later (Fig 8C and F). In $Rev3l^{-/-}$ cells in S-phase, 53BP1 foci accumulated at 5.6-fold higher levels 24 h post-UV.

To assess a potential defect in *Rev3l*-deficient cells to complete DNA repair by homologous recombination (HR), we quantified the level of resection indirectly by measuring levels of chromatin-bound RPA2 in individual cell nuclei 24 h after UV irradiation. In agreement with pRPA2 immunoblots, we observed that chromatin-bound RPA2 was 3.6-fold higher in cells lacking REV3L than in wild-type cells (Fig 8D and G). We also evaluated RAD51 focus formation, an indicator of HR initiation. Twenty-four hours after UV exposure, RAD51 foci were elevated 3-fold in *Rev3l*-deficient MEFs compared with wild-type cells (Fig 8E and H). This suggests that Polζ/REV3L may be required for efficient completion of DSB repair by HR.

Discussion

Polζ contributes to heterochromatin replication and repair

This study unveils a direct role for mammalian Polζ/REV3L in replication and repair in heterochromatin. Several lines of evidence point toward this conclusion, summarized in Fig 8I. First, inactivation of *Rev3l* drives an elevated number of breaks and structural variations in heterochromatic regions, particularly deletions that likely arise as a consequence of replication fork stalling or collapse occurring during DNA replication. Second, Polζ/REV3L influences replication fork progression in pericentromeric heterochromatin. REV3L-compromised cells exhibited a marked decrease in fork speed at pericentromeres which contain highly repeated DNA sequences embedded into compacted chromatin, making them difficult to replicate. A disruption of replication timing at specific loci mainly located in TTR may also be ascribed to heterochromatin replication and repair-associated functions of Polζ. Third, we show that Polζ directly interacts with specific components of heterochromatin including HP1 and SCAI. Fourth, we show that REV3L disruption is correlated with changes in epigenetic landscape and transcriptional control of developmentally regulated genes.

Polζ/REV3L interacts with heterochromatin proteins

We found that REV3L directly interacts, via a central PxVxL motif, with the heterochromatin component HP1. Dimerization of HP1 is necessary for interaction with REV3L, indicating that such binding occurs only when HP1 is engaged on chromatin through its reader interaction with H3K9me3 (Hiragami-Hamada *et al*, 2016). The PxVxL motif located at positions 800–804 in the human REV3L sequence is embedded in a larger region that is well-conserved in vertebrates but absent in invertebrate and fungal genomes. This indicates that the heterochromatin targeting of Polζ described in this study evolved in concert with larger genomes having more complex controls. The REV7 subunit also is associated with HP1α (Tomida *et al*, 2018), suggesting another mode of interaction of Polζ with heterochromatin. It is not yet known whether Polζ associates with the replisome to duplicate heterochromatin regions. The replisome

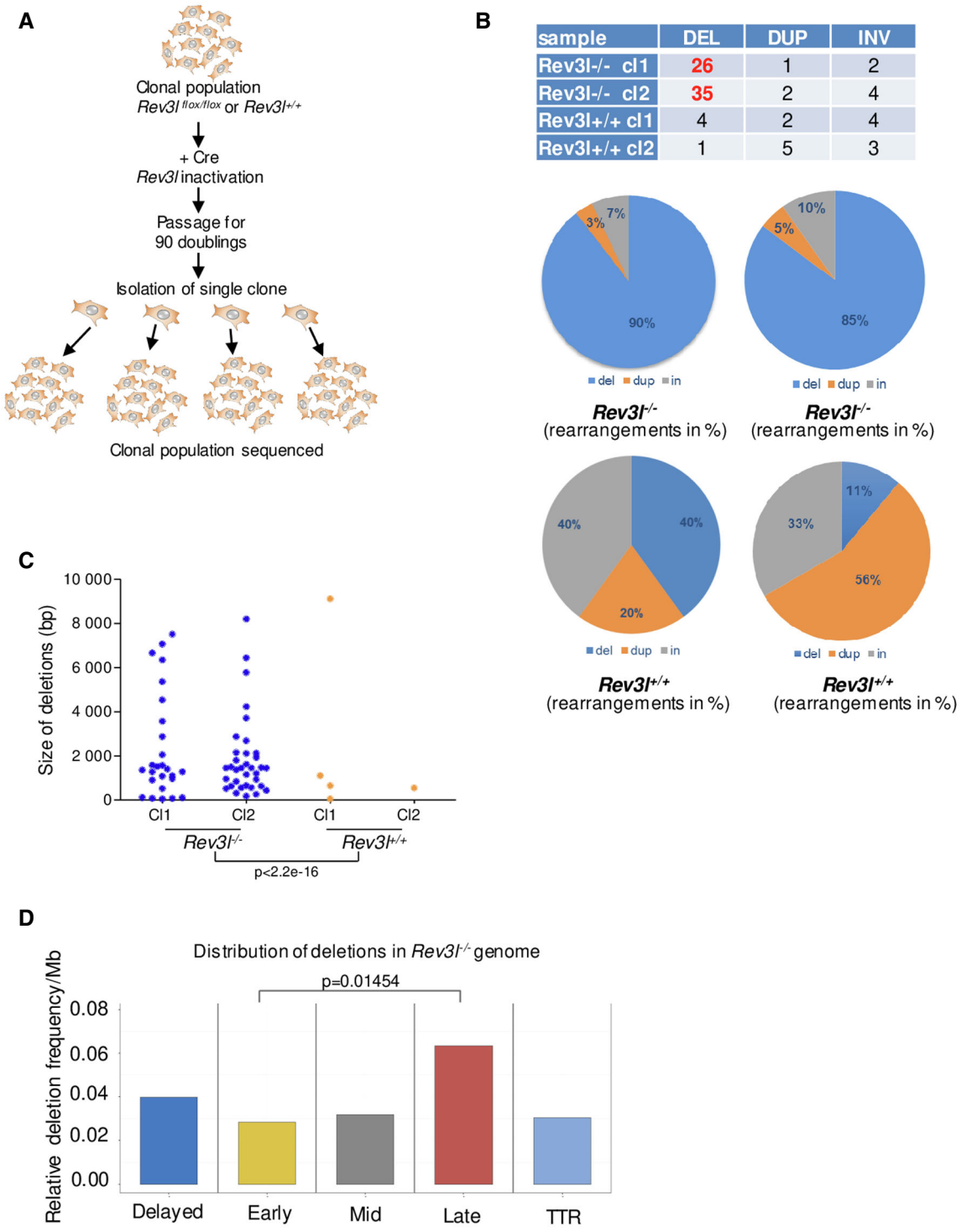


Figure 7. *Rev3l*-deficient cells show numerous genomic deletions.

A Experimental design for the whole genome sequencing of *Rev3l*-deficient cells. Populations of mouse embryonic fibroblasts $Rev3l^{+/+}$ and $Rev3l^{-/-}$ were isolated from $Rev3l^{+/+}$ and $Rev3l^{flox/flox}$ littermates, and *Rev3l* was excised using CRE recombinase. The resulting *Rev3l*-deficient cells and $Rev3l^{+/+}$ cells were expanded for 90 doublings. $Rev3l^{+/+}$ and $Rev3l^{-/-}$ subclones were then isolated for deep sequencing.

B Percentage of rearrangements (deletions, duplications, and inversions) identified in $Rev3l^{+/+}$ and $Rev3l^{-/-}$ subclones.

C Size distribution of the deletions detected in $Rev3l^{+/+}$ and $Rev3l^{-/-}$ subclones. Poisson test. Deletions were counted from four independent clones.

D Distribution of the deletions detected in $Rev3l^{-/-}$ subclones related to the replication timing in deletions per megabase. Difference between late and early regions is statistically significant ($P = 0.01454$, Poisson test with the correction for different length of intervals through the ratio). Data were calculated from two independent clones.

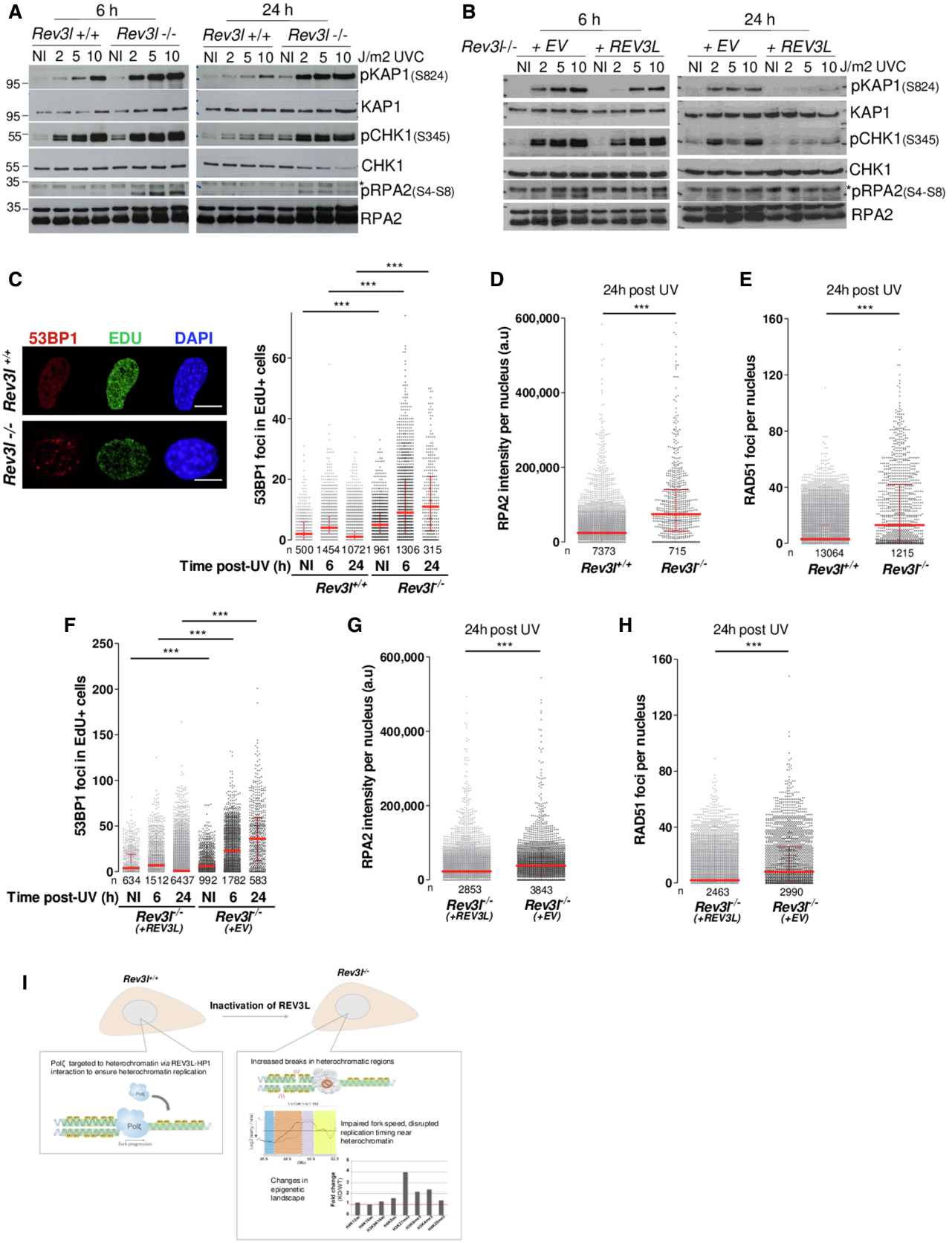


Figure 8.

Figure 8. REV3L loss compromises heterochromatin-associated DSB repair.

A Asynchronous *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs were UV-irradiated at the indicated doses and harvested 6 or 24 h later. Cell lysates were analyzed by immunoblotting with indicated antibodies.

B Asynchronous *Rev3l*$^{-/-}$ MEFs complemented with empty vector (EV) or REV3L were UV-irradiated at the indicated doses and harvested 6 or 24 h later. Cell lysates were analyzed by immunoblotting with indicated antibodies.

C–E Asynchronous *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ MEFs were pulse-labeled for 15 min using 10 μM EdU then UV-irradiated at 4 J/m^2. Cells pre-treated with CSK buffer were fixed, and 53BP1, RPA, and RAD51 were detected by immunofluorescence. Representative images in non-irradiated cells (NI) and quantification of 53BP1 foci in EdU-positive cells at different time points are shown (C). Scale bar = 10 μm. Quantification of the intensity of chromatin-bound RPA2 (D) and RAD51 foci (E) 24 h after UV irradiation was performed. The number of cells analyzed is indicated in (n) on each graph. ***$P < 0.001$ (Kruskal–Wallis test). The presented data are representative of 2 repeats. Bars represent the median ± interquartile range.

F–H Asynchronous *Rev3l*$^{-/-}$ MEFs complemented with empty vector (EV) or REV3L were processed as in (C). Quantification of 53BP1 foci in EdU-positive cells at different time points is shown (F). Quantification of the intensity of chromatin-bound RPA2 (G) and RAD51 foci (H) 24 h after UV irradiation was performed. The number of cells analyzed is indicated in (n) on each graph. ***$P < 0.001$ (Kruskal–Wallis test). The presented data are representative of two repeats. Bars represent the median ± interquartile range.

I Model that summarizes the results obtained in this study.

Source data are available online for this figure.

may change throughout the S-phase to assist conventional DNA polymerases to duplicate challenging areas of the genome such as compacted heterochromatin. Potentially, targeting of REV3L to heterochromatin could facilitate a proposed catalytic subunit switch with Polδ, mediated by PCNA and the shared POLD2 and POLD3 subunits. The interaction of the DONSON protein and the FANCM translocase with the replisome complex is dependent on replication timing and chromatin environment (Zhang *et al*, 2020). Further, progression of replication forks through pericentromeric heterochromatin is also facilitated by the shelterin subunit TRF2, in association with the helicase RTEL1 (Mendez-Bermudez *et al*, 2018). Investigation is warranted to define whether REV3L/Polζ works together with TRF2 and RTEL1 to facilitate fork progression in heterochromatic regions.

Polζ limits heterochromatin-associated DSBs

Our data also show that *Rev3l*-deficient MEFs exhibit an increased incidence of deletions, preferentially in late-replicating regions. This is consistent with the observed ~10-fold increase in the frequency of DNA double-stranded breaks and chromosome rearrangements in such MEFs (Lange *et al*, 2013, 2016), similar to the spectrum of changes in HR-deficient MEFs (Wittschieben *et al*, 2006). The present study shows that many spontaneous DSBs originate during replication of pericentromeric heterochromatic regions. We also provide evidence that *Rev3l* inactivation compromises repair of heterochromatin-associated DSBs. Low doses of UV radiation provoke replication-associated DSBs, leading to a massive increase of sustained 53BP1 foci in S-phase, but also markers of HR including pRPA2 and RAD51 foci. One hypothesis to explain the high level of these breaks can be the delay in repair caused by saturation of the repair machinery in *Rev3l* KO cells. Alternatively, Polζ/REV3L might be required for efficient HR-associated DNA synthesis to complete DSB repair (Sharma *et al*, 2012). There is evidence that Polζ-mediated DNA synthesis occurs during DSB repair in some settings in different organisms (Kane *et al*, 2012) (Martin & Wood, 2019). The contribution of specific DNA polymerases in HR-associated DNA synthesis may be subject to temporal control of replication and to chromatin composition. Indeed, our results showed that *Rev3l*-deficient cells fail to efficiently resolve DSBs in heterochromatin, with a pronounced persistence of S824-phosphorylated KAP1, known to facilitate chromatin decondensation and allows efficient DSB repair in heterochromatin (Goodarzi *et al*, 2008). Consistent with this, we demonstrated that REV3L interacts with the SCAI protein, a mediator of 53BP1-dependent repair of heterochromatin-associated DSBs (Hansen *et al*, 2016). By interacting with 53BP1, SCAI counteracts RIF1 function, facilitating BRCA1-mediated repair (Isobe *et al*, 2017). Future studies may determine the mechanisms by which Polζ/REV3L is involved in the repair of heterochromatin-associated DSBs by interaction with SCAI.

Absence of Polζ disrupts replication timing at specific loci

We found that REV3L loss perturbs the temporal replication program in mouse and human cells, which might be a consequence of fork progression alterations. The replication-timing defect affects only specific areas, mainly in temporal transition regions, corresponding to 5.7% of the whole genome in *Rev3l*-deficient mouse cells. This is a contrast to the change in replication patterns following inactivation of RIF1, with a shift in replication timing in over 40% of all replication segments (Cornacchia *et al*, 2012; Hayano *et al*, 2012). Our data indicate an involvement of REV3L/Polζ in the replication of particular TTRs that lie between early-replicating DNA and late-replicating DNA and, therefore, often replicate in mid-to-late S-phase. Consistent with this, pericentromeric heterochromatin is replicated in mid-to-late S-phase in mouse cells (Guenatri *et al*, 2004; Natale *et al*, 2018). Replication in TTR is achieved by either sequential activation of a series of origins (Guilbaud *et al*, 2011) or sequential activation of single long unidirectional fork that initiates at an adjacent early origin (Norio *et al*, 2005; Hiratani *et al*, 2008; Schultz *et al*, 2010). It has been reported that the replication machinery propagates at a slower velocity in some TTRs (Farkash-Amar *et al*, 2008; Desprat *et al*, 2009; Donley & Thayer, 2013), suggesting that one major consequence is an increase in the probability of replication fork stalling and breaks as has been proposed for fragile sites (Watanabe & Maekawa, 2010; Letessier *et al*, 2011; Ozeri-Galai *et al*, 2014). Consistent with this, *Rev3l*-deficient MEFs exhibit genomic deletions in perturbed TTR.

Potential contribution of REV3L/Polζ to the mutation rate in late-replicating heterochromatic regions

A growing body of evidence suggests that the replication-timing program strongly influences the spatial distribution of mutagenic

events during both species and cancer evolution with an increasing gradient of single-nucleotide substitutions that correlate with late-replicating heterochromatic regions (Stamatoyannopoulos *et al*, 2009; Koren *et al*, 2012; Schuster-Bockler & Lehner, 2012). Given the error-prone activity of Polζ, our results suggest that REV3L/Polζ may contribute to this increased mutation rate in heterochromatic regions which are relatively poor in transcribed genes. This is consistent with a bioinformatic analysis showing that mutations related to Polζ signature increase in late-replicating regions of the human genome (Seplyarskiy *et al*, 2015). Moreover, experiments in yeast established that late-replicating regions of the genome also have higher rates of spontaneous mutagenesis than early-replicating regions. Deletion of yRev1 significantly lowers mutation rate specifically in these late-replicating regions, suggesting that Rev1/Polζ complex is involved in the gradient of mutagenesis in lower eukaryotes (Lang & Murray, 2011).

REV3L loss impacts the epigenetic landscape and transcriptional program

The majority of disturbed regions in *Rev3l*-deficient mouse cells (> 80%) replicate later than in wild-type cells, suggesting heterochromatinization of these regions that shifts their replication timing. We further found that the inactivation of *Rev3l* affects epigenetic landscape and transcriptional program in mouse cells, with a substantial increase in H3K27me3 levels and, to a lesser extent, H3K9me3 and H3K4me3. This may be linked to DSB repair processes. Heterochromatin factors including KAP1, HP1, and the H3K9 methyltransferase Suv39h1 accumulate rapidly and transiently at DNA damage sites, in both euchromatic and heterochromatic regions (Lemaitre & Soutoglou, 2014; Nikolov & Taddei, 2016). Moreover, polycomb proteins involved in gene silencing at facultative heterochromatin are also recruited to DSB sites, where they are thought to switch off transcription to facilitate DSB repair (Vissers *et al*, 2012; Ui *et al*, 2015). Thus, continued DSB formation in *Rev3l*-deficient cells might result in a "heterochromatinization" at particular genomic regions. Interestingly, the TLS polymerase REV1, a key partner of Polζ, also influences the epigenetic landscape. During DNA replication, REV1 is required for the maintenance of repressive chromatin marks and gene silencing in the vicinity of G4 structure in DT40 chicken cells (Sarkies *et al*, 2010). Moreover, REV7 has been shown to function in epigenetic reprogramming by interacting with the G9a/G9a-like protein (GLP) histone lysine methyltransferase complex which catalyzes H3K9 mono- and dimethylation (Pirouz *et al*, 2013).

Our transcriptome analysis revealed that *Rev3l*-deficient cells repress numerous genes, known to contain bivalent promoters such as multiple *Hox* genes which are regulated by H3K27me3/H3K4me3 (Sachs *et al*, 2013). Whether REV3L/Polζ contributes to H3K27me3 regulation during DNA replication is largely unexplored. Of note, the large majority of the genes whose expression is down-regulated by *Rev3l* deletion do not fall within disturbed regions of replication timing. One possibility is that gene repression is due to the heterochromatinization of the loci that undergo continued DSBs repair. This implies that DSBs induced by *Rev3l* inactivation occur not only in heterochromatin, but also in euchromatin. We also found that several imprinted genes are down-regulated in the absence of *REV3L*. Rajewsky and colleagues suggested that the overall phenotype of $Rev3l^{-/-}$ embryos appears comparable to androgenetic embryos (Fundele & Surani, 1994), raising the possibility that Polζ participates in imprinting establishment (Esposito *et al*, 2000). Thus, our results on the transcriptional program in *Rev3l*-deficient mouse cells may partially explain the embryonic lethality observed in *Rev3l* KO mice.

Conclusion

Unique among translesion synthesis DNA polymerases, Polζ is essential during embryogenesis and cell proliferation (Wittschieben *et al*, 2000; Lange *et al*, 2012, 2018). However, the mechanisms of essentiality under normal growth conditions have not been clear. This study shows that Polζ prevents spontaneous chromosome break formation, rearrangements, and deletions in pericentromeric regions by facilitating fork progression, and may also operate in heterochromatin-associated DSB repair. Altogether, these results reveal a new function of Polζ in preventing chromosome instability during replication of heterochromatic regions possibly at the expense of increased mutations in late-replicating regions found in numerous tumors.

Materials and Methods

Cell culture and treatment

Immortalized mouse embryonic fibroblast (MEF) $Rev3l^{-/-}$ and $Rev3l^{+/+}$ were previously described (Lange *et al*, 2012) and were cultivated in D-MEM (Dulbecco's modified Eagle medium; Gibco) containing 100 U/ml penicillin, 100 μg/ml streptomycin, and 10% fetal calf serum (FCS) in an atmosphere containing 5% CO_2. MEFs expressing functional REV3L with a 3X-Flag epitope tag at the endogenous locus were derived from knock-in mice as described below. $Rev3l^{-/-}$ MEFs complemented with POZ empty vector (clone 4–5 POZN Cl2) or POZ-hREV3L (clone 4–5 POZRev3l2 Cl H11) have been previously described (Lange *et al*, 2016). Suv39h double-null MEFs kindly provided by T. Jenuwein and NIH3T3 cells (ATCC) were cultivated in D-MEM supplemented with 10% FCS, 100 U/ml penicillin, and 100 μg/ml streptomycin. hTERT RPE-1 cells (ATCC) were grown in D-MEM/F-12 (Gibco) supplemented with 10% FCS, 100 U/ml penicillin, and 100 μg/ml streptomycin. HeLa cells (ATCC) were cultivated in D-MEM containing sodium pyruvate, penicillin/streptomycin, and 10% FCS. All cell lines were incubated at 37°C in a 5% CO_2 atmosphere and were regularly tested for mycoplasma.

For UVC irradiation (254 nm), cells were rinsed in pre-heated phosphate buffer saline (PBS) and irradiated without any medium at a fluency of 0.65 $J/m^2/s$. Aphidicolin (Sigma) stock solutions was at 3 mM in DMSO, and trichostatin A (Sigma) stock solution was at 2 mg/ml in methanol.

Generation of knock-in mice expressing 3x-FLAG tag REV3L

Ethics statement

All animal work in this study was done according to The University of Texas, MD Anderson Cancer Center Institutional Animal Care and Use Committee guidelines, and approved by the MD Anderson Animal Care and Use Committee (IACUC).

Construction of the targeting vector

The targeting vector construction was designed and performed by genOway (Lyon, France). The *Rev3l* targeting vector was constructed from 129 Sv/Pas mouse strain genomic DNA with a long (5.8 kb) homology arm upstream of exon 1, and a short (1.5 kb) homology arm downstream of exon 1. A 3x-FLAG peptide sequence was inserted in frame with the ATG codon. A positive selection neomycin gene was flanked by loxP sites. The targeting vector also incorporated a diphtheria toxin-negative selection cassette.

Screening of Rev3l targeted knock-in ES cell clones

Linearized targeting vector was transfected into 129 Sv ES cells (genOway, Lyon, France) according to genOway's electroporation procedures (i.e., 5×10^6 ES cells with 40 μg of linearized plasmid, 260 V, 500 μF). Positive selection started 48 h after electroporation in medium containing 200 μg/ml of G418 (150 μg/ml of active component, Life Technologies, Inc.). *Rev3l*-resistant clones were isolated and amplified in 96-well plates. Duplicates of 96-well plates were made. The set of plates containing ES cell clones amplified on gelatin were genotyped by both PCR and *Southern blot analysis*.

For PCR analysis, one primer pair was designed to validate the presence of the 3xFLAG tag within the 5′ homology arm. This primer pair was designed to specifically amplify the targeted locus:

Forward (Neo cassette): 5′-CCTGCTCTTTACTGAAGGCTCTTTAC TATTGC-3′
Reverse: 5′-GGAACCCACAGTGGTTGTCCTAGTGC-3′

PCR products were sequenced in order to validate the presence of all elements. The targeted locus was confirmed by Southern blotting using a 3′ probe (Fig EV4A). Six clones were identified as correctly targeted and containing the 3x-FLAG tag at the *Rev3l* locus.

Generation of mosaic mice and breeding scheme

Clones were microinjected into C57BL/6J blastocysts and gave rise to male mosaics with a significant ES cell contribution (as determined by agouti coat color). Mice were bred to C57BL/6 mice expressing the Cre recombinase to remove the Neo cassette.

Genotyping of the knock-in mouse line

The following primers were used to monitor the Cre-mediated excision event and served for genotyping.

Forward: 5′-ACGAGTTCGCGGGTCCTTAGAGGTC-3′
Reverse: 5′-ACTTTCTACAGCCACAGCATCTCCGG-3′

The wild-type allele generates a product of 195 bp, the recombined allele 1,958 bp, and the Cre-excised allele 289 bp.

Heterozygous mutant mice (which were maintained as pure C57BL/6 mice) were used to produce MEFs from embryos isolated at days 13.5 dpc. Seven heterozygous clones were obtained and T-antigen immortalized as described (Lange *et al*, 2012).

Plasmid construction

The human REV3L full-length cDNA was kindly provided by C. Lawrence (University of Rochester Medical Center, Rochester, NY, USA). To generate the series of GFP-*Rev3l* plasmids, constructs were obtained by polymerase chain reaction (PCR) amplification from *Rev3l* cDNA as a XhoI–EcoRI fragment and cloned into peGFP-C3 expression vector (Clontech). Point mutants were generated by PCR-based methods using a QuikChange® site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions.

Full-length REV3L with an N-terminal FLAG-HA dual-epitope tag was constructed in the pOZN vector as described previously (Tomida *et al*, 2015).

The cDNA encoding REV3L amino acids 700–900 was PCR amplified from REV3L/pETDuet-1 (Tomida *et al*, 2015) as a XhoI–NotI fragment with 5′ REV3L (XhoI) primer and 3′ REV3L (NotI) primers (5′-CC GCTCGAGATGAATACATTGGGCAAAAATTCTTTC and 5′-TAAAAGCG GCCGCTTATCCAAAGTGACAGTCTATAAAAC) to clone into pGEX6P-1 (GE Healthcare). The V802D mutation was introduced into pGEX6P-1 using a site-directed mutagenesis kit. The GST fusion protein was expressed and purified from *E. coli* as described (Tomida *et al*, 2015).

The HP1α (CBX5) and HP1γ (CBX3) cDNAs were obtained from Thermo Fisher Scientific. The cDNAs were cloned into vectors pETDuet-1 and pCDH-EF1α-V5-MCS-IRES-Puro for His and V5-tagging, respectively. pCDH-EF1α-V5-MCS-IRES-Puro was cloned into pCDH-EF1α-V5-REV7-IRES-Puro from pRSFDuet-1 (Novagen) XhoI–NotI fragment. pCDH-EF1α-V5-REV7-IRES-Puro was initially PCR amplified from pOZN REV7 (Tomida *et al*, 2018) with 5′ V5-1st primer and 3′ pOZ 3′ primers (5′-ATTCTACGCTCGAGATGAC CACGCTCACACGACAAG and 5′-CGGAATTGATCCGCTAGAG). Three subsequent rounds of PCR used the following primer sets: 5′ V5 2nd primer and 3′ pOZ 3′ primers (5′-CTCTCCTCGGTCTCGATTC TACGCTCGAGATGACCACGC and 5′-CGGAATTGATCCGCTAGAG); 5′ V5 3rd primer and 3′ pOZ 3′ primers (5′-AGCCTATCCC TAACCCTCTCCTCGGTCTCGATTCTAC and 5′-CGGAATTGATCCGC TAGAG); 5′ V5 4th primer (EcoRI) and 3′ pOZ 3′ primers (5′-GG AATTCATGGGTAAGCCTATCCCTAACCCTCTCCTCG and 5′-CGGA ATTGATCCGCTAGAG). The resulting V5-REV7 fragment (XhoI–NotI) was cloned into pCDH-EF1α-MCS-IRES-Puro (System Biosciences). Mutations were introduced using a site-directed mutagenesis kit (Stratagene). His-XPA was as described (Manandhar *et al*, 2017). All constructs were verified by DNA sequencing. Details on all constructs are available upon request.

The cDNA encoding SCAI was cloned from HeLa mRNA by RT–PCR using the following primers: 5′-ATGCATCTCGAGGTCAGAG GAGCCCGGCAGCCCCAGCAGC-3′ and 5′-ATGCATGAATTCTTAA TAGTCATCAATGGTATTCTCAAAGAA-3′. The resulting SCAI construct (XhoI–EcoRI) was verified by Sanger sequencing then cloned into peGFP-C3 (Clontech).

Transfection

For transient transfection, plasmids were transfected using jetPEI (Polyplus), according to the manufacturer's instructions, and cells were processed 24–48 h later.

siRNAs were purchased from Dharmacon: siRNA *mRev3l* (ON-TARGETplus Smart pool), non-targeting siRNA (ON-TARGETplus non target pool). Cells were transfected with 30 nM of siRNAs using INTERFERin (Polyplus) according to the manufacturer's instructions and were processed 72 h later.

Protein extraction, Immunoprecipitation, and Western blotting

Histone extraction was carried out in 10 volumes of CSK buffer (200 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 10 mM Tris–HCl pH 7, 1 mM EGTA, 0.5% Triton X-100, 1 mM PMSF, 10 mM Na-Butyrate) during 5 min at 37°C. After centrifugation, the pellet was resuspended in 0.4 mM H_2SO_4 supplemented with Na-Butyrate and PMSF for 1 h, then centrifuged before adding acetone to supernatant, and incubated overnight in cold room. After centrifugation, the pellet was resuspended in distilled water and histone extracts were kept at −80°C.

For immunoprecipitation, cells were lysed in NETN buffer (50 mM Tris–HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.5% NP-40, protease inhibitor cocktail and anti-phosphatases) for 30 min on ice and sonicated twice at 29% amplitude for 10 s. Samples were cleared by centrifugation for 5 min at 4°C. Immunoprecipitations were performed with Anti-FLAG M2 Agarose Beads (Sigma-Aldrich) overnight at 4°C on a wheel. Beads were extensively washed in NETN, with 300 mM NaCl for the final wash, and denatured in 2× Laemmli buffer.

For GFP-Trap, cells transfected with plasmids expressing GFP, GFP-REV3L mutants, GFP-SCAIwt, or GFP-SCAI mutants were collected in NETN buffer, incubated for 30 min on ice, sonicated twice at 29% Amplitude for 10 s then centrifuged at 4°C. The supernatant was incubated with GFP-Trap beads (Chromotek) for 2 h at 4°C under rotation. Beads were washed and eluted in 2× Laemmli buffer.

Proteins were separated on 5, 8, or 15% acrylamide SDS–PAGE, transferred on Nylon membrane (GE Healthcare), and detected with the indicated antibodies and ECL reagents (Thermo Fisher Scientific).

GST pull-down assay

GST-tagged REV3L fragments and His-tagged test proteins were incubated together with Glutathione-Sepharose 4B (GS4B) beads (GE Healthcare) at 4°C for 4 h in 500 µl of buffer 0.1B (100 mM KCl, 20 mM Tris–HCl [pH 8.0], 5 mM $MgCl_2$, 10% glycerol, 1 mM PMSF, 0.1% Tween-20, 10 mM 2-mercaptoethanol). The beads were washed three times with 0.1B and eluted with 30 µl of 2× SDS loading buffer (100 mM Tris–HCl [pH 6.8], 4% SDS, 0.2% bromophenol blue, 20% glycerol, 200 mM DTT).

Flow cytometry

In order to study the S-phase progression in *Rev3L*$^{-/-}$ and *Rev3L*$^{+/+}$ fibroblasts, cells were incubated with 50 µM bromodeoxyuridine (BrdU) for 15 min prior to be harvested at different time points. Cells were collected, washed in PBS, and fixed in 80% ice-cold ethanol overnight at −20°C. Total DNA was stained in 2.5 µg propidium iodide (IP) and 50 µg/µl RNAse (Sigma) in PBS for 30 min at room temperature. Samples were analyzed on a C6 flow cytometer using the C6 Flow software (BD Accuri).

Immunofluorescence

For the visualization of DNA replication sites, asynchronous cells seeded on glass coverslips were incubated with 30 µM of thymidine analog (EdU or BrdU) for 15 min, pre-extracted with CSK100 buffer (100 mM NaCl, 300 mM sucrose, 10 mM PIPES [pH 6.8], 3 mM $MgCl_2$, 1 mM EGTA, 0.5% Triton X-100, protease inhibitors) for 5 min at 4°C and fixed with 4% paraformaldehyde. For BrdU detection, cells were treated with 2 M HCl at 37°C for 20 min and then neutralized in 0.1 M borate buffer (pH 8.5). After washing in PBS, cells were incubated for 1 h at room temperature with anti-BrdU mAb (347580, BD Biosciences). After washes, secondary antibodies coupled to Alexa Fluor were incubated for 30 min at room temperature. After washing three times with PBS, cells were mounted with mounting medium (Dako) supplemented with DAPI (Sigma). For EdU-mediated foci visualization, EdU was detected using the Click-iT EdU Alexa Fluor 488 Imaging Kit (Life Technologies) according to manufacturer's instructions.

For detection of heterochromatin-associated proteins and repressive histone marks, cells were pre-extracted or not with CSK100 buffer as described above for 5 min on ice with gentle agitation then fixed with 4% paraformaldehyde. GFP-REV3L mutants were detected by autofluorescence, and interesting proteins were visualized by immunodetection with indicated primary antibodies for 1 h then processed as described above. Images were acquired using Axio Imager Z1 microscope (Zeiss) or confocal Leica SPE.

For quantification of immunofluorescence in UV-irradiated cells, MEFs growing on coverslips were incubated with 30 µM EdU for 15 min just before UV irradiation at 4 J/m^2. The soluble proteins were extracted by incubating cells in CSK100 for 5 min then fixed with 4% paraformaldehyde. After washes in PBS, EdU was detected using the Click-iT EdU Alexa Fluor 488 Imaging Kit (Life Technologies) according to the manufacturer's instructions. Subsequently, coverslips were blocked for 30 min with 3% BSA in PBS and then incubated 1 h 30 at room temperature with indicated primary antibodies in PBS containing 3% BSA, 0.1% Tween-20. The coverslips were further washed with PBS and incubated with Alexa Fluor 555 for 1 h at room temperature. The DNA was stained with DAPI, and the coverslips were mounted in fluorescent mounting medium (DAKO). Images were acquired on an Axio Imager Z1 microscope equipped with a motorized stage, EC Plan-Neofluar x20 dry objective, and a digital monochrome Hamamatsu ORCA-ER camera. For each condition, 20–50 images were acquired under non-saturating conditions. Identical settings were applied to all samples within one experiment. Images were analyzed with the Axio Vision software (Zeiss). Specific module was generated for nuclei segmentation based on DAPI signal according to intensity threshold, generating a mask that identified each individual nucleus as an individual object. Focus segmentation for 51BP1 and RAD51 was performed using an integrated spot-detection module. After segmentation and pixel quantification, the quantified values for each cell/foci (mean and total intensities, area, number of foci) were extracted and exported to a home-made software as well as to CellProfiler software.

DNA FISH on metaphase chromosome spreads

A mouse major satellite repeat probe was generated by PCR with 5′-ATATGTTGAGAAAACTGAAAATCACG-3′ and 5′-CCTTCAGTG TGCATTTCTCATTTTTCAC-3′ primers and murine genomic DNA as a template. The probe was labeled by nick translation with biotin-16-dUTP. *Rev3l*$^{-/-}$ and *Rev3l*$^{+/+}$ MEFs were treated with 0.15 µg/ml of colcemid for 3 h before fixation to arrest cells in metaphase.

Trypsinized cells were rinsed in PBS and incubated in 75 mM KCl for 15 min then fixed with ethanol: acetic acid (3:1 ratio). The cells were then dropped onto glass slides and left to dry. Samples were incubated with 0.1 mg/ml RNase A solution at 37°C for 1 h in humid chamber then washed once for 5 min in 2× SSC. Slides were dehydrated in 70, 85, and 100% (v/v) ethanol for 5 min each and then air-dried. Metaphase chromosome spreads were hybridized with mouse major satellite probe in hybridization buffer (10% dextran sulfate, 50% formamide, 2× SSC, 1% Tween-20), denatured on a hot plate at 72°C for 4 min and then incubated overnight at 37°C in a humid chamber. Following hybridization, slides were washed with 50% formamide and 2× SSC pH 7 solution then washed in 2× SSC and incubated in BlockAid blocking solution (Thermo Fisher Scientific, B10710) and 0.05% Tween-20. Immunodetection was performed by successive incubation (30 min at 37°C in BlockAid solution) with Alexa Fluor 555 streptavidin (S32355, Thermo Fisher Scientific), rabbit anti-streptavidin biotin conjugated followed by Alexa Fluor 555 streptavidin. Samples were washed three times with 1× PBS after each incubation. After final wash, samples were mounted with VectaShield (Vector Laboratories, H-1000). Images were acquired using Axio Imager Z1 microscope (Zeiss).

DNA combing-coupled FISH

HeLa cells transfected with non-targeting siRNA (siNT) or siRNA against *Rev3l* were pulse-labeled for 30 min with IdU followed by 30 min with CldU (100 μM final concentration), and a final 10-min incubation with 1 mM thymidine. Genomic DNA was combed onto in-house silanized coverslips using a Genomic Vision apparatus. FISH and immunostaining of analogues were performed essentially as described in Mendez-Bermudez *et al* (2018). Slides were denatured in 1 N NaOH for 6 min at room temperature, rinsed in 1× PBS at 4°C for 5 min, and dehydrated in 70% ethanol (v/v) for 5 min at −20°C, followed by 85 and 100% ethanol for 5 min each at room temperature.

For SatIII hybridization, coverslips were incubated overnight at 37°C with the PNA probe (SatIII: Biotin-O-TTCCATTCCATTCCATTCCA; Eurogentec) diluted in hybridization mix (10% dextran sulfate, 50% formamide, 2× SSC, 1% Tween-20) to a final concentration of 1 μM and denatured for 5 min at 95°C. Coverslips were washed three times in 50% formamide, 2× SSC pH 7, three times in 2× SSC and one time in PBS, before blocking for 30 min at 37°C in BlockAid blocking solution (Thermo Fisher Scientific, B10710). Immunodetection was performed by successive incubations (30 min at 37°C in BlockAid solution) with: (i) Alexa Fluor 488 streptavidin (Thermo Fisher Scientific, S32354, 1/100), (ii) rabbit anti-streptavidin-biotin conjugated (Rockland, 200-406-095, 1/50), (iii) Alexa Fluor 488 streptavidin (1/100) + mouse anti-BrdU (BD Biosciences, 347580, 1/5) + rat anti-BrdU (Abcam Ab6326, 1/50), and (iv) goat anti-mouse Cy5.5 (Abcam ab6947, 1/100) + goat anti-rat Alexa Fluor 555 (Thermo Fisher Scientific, A21434, 1/100). Coverslips were washed three times with 1× PBS after each incubation, with the addition of a 6-min wash in NaCl 0.5 M, Tris 20 mM pH 7.8, Tween-20 0.5% after step 3. After final wash, samples were mounted with VectaShield (Vector Laboratories, H-1000).

For global replication analysis, coverslips were blocked in BlockAid solution and incubated with: (i) mouse anti-BrdU + rat anti-BrdU, (ii) goat anti-mouse Alexa Fluor 488 (Thermo Fisher Scientific, A11029, 1/50) + goat anti-rat Alexa Fluor 555 (Thermo Fisher Scientific, A21434, 1/50), (iii) mouse anti-ssDNA (DSHB, AB10805144, 1/25), (iv) goat anti-mouse Cy5.5 (Abcam ab6947, 1/100), and (v) donkey anti-goat Cy5.5 (Abcam ab6951, 1/100).

Images were acquired on an epifluorescence microscope Axio Imager.Z2(Zeiss) equipped with a motorized stage and a 63× objective lens (PL APO, NA 1.4 Oil DIC M27) connected to a charge-coupled device camera (Cool-SNAP HQ2; Roper Scientific). MetaMorph software (Roper Scientific) was used for image acquisition and analysis.

Only unbroken replication signals, as evidenced by embedment in total DNA staining or SatIII FISH signal, were taken into consideration. Fork speed (in kb/min) was calculated by divided the length of the IdU or CldU track (in kb) by the labeling time (in min). Fork ratio was calculated for each unbroken bicolor replication signal as the ratio between max (IdU length, CldU length) and min (IdU length, CldU length). Statistical analysis was performed with the nonparametric Mann–Whitney rank-sum test using GraphPad Prism 6 software.

Proximity Ligation Assay (PLA)

Mouse embryonic fibroblast (MEF) Flag-REV3L were seed on coverslips for 24 h and washed with PBS; then, protein pre-extraction was carried out with CSK50 (50 mM NaCl for PLA Flag-Histones) or CSK100 (100 mM NaCl for PLA DSB markers) for 5 min on ice, fixed with 4% paraformaldehyde, and blocked in BSA-PBS 3% for 30 min. Coverslips were incubated with primary antibodies for 1 h at RT. Proximity ligation was performed using the Duolink® In Situ Red Starter Kit Mouse/Rabbit (Sigma-Aldrich) according to the manufacturer's protocol. The oligonucleotides and antibody-nucleic acid conjugates used were those provided in the Sigma-Aldrich PLA kit (DUO92101). Images were quantified by counting the number of foci per nucleus using ImageJ software. Statistical analysis was performed with the nonparametric Mann–Whitney rank-sum test using GraphPad Prism 5 software.

Southern blotting

DNA methylation status at major satellites was performed on 5 μg of genomic DNA (gDNA) extracted from *Rev3L*$^{-/-}$ and *Rev3L*$^{+/+}$ fibroblasts and digested with the methylation-sensitive enzyme MaeII (NEB) overnight at 65°C. After precipitation with EtOH, the digested gDNA was run on an agarose gel overnight and successively depurinated with a 250 mM HCl solution, denatured with a solution containing 0.5 M NaOH and 1.5 M NaCl. After neutralization, the gDNA was transferred onto a Hybond-N^{+} membrane (GE Healthcare Amersham) overnight. The membrane was prehybridized with hybridization buffer (6× SSC, 5× Denhardt's, 0.5% SDS) containing 100 μg/ml salmon sperm DNA for 1 h at 65°C and incubated overnight at 65°C with the hybridization buffer containing radiolabeled probes against major satellites (5′-GTGAAATATGGCGAGGAAAACT-3′). The membrane was washed once with 5× SSC, 0.1% SDS at 65°C for 10 min, twice with 3× SSC, 0.1% SDS at 65°C for 10 min and twice with 2× SSC, 0.1% SDS at room temperature for 10 min before visualization using X-ray film.

RNA extraction and RT–qPCR

Total RNAs were extracted from cell cultures and purified using RNeasy Mini Kit (QIAGEN), and concentration was measured with a Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA). Total RNAs were then reverse transcribed using Superscript III (Invitrogen). cDNAs were used for quantitative PCR using TaqMan Gene Expression Assays (Applied Biosystems) on an ABI Prism 7500 apparatus (Applied Biosystems). *GAPDH* was used as an internal control. The comparative threshold (ΔCt) method was used.

ChIP-qPCR

Rev3L$^{-/-}$ and *Rev3L*$^{+/+}$ fibroblasts were seeded at 1.5 × 10^6 and 10^6 cells per dish, respectively. After 24 h, cells were cross-linked with 1% formaldehyde for 10 min at RT and formaldehyde was quenched with Glycine 0.1 M for 5 min. Cells were washed twice with PBS + protease inhibitor, scraped in PBS + protease inhibitor, and then centrifuged at 800 × *g* –4°C for 5 min. Cell pellets were lysed in 1 ml of Cell Lysis Buffer (20 mM Tris–HCl [pH 8], 85 mM KCl, 0.5% NP-40) for 15 min on ice and centrifuged at 800 × *g* at 4°C for 5 min. Pellets were then resuspended in 1 ml of Nuclei Lysis Buffer (Tris–HCl pH 8 50 mM, EDTA pH 8 10 mM, SDS 1%), sonicated using QSonica, Amplitude 40% for 10 cycles of 15 s ON/45 s OFF, and samples were centrifuged at 10,000 × *g* –4°C for 10 min. For each IP, supernatants were diluted 10 times in Dilution Buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris–HCl [pH 8], 150 mM NaCl) and incubated with 5 μg antibody, 20 μl Dynabeads Magnetic beads and 1% protease inhibitor cocktail overnight at 4°C with rotation. After magnetic separation, supernatant was removed and beads were washed successively with Low Salt Immune Complex Wash Buffer, High Salt Immune Complex Wash Buffer, and LiCl Immune Complex Wash Buffer followed by magnetic separation. The last wash was performed in 200 μl TE Buffer for 3min at 4°C with rotation; then, beads were resuspended in 100 μl TE Buffer. Samples were then deproteinated by adding Proteinase K and incubated at 55°C overnight with shaking. After proteinase K inactivation, beads were removed by magnetic separation to keep the supernatant. DNA purification was performed by phenol/chloroform extraction followed by ethanol precipitation and DNA resuspension in EB buffer. DNA samples were analyzed by quantitative real-time PCR on the Applied Biosystems Real Time PCR system (7500 system) using the Maxima SYBR Green/Rox qPCR Mastermix (Thermo Fisher Scientific). The primers used are listed below.

Replication-timing experiments and microarrays

Cells were incubated with 50 μM of BrdU for 1.5 h and collected, washed three times with DPBS (Dulbecco's phosphate-buffered saline, Gibco Life Technologies), and then fixed in ethanol 75%. Cells were resuspended in DPBS with RNAse (0.5 mg/ml) and then with propidium iodide (50 μg/ml) followed by incubation in the dark at room temperature for 30 min with low agitation. Two fractions of 100,000 cells, S1 and S2 corresponding to early and late S-phase fractions, respectively, were sorted by flow cytometry using INFLUX 500 (Cytopeia BD Biosciences). Whole DNA was extracted with a lysis buffer (50 mM Tris [pH 8], 10 mM EDTA, 300 mM NaCl, 0.5% SDS) and 0.2 mg/ml of Proteinase K for 2 h at 65°C. Neo-synthesized DNA was immunoprecipitated with BrdU antibodies (Anti-BrdU Pure, BD Biosciences, # 347580) as previously described (Fernandez-Vidal *et al*, 2014). To control the quality of enrichment of early and late fractions in S1 and S2, qPCR was performed with BMP1 oligonucleotides (early control) and with Dppa2 oligonucleotides (late control; data not shown, (Hiratani *et al*, 2008)). Microarray hybridization requires a minimum of 1,000 ng of DNA. To obtain sufficient specific immunoprecipitated DNA for this hybridization step, whole genome amplification was conducted (WGA, Sigma) on immunoprecipitated DNA. A post-WGA qPCR was performed to preserve specific enrichment in both S1 and S2 fractions. Early and late amplified neo-synthesized DNAs were then labeled with Cy3 and Cy5 ULS molecules, respectively (Genomic DNA labeling Kit, Agilent). The hybridization was performed according to the manufacturer instructions on 4×180K mouse microarrays (SurePrint G3 Mouse CGH Microarray Kit, 4x180K, AGILENT Technologies, reference genome: mm9). Microarrays were scanned with an Agilent High-Resolution C Scanner using a resolution of 3 μm and the autofocus option. Feature extraction was performed with the Feature Extraction 9.1 software (Agilent Technologies). For each experiment, the raw data sets were automatically normalized by the Feature extraction software. Analysis was performed with the Agilent Genomic Workbench 5.0 software. The $\log_2$-ratio timing profiles were smoothed using the Triangular Moving Average option of the Agilent Genomic Workbench 5.0 software with the linear algorithm and 500 kb windows.

Identification of replication domains and changes in replication-timing profiles

To determine the replication domains in control cells and in *Rev3L*$^{-/-}$ MEFs, the algorithms from the STAR-R software were exploited as previously described (Hadjadj *et al*, 2020). A comparison was conducted between early and late domains from both cell lines in order to determine segments where replication-timing changes.

GC content

For the GC content, two steps were required: (i) Loading of the DNA sequence from interval files was performed with the Extract Genomic DNA tool with the mm9 genome annotation and (ii) calculation of the GC content for each interval by the GeeCee EMBOSS tool installed on the GALAXY Web site.

Intersection and coverage with H3K27ac, H3K4me3, and H3K27me3 marks

Data are loaded from the UCSC Web site. Peak coordinates of H3K27me3 come from the Broad Institute in the mm8 genome annotation (http://genome-euro.ucsc.edu/cgi-bin/hgTrackUi?hgsid=199709818_uvGUfB28yATavzb9eAlxuHR5KGbQ&c=chr12&g=broadChromatinChIPSeq). The conversion of coordinates to mm9 genome annotation was performed with the Lift-over tool available in the GALAXY Web site (https://usegalaxy.org/). The peak coordinates of H3K4me3 and H3K27ac were obtained from the Ludwig Institute for Cancer Research via the UCSC Web site (http://genome-euro.ucsc.edu/cgi-bin/hgTrackUi?hgsid=208559602_SARngKJBy54VXgefrlgnd

8rJaWFD&c = chr12&g = wgEncodeLicrHistone) in MEF mouse cell line and then used knowingly for intersection and coverage with the Intersection and Coverage tools installed on the GALAXY Web site. The statistical analysis was performed with the R package.

LINE-1 content

We obtained the line1 coordinates from the RepeatMasker track via the UCSC Web site. Intersection and coverage were performed with the Intersection and Coverage tools installed on the GALAXY Web site.

Structural variant analysis

Single cells from *Rev3L*$^{-/-}$ and *Rev3L*$^{+/+}$ fibroblasts were isolated, amplified, and subjected to whole genome sequencing. The genomes were sequenced according to the manufacturer protocols (BGI Tech solutions, Hong Kong, Co., Ltd) with a mean coverage of 30× using 150 bp paired-end reads with BGISEQ-500 sequencer. Reads were then aligned to the mm9 mouse genome using BWA mem software (Li & Durbin, 2009) and sorted with SAMtools (Li *et al*, 2009). Then, we removed duplicates from the BAM file according to the GATK Best Practices pipeline (Van der Auwera *et al*, 2013). To identify structural variants (SV) from genomic data, we used GRIDSS2 (Cameron *et al*, 2017) and Manta (Chen *et al*, 2016) software and compared replicates 1 and 2 between each other for *Rev3l*$^{+/+}$ and *Rev3l*$^{-/-}$ samples. Resulting structural variants from two methods were intersected using BEDTools for each sample (Quinlan, 2014) with overlap at least 90% and used for further analysis.

Gene expression profiling data analysis

RNA was extracted using Qiagen column (RNeasy kit) according to the manufacturer's protocols. The quantity and purity of the total RNA were evaluated using a NanoDrop spectrophotometer. Quality of RNA samples was assessed using Lab-on-a-chip Bioanalyser 2000 technology (Agilent Technologies), based on the 28S/18S ribosomal RNAs ratio. All samples included in this study displayed a ratio of ribosomal RNAs between 1.5 and 2.

Labeling of RNA samples was done according to Agilent oligo Cy5 or Cy3 probes labeling protocol using the Agilent Low Input QuickAmp labeling kit (ref 5190-2306) adapted for small amount of total RNA (100 ng total RNA per reaction). Hybridization was performed using the Agilent Hybridization Protocol (Gene expression Hyb kit Large - ref 5188-5280). Scanning was performed with an Agilent G2505C DNA Microarray scanner using default parameters: 20 bits mode, 3 μm resolution, at 20 C in low ozone concentration environment. Microarray images were analyzed by using Feature extraction software version 10.7.3.1 from Agilent technologies and Agilent normalization protocol GE2_107_Sep09 with 028005_D_F_20130207 as design. Default settings are used. A quantile array normalization was performed on raw MedianSignal values with in-house scripts using Bioconductor LIMMA package. Controls and flags spots were filtered. Missing values were completed with KNN imputation method. Then, the median of all probes for a given transcript was computed to summarize the data. Normalized data (intensities in log2) were imported in BrB Arrays Tools (https://brb.nci.nih.gov/BRB-ArrayTools/). Class Comparison was performed as weighted *t*-test and an FDR adjustment ($< 10^{-5}$) for multiple testing between Rev3l wt and KO samples (Class Comparison in BrB Arrays Tools 4.6.0) [https://brb.nci.nih.gov/BRB-ArrayTools]. Gene expressions altered by at least threefold between conditions were considered.

Antibodies

Antibody	Source	Identifier
Mouse Monoclonal anti-Flag M2	Sigma-Aldrich	F1804
Mouse Monoclonal anti-flag M2 affinity gel	Sigma-Aldrich	A2220
Rabbit Polyclonal anti-Flag	Sigma-Aldrich	F7425
Rabbit Polyclonal anti-Histone H3K4me3	Abcam	ab8580
Goat anti-Histone H3	Abcam	ab12079
Rabbit polyclonal anti-Histone H4	Active Motif	39269
Rabbit Polyclonal anti-ATRX	Santa Cruz	sc-15408
Mouse Monoclonal anti-BrdU (for FACS) BU20a	Dako	MO744
Purified Mouse Anti-BrdU (for IF) clone B44	BD Biosciences	347580
Rabbit Polyclonal anti-Histone H3K9me3	Abcam	ab8898
Rabbit Polyclonal anti-acetyl Histone H3 (Lys 9/14)	Cell Signaling technology	9677
Mouse Monoclonal anti-Histone H3K27me3	Abcam	ab6002
Mouse Monoclonal anti-Histone H4K20me3	Active Motif	39672
Rabbit polyclonal anti-Histone H4 (acetyl K12)	Abcam	ab46983
Rabbit monoclonal anti-Histone H4 (acetyl K16)	Abcam	ab109463
Rabbit monoclonal anti-Histone H4 (acetyl K5)	Abcam	ab51997
Mouse monoclonal anti-HP1α	Millipore	05-689
Mouse monoclonal anti-HP1β	Abcam	ab10478
Mouse Monoclonal anti-phospho Histone H2A.X (ser139)	Millipore	05-636
Rabbit Polyclonal anti-53BP1	Abcam	ab21083
Rabbit monoclonal anti-HA Tag	Cell Signaling Technologies	3724
Mouse monoclonal anti-His	Cell Signaling Technologies	2366
Mouse monoclonal anti-V5	Invitrogen	R962-25
Mouse monoclonal anti-HP1γ	Millipore Sigma	MAB3450
Anti-GST-HRP conjugate	Millipore Sigma	RPN1236
Rabbit monoclonal anti-Mad2L2/REV7	Abcam	ab180579
Mouse monoclonal anti-Chk1 (G-4)	Santa Cruz	sc-8408

Table (continued)

Antibody	Source	Identifier
Rabbit anti-pChk1 S345 (133D3)	Cell Signaling	2348S
GFP-Trap	Chemotek	gta-20
Rabbit anti-RPA2	Bethyl	A300-244A
Rabbit anti-pRPA2 S4/S8	Bethyl	A300-245A
Rabbit anti-KAP1	Abcam	ab10484
Rabbit anti-pKAP1 (Ser824)	Bethyl	A300-767A
Rabbit anti-SCAI antibody EPR4128	Abcam	124688
Rabbit anti-Rad51 antibody EPR4030	Abcam	ab133534
Alexa Fluor 555 streptavidin	Thermo Fisher Scientific	S32355
rabbit anti-streptavidin biotin conjugated	Rockland	200-406-095
Alexa Fluor 488 streptavidin	Thermo Fisher Scientific	S32354
rat anti-BrdU	Abcam	ab6326
goat anti-mouse Cy5.5	Abcam	ab6947
anti-rat Alexa Fluor 555	Thermo Fisher Scientific	A21434
anti-ssDNA	DSHB	AB10805144
anti-goat Cy5.5	Abcam	ab6951

Primers

Gene	Forward	Reverse
Primers for ChIP-qPCR		
Hoxb2	CCGAGTGAGTCCGTTTGGTT	GACTCCAGCCATGAGTAGCC
Hoxb8	GCCGGATGCAAAATACCGAC	GAGAAGATGTGGGGTGGGTG
Hoxb9	CCCTTTCACAGGTGAGTCCC	GGTGTCCACAGGAAGAGCAA
WT1	GAGGGAGGGAGATGAGAGGC	CCTGCATCTCAGGGCACTTT
FoxG1	TGAGGACAGGCCAGGAAAAC	GTCAAGGCTTCCATGTGTGC
Gata6	CGACTTGGGAGGACCTGTTG	TGGTACATTTCCTCCGGCTG
Primers for screening KI mouse line		
Rev3l	ACGAGTTCGCGGGTCCTTA GAGGTC	ACTTTCTACAGCCACAGCATCTC
Primers for major satellite probe (FISH)		
Maj sat	ATATGTTGAGAAAACTGAAA ATCACG	CCTTCAGTGTGCATTTCTCATT TTTCAC

List of TaqMan Gene Expression Assays used in this study.

(TaqMan Gene Expression Assays consist of forward and reverse primers with TaqMan minor groove binder probe for each gene).

Gene	Catalog #
Hoxb2	Mm04209931-m1
Hoxb8	Mm00439368-m1
Hoxb9	Mm01700220-m1
WT1	Mm01337048-m1

Table (continued)

Gene	Catalog #
Dlk1	Mm00494477-m1
Slcl8a4	Mm0045056-m1
Dcn	Mm0051435-m1
H19	Mm011567221-g1
GAPDH	Mm999999915-g1

Data availability

- Microarray data: EMBL-EBI E-MTAB-8338 (https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-8338/).
- Replication-timing data: GSE178927 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE178927).

Expanded View for this article is available online.

Acknowledgements

We thank T. Jenuwein (Max Planck Institute, Germany) for providing *Suv39h dn* cells. We are also grateful to Niels de Wind (Leiden University, Netherland) and Julien Duxin (Copenhagen University, Denmark) for sharing unpublished results on HP1 and SCAI, respectively. P.L.K. and E.D. express their gratitude to Michelle Debatisse and Stephane Koundrioukoff for their help and valuable advice on the application of the DNA combing-FISH technique to study replication dynamics in pericentromeric heterochromatin. We thank Nikola Djordjevic, Charlène Garandeau, Loélia Babin, and Kei-ichi Takata for experimental help and all members of the Kannouche lab for helpful discussions, especially Said Aoufouchi. This work was performed thanks to Gustave Roussy core facilities (Cell imaging, Genomics and Bio-informatics). The P.L.K lab was supported by La Ligue Nationale contre le Cancer (Equipe labellisée). This work was supported by grants from Institut National du Cancer (INCa) PLBIO16-011 (P.L.K. and J-C.C), INCa-DGOS-Inserm 12551, ARC-ARCPGA12019120001055_1578, and ANR-14-CE10-0008-01 (P.L.K. and J-C.C). S.A.S received support from the Ministère de l'Enseignement Supérieur et de la Recherche and the Foundation ARC and B.B received support from the Ministère de l'Enseignement Supérieur et de la Recherche and the Foundation pour la Recherche Médicale. The J-C.C lab is supported by the IdEX Université de Paris ANR-18-IDEX-0001. R.D.W. and J.T were funded by National Institutes of Health grant CA193124, Department of Defense grant W81XWH-17-10239, and the J. Ralph Meadows Chair in Carcinogenesis Research.

Author contributions

BBY performed experiments for protein interaction (co-IP and PLA) and ChIP-qPCR; SA-S performed GFP-fused protein localization experiments and FACS analyses. QD performed FISH on metaphase spread, and ED carried out experiments with DNA combing. JG and RC performed immunofluorescence analysis on UV-irradiated cells. JCC, ND, GB, and DG designed experiments for replication timing, and JCC performed genome-wide analysis. PD performed transcriptome analysis. AY, SN, and PLK designed experiments for whole genome sequencing and AY and SN performed mutational analysis. JT and SB carried out GST pull-down and Co-IP experiments involving HP1 protein. SSL designed the endogenous tagged Rev3l mouse and derived all MEF cell lines used in this research. MTM-G helped in the design of cellular studies, and CP brought technical

support in many experiments. PLK and RDW wrote the manuscript and all authors reviewed it. PLK, RDW, JCC, and JT conceived and planned the study.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Aladjem MI (2007) Replication in context: dynamic regulation of DNA replication patterns in metazoans. *Nat Rev Genet* 8: 588 – 600

Allshire RC, Karpen GH (2008) Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nat Rev Genet* 9: 923 – 937

Baranovskiy AG, Lada AG, Siebler HM, Zhang Y, Pavlov YI, Tahirov TH (2012) DNA polymerase delta and zeta switch by sharing accessory subunits of DNA polymerase delta. *J Biol Chem* 287: 17281 – 17287

Bergoglio V, Boyer AS, Walsh E, Naim V, Legube G, Lee MY, Rey L, Rosselli F, Cazaux C, Eckert KA *et al* (2013) DNA synthesis by Pol eta promotes fragile site stability by preventing under-replicated DNA in mitosis. *J Cell Biol* 201: 395 – 408

Bhat A, Andersen PL, Qin Z, Xiao W (2013) Rev3, the catalytic subunit of Polzeta, is required for maintaining fragile site stability in human cells. *Nucleic Acids Res* 41: 2328 – 2339

Burgers PMJ, Kunkel TA (2017) Eukaryotic DNA replication fork. *Annu Rev Biochem* 86: 417 – 438

Cameron DL, Schroder J, Penington JS, Do H, Molania R, Dobrovic A, Speed TP, Papenfuss AT (2017) GRIDSS: sensitive and specific genomic rearrangement detection using positional de Bruijn graph assembly. *Genome Res* 27: 2050 – 2060

Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Kallberg M, Cox AJ, Kruglyak S, Saunders CT (2016) Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* 32: 1220 – 1222

Cornacchia D, Dileep V, Quivy JP, Foti R, Tili F, Santarella-Mellwig R, Antony C, Almouzni G, Gilbert DM, Buonomo SB (2012) Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. *EMBO J* 31: 3678 – 3690

Despras E, Sittewelle M, Pouvelle C, Delrieu N, Cordonnier AM, Kannouche PL (2016) Rad18-dependent SUMOylation of human specialized DNA polymerase eta is required to prevent under-replicated DNA. *Nat Commun* 7: 13326

Desprat R, Thierry-Mieg D, Lailler N, Lajugie J, Schildkraut C, Thierry-Mieg J, Bouhassira EE (2009) Predictable dynamic program of timing of DNA replication in human cells. *Genome Res* 19: 2288 – 2299

Dimitrova DS, Berezney R (2002) The spatio-temporal organization of DNA replication sites is identical in primary, immortalized and transformed mammalian cells. *J Cell Sci* 115: 4037 – 4051

Donley N, Thayer MJ (2013) DNA replication timing, genome stability and cancer: late and/or delayed DNA replication timing is associated with increased genomic instability. *Semin Cancer Biol* 23: 80 – 89

Esposito G, Godindagger I, Klein U, Yaspo ML, Cumano A, Rajewsky K (2000) Disruption of the Rev3l-encoded catalytic subunit of polymerase zeta in mice results in early embryonic lethality. *Curr Biol* 10: 1221 – 1224

Eustermann S, Yang JC, Law MJ, Amos R, Chapman LM, Jelinska C, Garrick D, Clynes D, Gibbons RJ, Rhodes D *et al* (2011) Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nat Struct Mol Biol* 18: 777 – 782

Farkash-Amar S, Lipson D, Polten A, Goren A, Helmstetter C, Yakhini Z, Simon I (2008) Global organization of replication time zones of the mouse genome. *Genome Res* 18: 1562 – 1570

Fernandez-Vidal A, Guitton-Sert L, Cadoret JC, Drac M, Schwob E, Baldacci G, Cazaux C, Hoffmann JS (2014) A role for DNA polymerase theta in the timing of DNA replication. *Nat Commun* 5: 4285

Fundele RH, Fundele RH, Surani MA (1994) Experimental embryological analysis of genetic imprinting in mouse development. *Dev Genet* 15: 515 – 522

Gan GN, Wittschieben JP, Wittschieben BO, Wood RD (2008) DNA polymerase zeta (pol zeta) in higher eukaryotes. *Cell Res* 18: 174 – 183

Goodarzi AA, Noon AT, Deckbar D, Ziv Y, Shiloh Y, Lobrich M, Jeggo PA (2008) ATM signaling facilitates repair of DNA double-strand breaks associated with heterochromatin. *Mol Cell* 31: 167 – 177

Guenatri M, Bailly D, Maison C, Almouzni G (2004) Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J Cell Biol* 166: 493 – 505

Guilbaud G, Rappailles A, Baker A, Chen CL, Arneodo A, Goldar A, d'Aubenton-Carafa Y, Thermes C, Audit B, Hyrien O (2011) Evidence for sequential and increasing activation of replication origins along replication timing gradients in the human genome. *PLoS Comput Biol* 7: e1002322

Hadjadj D, Denecker T, Guerin E, Kim SJ, Fauchereau F, Baldacci G, Maric C, Cadoret JC (2020) Efficient, quick and easy-to-use DNA replication timing analysis with START-R suite. *NAR Genom Bioinform* 2: lqaa045

Hansen RK, Mund A, Poulsen SL, Sandoval M, Klement K, Tsouroula K, Tollenaere MA, Raschle M, Soria R, Offermanns S *et al* (2016) SCAI promotes DNA double-strand break repair in distinct chromosomal contexts. *Nat Cell Biol* 18: 1357 – 1366

Hayano M, Kanoh Y, Matsumoto S, Renard-Guillet C, Shirahige K, Masai H (2012) Rif1 is a global regulator of timing of replication origin firing in fission yeast. *Genes Dev* 26: 137 – 150

Hiragami-Hamada K, Soeroes S, Nikolov M, Wilkins B, Kreuz S, Chen C, De La Rosa-Velazquez IA, Zenn HM, Kost N, Pohl W *et al* (2016) Dynamic and flexible H3K9me3 bridging via HP1beta dimerization establishes a plastic state of condensed chromatin. *Nat Commun* 7: 11310

Hiratani I, Ryba T, Itoh M, Yokochi T, Schwaiger M, Chang CW, Lyou Y, Townes TM, Schubeler D, Gilbert DM (2008) Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol* 6: e245

Isobe SY, Nagao K, Nozaki N, Kimura H, Obuse C (2017) Inhibition of RIF1 by SCAI allows BRCA1-mediated repair. *Cell Rep* 20: 297 – 307

Johnson RE, Prakash L, Prakash S (2012) Pol31 and Pol32 subunits of yeast DNA polymerase delta are also essential subunits of DNA polymerase zeta. *Proc Natl Acad Sci USA* 109: 12455 – 12460

Johnson RE, Washington MT, Haracska L, Prakash S, Prakash L (2000) Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature* 406: 1015 – 1019

Kane DP, Shusterman M, Rong Y, McVey M (2012) Competition between replicative and translesion polymerases during homologous recombination repair in *Drosophila*. *PLoS Genet* 8: e1002659

Koren A, Polak P, Nemesh J, Michaelson JJ, Sebat J, Sunyaev SR, McCarroll SA (2012) Differential relationship of DNA replication timing to different forms of human mutation and variation. *Am J Hum Genet* 91: 1033 – 1040

Lang GI, Murray AW (2011) Mutation rates across budding yeast chromosome VI are correlated with replication timing. *Genome Biol Evol* 3: 799 – 811

Lange SS, Bedford E, Reh S, Wittschieben JP, Carbajal S, Kusewitt DF, DiGiovanni J, Wood RD (2013) Dual role for mammalian DNA polymerase zeta in maintaining genome stability and proliferative responses. *Proc Natl Acad Sci USA* 110: E687 – 696

Lange SS, Bhetawal S, Reh S, Powell KL, Kusewitt DF, Wood RD (2018) DNA polymerase zeta deficiency causes impaired wound healing and stress-induced skin pigmentation. *Life Sci Alliance* 1: 1 – 15

Lange SS, Tomida J, Boulware KS, Bhetawal S, Wood RD (2016) The polymerase activity of mammalian DNA Pol zeta is specifically required for cell and embryonic viability. *PLoS Genet* 12: e1005759

Lange SS, Wittschieben JP, Wood RD (2012) DNA polymerase zeta is required for proliferation of normal mammalian cells. *Nucleic Acids Res* 40: 4473 – 4482

Lee YS, Gregory MT, Yang W (2014) Human Pol zeta purified with accessory subunits is active in translesion DNA synthesis and complements Pol eta in cisplatin bypass. *Proc Natl Acad Sci USA* 111: 2954 – 2959

Lemaitre C, Soutoglou E (2014) Double strand break (DSB) repair in heterochromatin and heterochromatin proteins in DSB repair. *DNA Repair* 19: 163 – 168

Letessier A, Millot GA, Koundrioukoff S, Lachages AM, Vogt N, Hansen RS, Malfoy B, Brison O, Debatisse M (2011) Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature* 470: 120 – 123

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754 – 1760

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078 – 2079

MacAlpine DM, Rodriguez HK, Bell SP (2004) Coordination of replication and transcription along a *Drosophila* chromosome. *Genes Dev* 18: 3094 – 3105

Maison C, Bailly D, Peters AH, Quivy JP, Roche D, Taddei A, Lachner M, Jenuwein T, Almouzni G (2002) Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat Genet* 30: 329 – 334

Makarova AV, Burgers PM (2015) Eukaryotic DNA polymerase zeta. *DNA Repair* 29: 47 – 55

Makarova AV, Stodola JL, Burgers PM (2012) A four-subunit DNA polymerase zeta complex containing Pol delta accessory subunits is essential for PCNA-mediated mutagenesis. *Nucleic Acids Res* 40: 11618 – 11626

Manandhar M, Lowery MG, Boulware KS, Lin KH, Lu Y, Wood RD (2017) Transcriptional consequences of XPA disruption in human cell lines. *DNA Repair* 57: 76 – 90

Martin SK, Wood RD (2019) DNA polymerase zeta in DNA replication and repair. *Nucleic Acids Res* 47: 8348 – 8361

McCulloch SD, Kunkel TA (2008) The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res* 18: 148 – 161

Mendez-Bermudez A, Lototska L, Bauwens S, Giraud-Panis MJ, Croce O, Jamet K, Irizar A, Mowinckel M, Koundrioukoff S, Nottet N *et al* (2018) Genome-wide control of heterochromatin replication by the telomere capping protein TRF2. *Mol Cell* 70: 449 – 461.e5

Murzina N, Verreault A, Laue E, Stillman B (1999) Heterochromatin dynamics in mouse cells: interaction between chromatin assembly factor 1 and HP1 proteins. *Mol Cell* 4: 529 – 540

Natale F, Scholl A, Rapp A, Yu W, Rausch C, Cardoso MC (2018) DNA replication and repair kinetics of Alu, LINE-1 and satellite III genomic repetitive elements. *Epigenet Chromatin* 11: 61

Nguyen VH, Lavenier D (2009) PLAST: parallel local alignment search tool for database comparison. *BMC Bioinformatics* 10: 329

Nikolov I, Taddei A (2016) Linking replication stress with heterochromatin formation. *Chromosoma* 125: 523 – 533

Norio P, Kosiyatrakul S, Yang Q, Guan Z, Brown NM, Thomas S, Riblet R, Schildkraut CL (2005) Progressive activation of DNA replication initiation in large domains of the immunoglobulin heavy chain locus during B cell development. *Mol Cell* 20: 575 – 587

Nozawa RS, Nagao K, Masuda HT, Iwasaki O, Hirota T, Nozaki N, Kimura H, Obuse C (2010) Human POGZ modulates dissociation of HP1alpha from mitotic chromosome arms through Aurora B activation. *Nat Cell Biol* 12: 719 – 727

Ozeri-Galai E, Tur-Sinai M, Bester AC, Kerem B (2014) Interplay between genetic and epigenetic factors governs common fragile site instability in cancer. *Cell Mol Life Sci* 71: 4495 – 4506

Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A *et al* (2001) Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* 107: 323 – 337

Pirouz M, Pilarski S, Kessel M (2013) A critical function of Mad2l2 in primordial germ cell development of mice. *PLoS Genet* 9: e1003712

Polak P, Karlic R, Koren A, Thurman R, Sandstrom R, Lawrence M, Reynolds A, Rynes E, Vlahovicek K, Stamatoyannopoulos JA *et al* (2015) Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature* 518: 360 – 364

Quinlan AR (2014) BEDTools: the swiss-army tool for genome feature analysis. *Curr Protoc Bioinformatics* 47: 11 – 34

Quivy J-P, Roche D, Kirschner D, Tagami H, Nakatani Y, Almouzni G (2004) A CAF-1 dependent pool of HP1 during heterochromatin duplication. *EMBO J* 23: 3516 – 3526

Sachs M, Onodera C, Blaschke K, Ebata KT, Song JS, Ramalho-Santos M (2013) Bivalent chromatin marks developmental regulatory genes in the mouse embryonic germline in vivo. *Cell Rep* 3: 1777 – 1784

Sarkies P, Reams C, Simpson LJ, Sale JE (2010) Epigenetic instability due to defective replication of structured DNA. *Mol Cell* 40: 703 – 713

Schultz SS, Desbordes SC, Du Z, Kosiyatrakul S, Lipchina I, Studer L, Schildkraut CL (2010) Single-molecule analysis reveals changes in the DNA replication program for the POU5F1 locus upon human embryonic stem cell differentiation. *Mol Cell Biol* 30: 4521 – 4534

Schuster-Bockler B, Lehner B (2012) Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature* 488: 504 – 507

Schwienbacher C, Angioni A, Scelfo R, Veronese A, Calin GA, Massazza G, Hatada I, Barbanti-Brodano G, Negrini M (2000) Abnormal RNA expression of 11p15 imprinted genes and kidney developmental genes in Wilms' tumor. *Cancer Res* 60: 1521 – 1525

Seplyarskiy VB, Bazykin GA, Soldatov RA (2015) Polymerase zeta activity is linked to replication timing in humans: evidence from mutational signatures. *Mol Biol Evol* 32: 3158 – 3172

Sharma S, Hicks JK, Chute CL, Brennan JR, Ahn JY, Glover TW, Canman CE (2012) REV1 and polymerase zeta facilitate homologous recombination repair. *Nucleic Acids Res* 40: 682 – 691

Sima J, Chakraborty A, Dileep V, Michalski M, Klein KN, Holcomb NP, Turner JL, Paulsen MT, Rivera-Mulia JC, Trevilla-Garcia C *et al* (2019) Identifying cis elements for spatiotemporal control of mammalian DNA replication. *Cell* 176: 816 – 830.e18

Sonoda E, Okada T, Zhao GY, Tateishi S, Araki K, Yamaizumi M, Yagi T, Verkaik NS, van Gent DC, Takata M *et al* (2003) Multiple roles of Rev3, the catalytic subunit of polzeta in maintaining genome stability in vertebrates. *EMBO J* 22: 3188 – 3197

Stamatoyannopoulos JA, Adzhubei I, Thurman RE, Kryukov GV, Mirkin SM, Sunyaev SR (2009) Human mutation rate associated with DNA replication timing. *Nat Genet* 41: 393 – 395

Stone JE, Kissling GE, Lujan SA, Rogozin IB, Stith CM, Burgers PM, Kunkel TA (2009) Low-fidelity DNA synthesis by the L979F mutator derivative of *Saccharomyces cerevisiae* DNA polymerase zeta. *Nucleic Acids Res* 37: 3774 – 3787

Taddei A, Maison C, Roche D, Almouzni G (2001) Reversible disruption of pericentric heterochromatin and centromere function by inhibiting deacetylases. *Nat Cell Biol* 3: 114 – 120

Thiru A, Nietlispach D, Mott HR, Okuwaki M, Lyon D, Nielsen PR, Hirshberg M, Verreault A, Murzina NV, Laue ED (2004) Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin. *EMBO J* 23: 489 – 499

Tomida J, Takata K, Lange SS, Schibler AC, Yousefzadeh MJ, Bhetawal S, Dent SY, Wood RD (2015) REV7 is essential for DNA damage tolerance via two REV3L binding sites in mammalian DNA polymerase zeta. *Nucleic Acids Res* 43: 1000 – 1011

Tomida J, Takata KI, Bhetawal S, Person MD, Chao HP, Tang DG, Wood RD (2018) FAM35A associates with REV7 and modulates DNA damage responses of normal and BRCA1-defective cells. *EMBO J* 37: e99543

Tsao WC, Eckert KA (2018) Detours to replication: functions of specialized DNA polymerases during oncogene-induced replication stress. *Int J Mol Sci* 19: 1 – 25

Ui A, Nagaura Y, Yasui A (2015) Transcriptional elongation factor ENL phosphorylated by ATM recruits polycomb and switches off transcription for DSB repair. *Mol Cell* 58: 468 – 482

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J *et al* (2013) From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 43: 11 10 11 – 11 10 33

Van Sloun PP, Varlet I, Sonneveld E, Boei JJ, Romeijn RJ, Eeken JC, De Wind N (2002) Involvement of mouse Rev3 in tolerance of endogenous and exogenous DNA damage. *Mol Cell Biol* 22: 2159 – 2169

Vissers JH, van Lohuizen M, Citterio E (2012) The emerging role of Polycomb repressors in the response to DNA damage. *J Cell Sci* 125: 3939 – 3948

Watanabe Y, Maekawa M (2010) Spatiotemporal regulation of DNA replication in the human genome and its association with genomic instability and disease. *Curr Med Chem* 17: 222 – 233

Wittschieben JP, Reshmi SC, Gollin SM, Wood RD (2006) Loss of DNA polymerase zeta causes chromosomal instability in mammalian cells. *Cancer Res* 66: 134 – 142

Wittschieben J, Shivji MK, Lalani E, Jacobs MA, Marini F, Gearhart PJ, Rosewell I, Stamp G, Wood RD (2000) Disruption of the developmentally regulated Rev3l gene causes embryonic lethality. *Curr Biol* 10: 1217 – 1220

Zhang J, Bellani MA, James RC, Pokharel D, Zhang Y, Reynolds JJ, McNee GS, Jackson AP, Stewart GS, Seidman MM (2020) DONSON and FANCM associate with different replisomes distinguished by replication timing and chromatin domain. *Nat Commun* 11: 3951

Nucleic Acids Research, 2019 **1**
doi: 10.1093/nar/gkz705

SURVEY AND SUMMARY

DNA polymerase ζ in DNA replication and repair

Sara K. Martin and Richard D. Wood*

Department of Epigenetics & Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Smithville, TX, USA and The University of Texas MD Anderson Cancer Center UT Health Graduate School of Biomedical Sciences

Received June 18, 2019; Revised July 24, 2019; Editorial Decision July 30, 2019; Accepted August 08, 2019

ABSTRACT

Here, we survey the diverse functions of DNA polymerase ζ (pol ζ) in eukaryotes. In mammalian cells, REV3L (3130 residues) is the largest catalytic subunit of the DNA polymerases. The orthologous subunit in yeast is Rev3p. Pol ζ also includes REV7 subunits (encoded by *Rev7* in yeast and *MAD2L2* in mammalian cells) and two subunits shared with the replicative DNA polymerase, pol δ. Pol ζ is used in response to circumstances that stall DNA replication forks in both yeast and mammalian cells. The best-examined situation is translesion synthesis at sites of covalent DNA lesions such as UV radiation-induced photoproducts. We also highlight recent evidence that uncovers various roles of pol ζ that extend beyond translesion synthesis. For instance, pol ζ is also employed when the replisome operates suboptimally or at difficult-to-replicate DNA sequences. Pol ζ also participates in repair by microhomology mediated break-induced replication. A *rev3* deletion is tolerated in yeast but *Rev3l* disruption results in embryonic lethality in mice. Inactivation of mammalian *Rev3l* results in genomic instability and invokes cell death and senescence programs. Targeting of pol ζ function may be a useful strategy in cancer therapy, although chromosomal instability associated with pol ζ deficiency must be considered.

INTRODUCTION

DNA polymerase ζ (pol ζ) is universally present in eukaryotes including fungi, plants, and animals. Genetic studies of the budding yeast *Saccharomyces cerevisiae* have long established that pol ζ is a biologically fundamental enzyme, necessary for mediating most damage-induced mutagenesis. In this summary, we highlight newer discoveries of multiple roles for pol ζ. Pol ζ is a multi-subunit enzyme (Figure 1). The catalytic subunit is called Rev3p in budding yeast *S. cerevisiae*, and has 1504 amino acid residues. In mammalian cells, the orthologous protein REV3L is twice the size (3130 amino acids in human cells). Although the mammalian and yeast enzymes have some basic biochemical similarities, there are also significant functional differences, as summarized below.

Pol ζ structure and composition

Rev3p/REV3L is a member of the 'B-family' of DNA polymerases. The other B-family DNA polymerases in eukaryotes (Pols α, δ and ϵ, Figure 2A) are components of the core DNA replication apparatus. The C-terminal portion of Rev3p/REV3L encompasses most of the conserved DNA polymerase domain (Figure 2B, C). A conserved N-terminal domain is predicted to form part of the overall polymerase fold (1,2) (Figure 2D).

Structures of the homologous regions of other B-family DNA polymerases show that the N-terminal domain caps the exonuclease domain, and contains residues that make direct contacts to DNA (2–5). In the other B-family polymerases, the DNA sequence encoding the N-terminal domain is directly adjacent to that encoding the exonuclease domain (Figure 2A). However, for Rev3p/REV3L and its orthologs, a large insertion separates the N-terminal domain from an inactive exonuclease domain (Figure 2B, C). In yeast pol δ, a ∼45 amino acid region immediately downstream of the exonuclease domain (Figure 2A, light green) also folds with the N-terminal domain (3). This region is conserved and structural modeling predicts that it folds similarly with the N-terminal domain in Rev3p (Figure 2B–D).

A single exon encodes the large central insertion that interrupts the domains of the catalytic core, and its variation across species results in large differences in the protein size. In mammalian *REV3L* genes this exon is relatively gargantuan (1386 amino acids), resulting in a protein double the size of yeast Rev3p (6) (Figure 2A).

*To whom correspondence should be addressed. Tel: +1 512 237 9431; Fax: +1 512 237 6532; Email: rwood@mdanderson.org

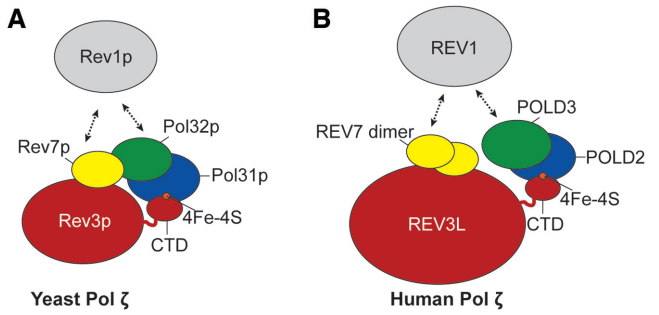


Figure 1. Pol ζ is a multi-subunit polymerase that interacts with the TLS master regulator, REV1. (**A**) Schematic of pol ζ from the budding yeast *Saccharomyces cerevisiae*. The catalytic subunit Rev3p binds to the accessory subunit Rev7p. The precise binding sites and dimeric state of Rev7p are unknown. Pol δ accessory subunits, Pol31p-Pol32p, bind to the C-terminal domain (CTD). Pol31p interaction is coordinated by the iron-sulfur cluster (4Fe-4S) in Rev3p. Rev7p and Pol32p interact directly. Rev7p and Pol32p mediate the interaction with Rev1p. (**B**) Schematic of human pol ζ. A dimer of REV7 can bind to the catalytic subunit, REV3L. POLD2 binds REV3L at the CTD, coordinated by the iron-sulfur (4Fe-4S) cluster. POLD3 binds to POLD2. It is unknown whether REV7 and POLD3 directly interact as Rev7p and Pol32p do in yeast. REV1 can bind human pol ζ through interactions with REV7 and POLD3.

While this single exon insert contains many stretches of predicted disorder, it is also home to a conserved positively charged domain that is necessary for efficient polymerase activity of the recombinant protein (7). A ∼250 aa region of the large insert of REV3L is highly homologous to the *NEXMIF (KIAA2022)* gene (6). This is annotated as Pfam domain DUF4683. The role this domain plays in REV3L or NEXMIF function is uncertain.

Like the other B-family polymerases, pol ζ contains multiple accessory subunits, including two subunits of pol δ: Pol31p/POLD2 and Pol32p/POLD3. The C-terminal domain of Rev3p/REV3L directly binds Pol31p/POLD2 (7–10) (Figure 1). The C-terminal domain contains two cysteine clusters: one that binds zinc (CysA) and one that binds an iron-sulfur [4Fe–4S] cluster (CysB) (8) (Figure 2A–C). CysB is essential for the binding of Pol31p/POLD2 to Rev3p/REV3L and pol ζ function (7,9,10). Pol32p/POLD3 binds indirectly through its interaction with Pol31p/POLD2. Pol ζ includes a third accessory subunit, Rev7p/REV7 (mammalian gene name *MAD2L2)* (11).

In mammalian REV3L there are two adjacent binding sites for REV7 with a consensus ϕϕxPxxxxPSR (where ϕ represents an aliphatic amino acid residue) (12,13). One of these mammalian binding sites resides in the large exon (Figure 2C). Crystal structures of REV7 bound to the two corresponding REV3L peptides have been obtained (14,15). Dimerization of REV7 appears to contribute to pol ζ biological function in mammalian cells (15,16). The exact interaction site or sites have yet to be discovered in yeast, and it is unknown whether Rev7p forms a dimer in yeast pol ζ. Rev7p greatly enhances pol ζ activity *in vitro* through an unknown mechanism (17). Similarly, Pol31p/POLD2 and Pol32p/POLD3 stimulate the activity of Rev3p/REV3L-Rev7p/REV7 (7,10).

A low resolution structure (23 Å) of yeast pol ζ obtained by negative stain electron microscopy suggests that all three regulator subunits form a hub, with Rev7p making direct contact to Pol32p (2) (Figure 1). Whether POLD3 interacts with REV7 in mammals has not been reported. The function of this hub in pol ζ remains largely unexplored. One important role is that it may serve to regulate recruitment of Rev1p/REV1, a major regulator of translesion DNA synthesis (TLS) (Figure 1). As will be discussed later, proper pol ζ function requires interaction with Rev1p/REV1. Rev1p/REV1 binds to pol ζ through direct interactions with both Rev7p/REV7 and Pol32p/POLD3 (12,18–23).

REV3L expression and isoforms

REV3L is ubiquitously expressed in many tissues to yield a low level of protein. There are two main isoforms of the human and mouse mRNA transcripts. In human cells, the well-documented alternative isoform has a 128 base pair insertion between positions 139 and 140 of the cDNA, resulting in three in-frame stop codons. An alternative translation start site in this insertion was proposed that would result in a REV3L protein with 78 fewer amino acids at the N-terminus (24). However, when it was experimentally tested whether N-terminal fragments of both isoforms could be expressed under their endogenous UTRs, only the reference isoform, and not the alternate isoform, was able to produce protein in vitro and in cellular translation systems (25). The alternative non-functional transcript isoform may be one way by which protein levels of REV3L are kept low in cells (25). Another means of limiting pol ζ levels may be proteolytic cleavage. A TASP1 cleavage site in REV3L is predicted by the eukaryotic linear motif resource (26) (Figure 2C). This would account for the prominent N-terminal proteolytic cleavage product of 60–70 kDa that is observed after REV3L overexpression (see Supplementary Figure S1 of (13)).

Pol ζ disruption leads to mitochondrial deficits

Mitochondrial function depends on the coordinated expression of many nuclear genes. Consequently, mitochondria are a large and sensitive target and they provide a readout for nuclear genome stability. It is known that disruption of many nuclear DNA repair and metabolism genes can lead to mitochondrial dysfunction, and this appears to be true for pol ζ disruption as well (27).

It is not clear, however, whether active pol ζ is physically present in mitochondria or is only resident in the nucleus. It has been proposed that Rev3p contains a mitochondrial localization signal, but the reference mammalian *REV3L* sequence does not contain a functional mitochondrial localization signal (27,28). A suggested mitochondrial localization signal was identified in the predicted product of the shorter alternative mRNA isoform of mammalian REV3L (27). Ectopic expression of a peptide containing the N-terminal sequence of the alternate isoform drove the peptide to mitochondria (27), but there is no convincing evidence to show that this mRNA isoform produces REV3L (commercial REV3L antibodies are not able to detect REV3L in

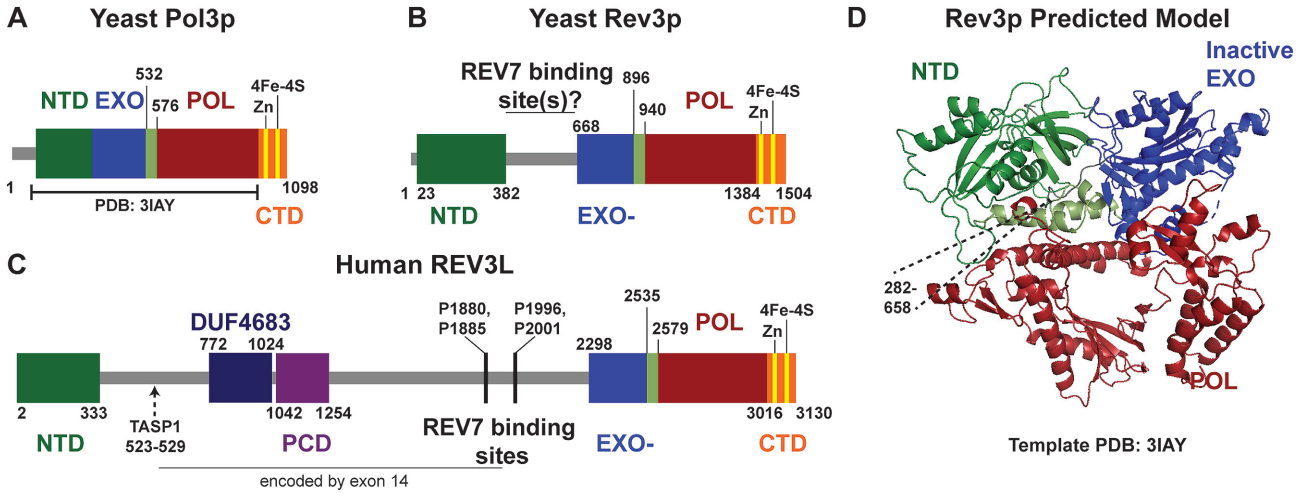


Figure 2. Rev3p/REV3L is a B-family DNA polymerase. Schematic of the catalytic subunits of (**A**) pol δ (Pol3p), (**B**) yeast pol ζ (Rev3p) and (**C**) human pol ζ (REV3L). All contain a conserved N-terminal domain (NTD, dark and light green), an exonuclease domain (EXO, blue), a B-family polymerase domain (POL, red), and a C-terminal domain (CTD, orange) with two metal binding sites: one that binds zinc (CysA) and one that binds an iron-sulfur cluster (4Fe-4S, CysB). The NTD is annotated by Interpro as domain IPR006133. The EXO domain is active in pol δ but inactive in pol ζ (EXO-). Rev7p binding likely occurs somewhere in between the NTD and the EXO- in Rev3p. In human REV3L, exon 14 encodes a large insertion that contains a positively charged domain (PCD, purple) and a domain of unknown function (DUF4683). Two REV7 binding sites (13) are well-conserved in mammals. (**D**) We generated a predicted model of Rev3p using PHYRE2 (125), with the known structure of Pol3p (3) as a template (PDB: 3IAY) given that Rev3p shares the highest sequence identity with pol δ. This shows the predicted folding of the NTD with the catalytic core of REV3L. The insert between the NTD and EXO- domains (282-658) was not included in the predicted model.

cells). There is currently no strong evidence for the presence of the other mammalian pol ζ subunits (or pol ζ function) within mitochondria.

Pol ζ as a major extender enzyme past DNA lesions

Damage removal is usually the first line of defense against UV radiation or chemical adducts in DNA. For example, base excision repair (BER) removes abasic sites, and nucleotide excision repair removes UV radiation-induced pyrimidine dimers. If lesions are not removed by the time of DNA replication, they can block replication fork progression. In order for replication to reach completion the lesion must be bypassed. When replicative DNA polymerases stall at a lesion, there are two well-characterized routes of lesion bypass (Figure 3). In template switching, the lesion is circumvented completely by synthesis using the undamaged sister strand. In TLS, the lesion is directly bypassed by specialized DNA polymerases.

The route to either template switching or TLS is directed by posttranslational modifications to proliferating cell nuclear antigen (Pol30p/PCNA) (29–31). The homotrimeric Pol30p/PCNA forms a sliding clamp around the DNA, serving as an anchor for DNA polymerases. When a replication fork stalls at a lesion, Pol30p/PCNA becomes monoubiquitinated at K164 and promotes TLS. Alternatively, Pol30p can be polyubiquitinated at K164 (in yeast, by Mms2p/Ubc13p/Rad5p complex), and this shunts bypass to the template switching pathway.

TLS DNA polymerases and replicative DNA polymerases have contrasting biochemical properties that serve their respective functions. The high fidelity of replicative DNA polymerases is boosted by a functional exonuclease domain that corrects most errors. In contrast, TLS polymerases generally have low fidelity and lack a functional exonuclease domain. In Y-family DNA polymerases, this low fidelity has been attributed to a more spacious active site, which allows TLS DNA polymerases to incorporate bases across from lesions that cannot be accommodated by the active sites of replication polymerases (32). Another distinction between TLS polymerases and replicative polymerases is processivity, the amount of DNA synthesized in a single event before dissociation. TLS polymerases are less processive than replication polymerases. For example, Y-family TLS pol η, even when stimulated by the physiologically relevant monoubiquitinated PCNA, has processivity dramatically lower than pol δ (33). Low processivity may be selectively advantageous for organisms, preventing error prone polymerases from uninterrupted synthesis of large swaths of DNA which could leave genomes riddled with errors. Given the introduction of errors that occur with TLS, it makes sense that this is a tightly regulated process.

A large part of this regulation centers on monoubiquitinated Pol30p/PCNA, which serves as a docking site for key TLS players. For instance, the master TLS regulator Rev1p/REV1, binds monoubiquitinated Pol30p/PCNA facilitated by interactions through its BRCT, PAD and UBM domains (34,35). Rev1p/REV1 interacts with pol ζ via Rev7p/REV7 and Pol32p/POLD3 as described earlier (12,18–23). Rev1p/REV1 also interacts with Y-family TLS polymerases (36–38). This structural function of Rev1p/REV1, mediating key protein–protein interactions, is widely appreciated as essential to functional TLS (29–31,39,40). As discussed later, in certain scenarios Rev1p also contributes to TLS through its catalytic activity as a deoxycytidyl transferase (41).

Two models have been proposed for the timing of TLS: it may occur at the replication fork or after replication

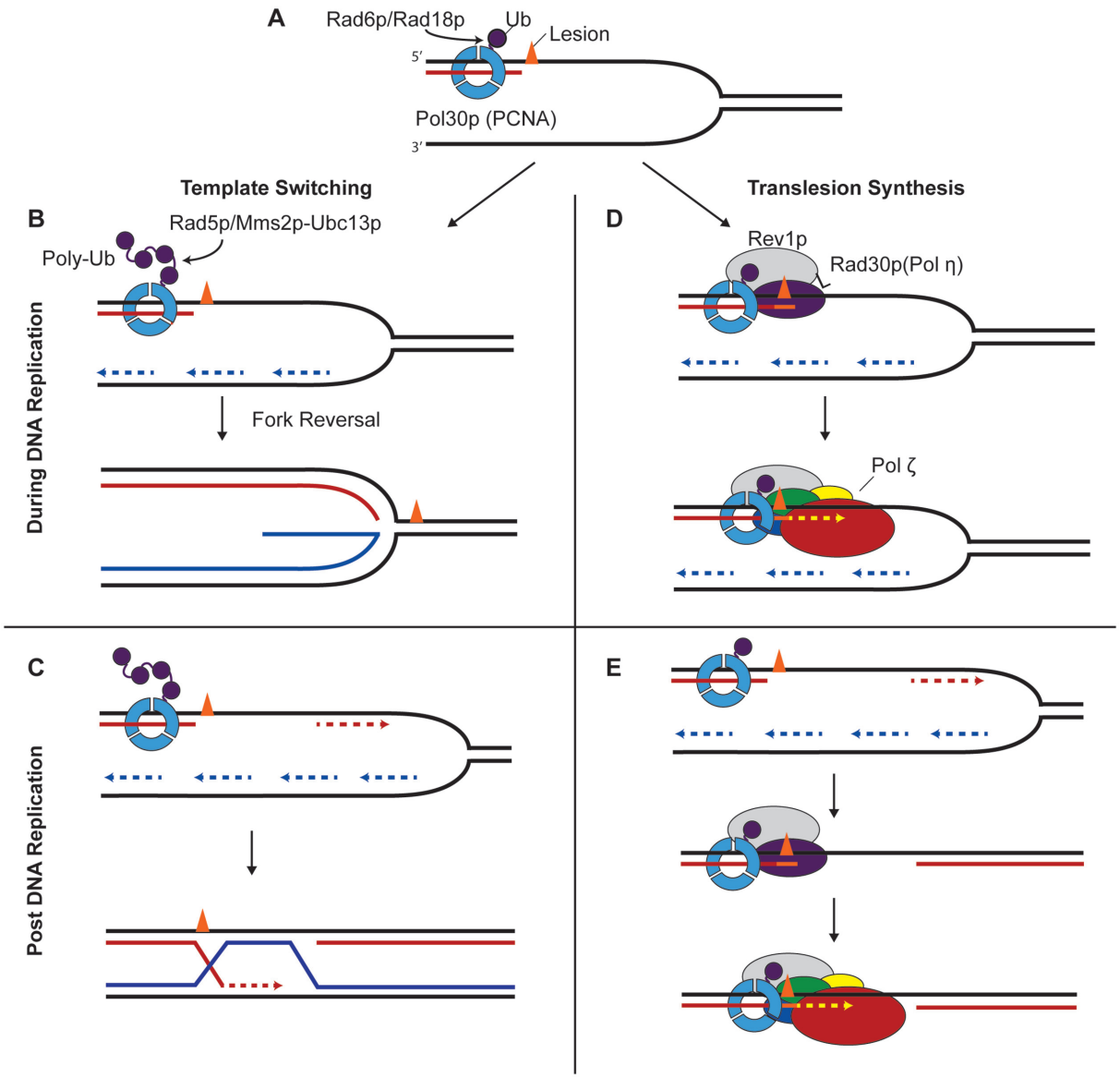


Figure 3. Models for lesion bypass. (**A**) When a replication fork stalls at a lesion, PCNA becomes monoubiquitinated by Rad6p/Rad18p at K164. This is the initial step in both lesion bypass pathways, translesion synthesis (TLS) and template-switching. For both pathways models have been proposed for bypass occurring at the replication fork (co-replication) or after synthesis has reprimed downstream of the lesion (post-replication). Here, bypass is depicted on the leading strand. (**B**, **C**) Template switching is instigated by polyubiquitination of K164, by Mms2p/Ubc13p/Rad5p. This pathway involves switching to the undamaged sister strand as a template for synthesis. This pathway may occur at the fork – one mechanism would be fork reversal as depicted here – or template switching may occur after the replication fork has moved on. (**D**, **E**) Rev1p binds monoubiquitinated Pol30p/PCNA (Ub-PCNA) facilitated by interactions through its BRCT, PAD, and UBM domains. A TLS pol such as pol η is recruited through interactions with Rev1 and Ub-PCNA. After recruitment, pol η (or another TLS pol) can insert a nucleotide opposite the lesion. Next pol ζ is recruited to Ub-PCNA by its interaction with Rev1p, and continues extension past the site of the lesion. In a co-replication model pol ζ would hand synthesis off to a replication polymerase. In the post-replication or 'gap filling' model, pol ζ may fill in a gap either on its own or might hand off to replication polymerases to finish synthesis in the gap.

has continued through repriming past the lesion (Figure 3). These models are not mutually exclusive, and there is indirect evidence for both outcomes (42). In both models, monoubiquitination of Pol30p/PCNA instigates TLS, followed by association of Rev1p/REV1 and recruitment of a TLS polymerase.

Pol ζ has been implicated in direct bypass of several types of lesions as thoroughly reviewed previously (43). Correspondingly, cells with defective pol ζ are sensitive to an array of DNA damaging agents such as UV radiation, cisplatin and aflatoxin, which result in distorting or 'bulky' adducts on DNA (13,44,45). However, depletion of *Rev3l* in human cells does not alter survival or mutagenesis after BPDE treatment, another chemical that induces bulky adducts (46). Intriguingly, REV1, REV3 and REV7 were recruited to sites of DNA-protein crosslinks, implying that pol ζ may serve some function at these lesions (47). Additionally, pol ζ has been linked to interstrand crosslink repair (48,49), although biochemically Y-family polymerases appear to be intrinsically more efficient at bypassing inter-

strand links than pol ζ (50). Thus, pol ζ function at interstrand crosslinks have yet to be fully resolved.

On synthetic templates, Pol ζ can mediate insertion opposite some DNA adducts, as well as subsequent extension. Sometimes lesion bypass may be coordinated by two specialized polymerases. Pol ζ may play a role during lesion bypass by extending the primer-terminus after a nucleotide has been incorporated opposite a lesion by another TLS polymerase (51). For example, pol η and pol ζ have complementary biochemical activities that imply they are cooperating partners for lesion bypass. *In vitro*, pol η efficiently inserts a nucleotide opposite a 1,2 (dGpG) cisplatin intrastrand crosslink, while pol ζ cannot (7). In contrast, pol η starts to falter in extension past a lesion site, while pol ζ can efficiently extend past the lesion (7).

This introduces several important mechanistic questions. Does pol ζ continue significant synthesis after extending from a lesion? If the lesion is bypassed during DNA replication, when does pol ζ hand the reins back to the higher fidelity replicative polymerases? If lesion bypass occurs in a gap left after replication, can pol ζ fill the gap on its own?

Extensive DNA synthesis beyond a lesion by pol ζ

There is evidence that pol ζ is capable of extensive synthesis downstream of UV-induced lesions in yeast. In yeast, Pol ζ-mediated synthesis can be monitored in vivo by mapping *REV3*-dependent mutations. When a plasmid containing a single tetrahydrofuran (a stable AP site mimic) is introduced into BER-deficient yeast cells, sequencing reveals mutations up to 200 bp away from the site of the lesion. These mutations depend on *REV3* function, indicating that pol ζ performs long range synthesis past the site of a lesion (52). In order to study *REV3*-dependent synthesis past the site of a UV radiation-induced photoproduct, reversion of an inactivating point mutation in the *URA3* gene was used. A TC sequence at the reversion site is a target for photoproduct formation by UV radiation. Following UV irradiation, *REV3*-dependent mutations were detected up to 1 kb away from the UV-induced target site. Single-stranded gaps in this size range are observed after UV irradiation in NER-deficient yeast, implying that pol ζ may be responsible for synthesis in a large portion of such gaps (52,53). Synthesis could occur during genome replication or afterwards. In either case, these results suggest that mutagenic DNA synthesis by pol ζ extends significantly beyond a site of UV radiation induced damage.

Unique biochemical characteristics of pol ζ

Pol ζ is the only specialized DNA polymerase in eukaryotes that is a member of the B-family; the other specialized DNA polymerases belong to the A, Y, X and AEP families. The error rate of pol ζ is intermediate between replication and TLS DNA polymerases. It is 10 to 100-fold less accurate than its fellow B-family members in the DNA replication apparatus, but 5 to 30-fold more accurate than Y family TLS polymerases (54). However, much of this analysis was conducted using Rev3p and Rev3p-Rev7p, before pol ζ was discovered to include Pol31p and Pol32p. The relative processivity of pol ζ has yet to be evaluated systematically. Given *in vivo* evidence for pol ζ long-range synthesis, it would be useful to measure its processivity in a single round of DNA synthesis in comparison to Y family polymerases. These unique biochemical features of pol ζ may be relevant not only to its function in lesion bypass, but to more recently revealed pol ζ functions.

Pol ζ synthesis associated with DNA replication stress

In yeast, pol ζ is responsible for a substantial portion (half or more) of 'spontaneous' mutations in cells that are not challenged with external DNA damaging agents (55–57). This implies that cells use pol ζ to copy DNA containing endogenous lesions (such as AP sites) and in response to other stresses. One source of DNA replication stress can be induced through depletion of dNTP pools by incubation with hydroxyurea (HU). Such exposure to HU induces point mutations in yeast, which are largely dependent on *REV3* (58). Additionally, DNA synthesis by pol ζ was not dramatically altered by changes in dNTP levels *in vitro* or *in vivo*, unlike other DNA polymerases (59). Together this data is consistent with a model whereby pol ζ synthesis plays a role in dealing with replication stress, including circumstances where dNTP levels are low enough to hinder chromosomal replication. However, there are multiple pathways for cells to deal with nucleotide depletion since deletion of *REV3* in yeast does not confer enhanced sensitivity to HU (60). Interestingly, knockdown of ribonucleotide reductase M1 (which should result in lower dNTP pools) decreased the viability of *REV3L*-defective human cells more than *REV3L*-proficient controls (61). More research may help clarify the roles mammalian pol ζ may play when cells are under replication stress.

Pol ζ-dependent mutagenesis in yeast also occurs following other types of replication stress. Mutations in replicative polymerases can cripple the replisome and result in elevated mutagenesis. This has been called DRIM or defective-replisome-induced-mutagenesis. Importantly, the vast majority of this mutagenesis is dependent on *REV3* (58,62). Furthermore, the ability of Rev3p to interact with Pol31p–Pol32p is essential for DRIM, implying that pol ζ functions as a four-subunit complex in DRIM (63). Additionally, *REV3*-dependent mutations in DRIM increase when the template switching pathway is inhibited by deletion of *MMS2*. This illustrates the importance of PCNA directing repair choice to either pol ζ synthesis or the 'error-free' template switching when the replisome runs into trouble. Together this work demonstrates that there are multiple routes for cells to overcome various sources of replication stress, and that pol ζ synthesis is an important but mutagenic route.

A further source of replication stress arises when the replisome encounters DNA sequences that are challenging to replicate, such as sequences forming secondary structures (64,65) (Figure 4A). Spontaneous and DRIM pol ζ-dependent mutations often occur at predicted sites of hairpin formation in vivo (66). These same hairpin structures stall pol δ *in vitro*. This highlights that introduction of *REV3*-dependent mutations can be sequence-specific, and could explain variability in the *REV3*-dependent mutation rate when different regions are monitored. DRIM is also dependent on the deoxycytidyl transferase Rev1p. Mutations

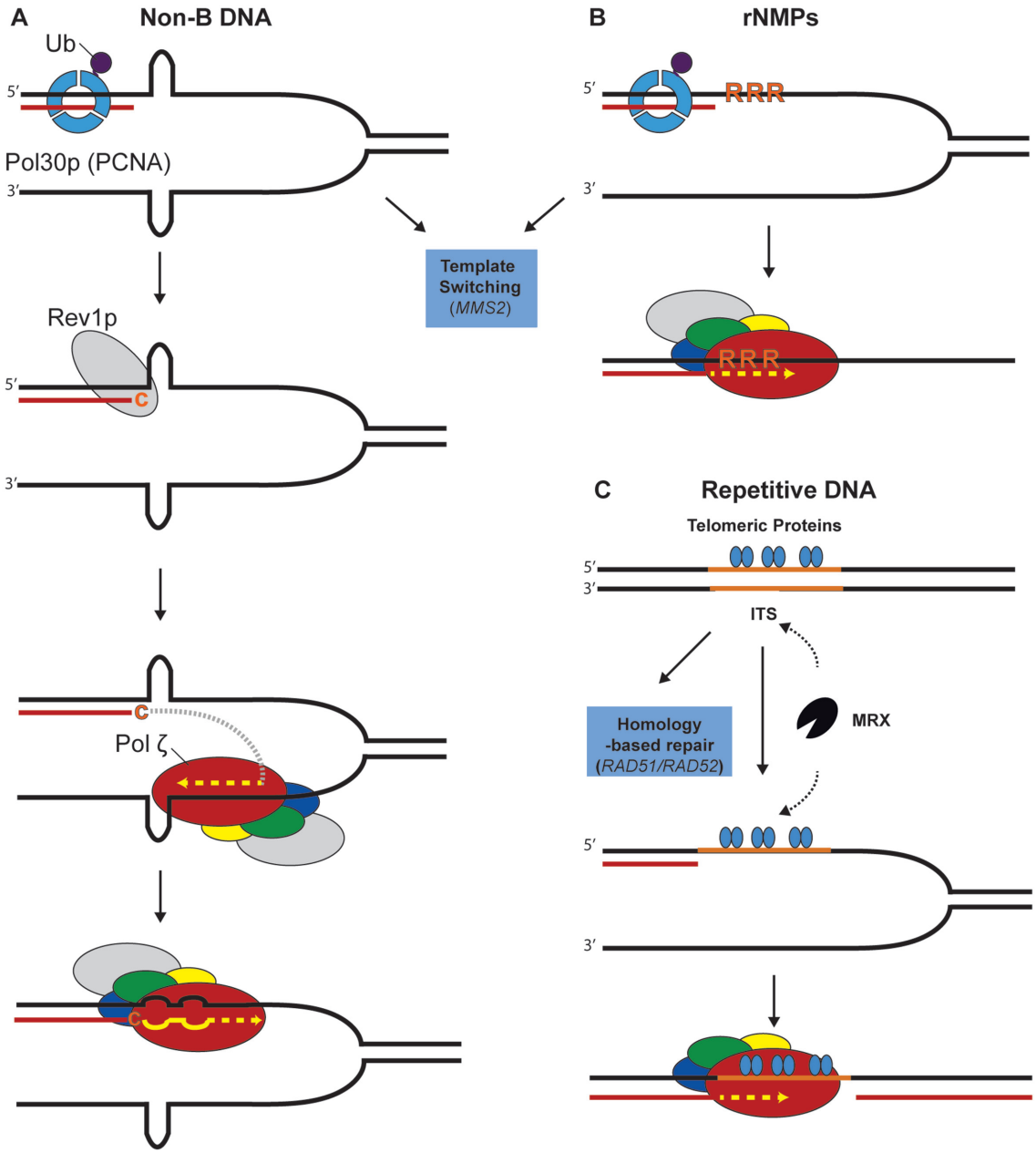


Figure 4. Models for *REV3*-dependent synthesis at difficult-to-replicate regions in the genome. (**A**) Model, after Northam *et al.*, for the role of pol ζ and Rev1p synthesizing non-B DNA (66). Hairpin structures in DNA can stall replicative polymerases. Ubiquitinated PCNA then recruits Rev1p. Rev1p first acts enzymatically and inserts a dCTP near the base of the hairpin, after which it serves to recruit pol ζ. Pol ζ then initiates synthesis on the opposing strand followed by repriming near the original site of the hairpin, copying from this mismatched sequence. Alternatively, hairpin structures can be bypassed by template switching. (**B**) Model for pol ζ bypass of incorporated ribonucleotides proposed by Lazzaro *et al.* (75). Incorporated ribonucleotides can be repaired by ribonucleotide excision repair or a Top1-dependent pathway. When incorporated ribonucleotides remain in the DNA during replication they have the potential to stall a replication fork. In yeast, pol ζ may bypass incorporated ribonucleotides. (**C**) Model for pol ζ replicating interstitial telomeric sequences (ITS), after Moore *et al.* (76). Telomeric proteins, at (ITS), may result in stalling of the replication fork. *REV3* is proposed to fill a gap caused by cleavage of the fork by the MRX endonuclease complex (Mre11p/Rad50p/Xrs2p) followed by end resection. Alternatively, cells can deal with ITS through homology dependent repair mechanisms (dependent on *RAD51* or *RAD52*).

in response to replication perturbations were reduced in a *REV1*-deletion strain as well as in a strain with a deletion in the *REV1* BRCT domain mediating protein-protein interactions. This is consistent with Rev1p having both a structural and enzymatic function in bypass of DNA structures that are difficult to replicate (66).

Bypass of incorporated ribonucleotides

Another major source of replication stress arises from ribonucleotides (rNMPs) erroneously incorporated into the genome during replication (67). Multiple pathways exist to repair and bypass rNMPs, which likely represent the most common error made during replication (68–72). One study has suggested that pol ζ bypass may be one of these pathways.

Most incorporated ribonucleotides are removed from DNA by ribonucleotide excision repair. This process is initiated by RNase H2 (and perhaps RNase H1). An alternative repair mechanism exists that is dependent on topoisomerase I (73,74). However, if unrepaired rNMPs are present during replication they may be bypassed. Incorporated ribonucleotides can also be bypassed by the *MMS2*-dependent template switching pathway discussed above. Yeast cells lacking RNase H1, RNase H2 and Mms2p are only mildly sensitive to HU, but co-deletion of pol ζ components in these cells dramatically sensitizes them to HU (75). This suggests that pol ζ provides one mechanism for cells to survive replication fork stalling induced by collision with rNMPs (Figure 4B).

Providing biochemical support for this model, two-subunit pol ζ (Rev3p-Rev7p) is dramatically more efficient than pol δ at bypassing four tandem rNMPs (75). It will be intriguing to see if the four-subunit pol ζ (Rev3p-Rev7p-Pol31p-Pol32p) can bypass rNMPs even more efficiently than Rev3p-Rev7p alone. In addition, it will be important to determine if this function is conserved for mammalian pol ζ.

Pol ζ and repetitive DNA sequences

Highly repetitive sequences are another impediment to replicative polymerases. An important example are telomere sequence repeats that are found embedded in chromosomes outside of the telomeres, called interstitial telomeric sequences (ITS). When an ITS was placed within an intron of *URA3*, an increased rate of point mutations and terminal sequence inversions was observed in adjacent gene segments (76). The point mutation rate was reduced 10-fold in a *rev3* deletion strain (76). These *REV3*-dependent mutations appeared not to result from a repair of a double strand break, since the mutations were not accompanied by indels in the ITS. The mutations were also dependent on Mre11p, a component of a nuclease that acts on stalled replication forks and facilitate their degradation. This led to the model that *MRE11*-dependent gaps are formed at ITS sequences and are then filled in by pol ζ (Figure 4C). Mutations adjacent to the ITS increased upon disabling homology-based repair by deletion of *RAD51* or *RAD52*, but not after disabling the template switching pathway by deletion of *MMS2*. This suggested that homology-based repair pathways are an alternative repair mechanism at ITS sequences, while canonical template switching is not.

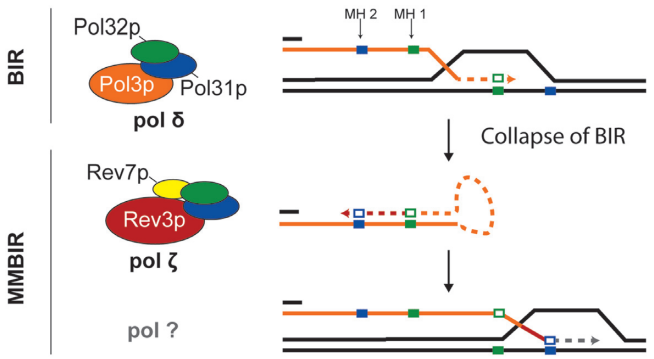


Figure 5. Model for the role of pol ζ in microhomology-mediated break-induced replication. In break-induced replication (BIR) a single-ended DNA break is repaired by extensive synthesis on the sister by pol δ. If the replication fork in BIR collapses, an alternative repair pathway called microhomology-mediated BIR (MMBIR) can be used. A proposed model for MMBIR involves looping of a strand, annealing to itself at a short region of homology (microhomology, MH), with pol ζ carrying out DNA synthesis from this MH-paired primer. The extended strand re-associates with the sister strand, annealing at a microhomology created during synthesis, and continues synthesizes using the sister template. This model was proposed by Sakofsky *et al.* (84).

REV3-dependent mutations in the neighborhood of other types of repetitive sequences have been observed. Mutations in a *URA3* reporter gene were observed as far as 8 kb past an Alu inverted sequence quasi-palindrome; these mutations were largely *REV3*-dependent (77). Mutations also occur in the neighbourhood of trinucleotide repeats (78,79). *REV3*-dependence of mutations in the chromosome adjacent to trinucleotide repeats was most apparent in strains with replisomes carrying crippling *pol3* or *pol2* variants (77,80). While it was demonstrated that double strand breaks do form at these repeats, it is not known whether the pol ζ mutations are dependent on the formation of a break.

It has been known for two decades that *REV3*-dependent mutagenesis occurs in reporter genes located in the neighborhood of a double strand break induced by the HO endonuclease (81,82). Nevertheless, *REV3* is not essential for the repair of the break, leaving the question of the function of pol ζ synthesis around double strand breaks (81,83). Pol ζ is necessary for a specific type of single-ended DNA double strand break processing, termed microhomology-mediated break induced replication (MMBIR, Figure 5) (84). Single-ended double strand breaks can arise from collapsed replication forks or degradation of telomeres. Cells can repair the break through extensive synthesis, involving pol δ, called break-induced replication (BIR, Figure 5). BIR uses the homolog as a template and a primer with a substantial stretch of homology. In yeast BIR has been monitored by the induction of a single ended break at the end of a single chromosome III by HO endonuclease, and tracked with auxotrophic reporters. Deletion of the *PIF1* helicase gene disables extensive synthesis, resulting in collapse of BIR and the appearance of complex mutations near the site of strand invasion (84). These complex mutations can be explained by synthesis arising from alternative templating at microhomologies. These mutations, in *pif1Δ* yeast, are thought

to arise from MMBIR, and such mutations were absent in strains with deletion of *REV3* (84). Point mutations downstream of the single-ended break repaired by MMBIR have not been measured. Such experiments could provide insight into the extent of pol ζ synthesis and its potential function during MMBIR.

An intriguing parallel arises from recent work in human cells, where restart of stalled replication forks in *BRCA2*-deficient cells was dependent on the MUS81 endonuclease and POLD3 (85). A model proposed from this work was that reversed forks (Figure 3B) were being cleaved by MUS81, forming a single-ended DNA break that was repaired with POLD3-dependent synthesis. Whether this repair is mediated by pol δ, pol ζ, or combination of the two remains unresolved.

Growth defects in pol ζ disrupted mammalian cells

Yeast genetics resulted in the discovery of *REV3* over four decades ago, and yeast studies are still at the forefront of unravelling the multifaceted functions of pol ζ as a specialized DNA polymerase. *REV3* mutants of *S. cerevisiae* are viable, without an enhanced formation of gross chromosomal rearrangements (86). In marked contrast to yeast, inactivation of Pol ζ has severe consequences for mammalian cells without exogenous sources of stress. Clarifying this essential function in mammalian systems is paramount to understanding the relevance of pol ζ to humans and cancer.

Without pol ζ function, mammalian genomes accumulate DNA breaks in every cell cycle, with ensuing chromosome breaks, gaps, rearrangements, and formation of micronuclei. In normal cells, this DNA damage impairs viability by initiating gene expression programs that lead to cell death or senescence. For example, primary mouse embryonic fibroblasts (MEFs) die or senesce within two cell divisions after disruption of *Rev3l*. This excessive cell death and premature senescence in primary MEFs can be prevented by immortalizing cells with T-antigen before disruption of *Rev3l*. Although cells can proliferate because immortalization blunts the implementation of cell death pathways, the *Rev3l* defect leads to a 10-fold increase in the frequency of DNA double-strand breaks and chromosome rearrangements (6,87). Similarly, some human cancer cell lines such as NALM-6 can propagate after inactivation of *REV3L*, but with increased sensitivity to many DNA damaging agents (88). A *REV3l*$^{-/-}$ mutant was also derived from TK6, a p53$^{+/+}$ EBV-immortalized lymphoblastoid cell line (89). Presumably such cell lines are also compromised in checkpoints or cell death pathways. An interesting parallel arises from studies of yeast strains with defective telomere maintenance. Such strains with a *rev3* deletion cannot proliferate and undergo senescence (90). One proposal is that short inverted repeat-induced synthesis is used to maintain chromosome ends, taking advantage of palindrome sequences (90). This process, related to MMBIR, may require DNA synthesis mediated by pol ζ.

Mouse embryos with a homozygous disruption of *Rev3l* do not develop further than mid-gestation. Some strains lose viability even earlier, and so several investigations have turned to conditional disruption models. A broad inactivation of *Rev3l* in hematopoietic progenitors does not give rise to propagating B or T cells in adult mice (91). Disruption is better tolerated in mature resting B cells expressing CD19 or CD21 (92,93).

Adult mice are able to survive after ablation of *Rev3l* in epithelial keratinocytes (using Cre recombinase driven by a keratin-5 promoter) (87). However, these animals display major skin abnormalities, including reduced epidermal cellularity, irregular hair cycling, and gradual hair loss (87). Keratinocytes from the mice exhibit highly enhanced baseline levels of DNA damage stress and do not survive if explanted into cell culture. Remarkably, these mice are the most sensitive to UV radiation of any known single mutant mouse model, much more sensitive than those with defects in nucleotide excision repair (87).

The greatly enhanced UV radiation sensitivity of mice with *Rev3l*-defective keratinocytes arises from two sources. First, the cells are defective in TLS of UV radiation-induced DNA damage. Second, the pol ζ defect impairs the proliferation of normal keratinocytes as the skin attempts to recover from UV-irradiation. To verify that pol ζ is necessary for continued proliferation of primary cells in the mouse, even in the absence of exogenous DNA damage, wound healing was examined. This process depends on rapid proliferation of cells to regenerate the wounded area. *Rev3l*-defective epidermis was found to have a proliferative defect in wound healing (94). Pronounced epidermal skin pigmentation occurred following UV irradiation or wound healing in *Rev3l*-defective adult mice (87,94). This is striking because, unlike humans, adult mice do not normally have melanocytes resident in the basal epidermis. The explanation is related to the DNA damage-dependent stress response developed by *Rev3l*-defective keratinocytes as they attempt to proliferate. Stress signals from the keratinocytes then induce p53-dependent migration of melanocytes to the epidermis (94). In keratinocytes, p53 directly upregulates the transcription of genes that stimulate production of eumelanin and increase pigmentation (95–98).

Pol ζ as a tumor suppressor

One consequence of the genome protective function of pol ζ is that *Rev3l* is a suppressor of spontaneous tumor formation. There are several examples. In *Tp53*-defective mice, a mosaic deletion of *Rev3l* in the lymphocyte lineage caused an increased incidence and reduced latency of lymphomas (91). Conditional partial disruption of *Rev3l* in the mammary gland enhances the incidence of mammary gland tumorigenesis (91). About 90% of mice with disruption of *Rev3l* in keratin-5 expressing cells acquired squamous cell carcinomas after 1 year (87). It is likely that the increased frequency of tumors is due to oncogenic chromosome rearrangements in *Rev3l*-defective tissues. For example, cancers arising from *Tp53*-defective and *BRCA1* defective mouse epithelia are driven by amplifications or translocations that result in elevated oncoprotein expression or oncoprotein-containing fusions (99).

The polymerase activity of pol ζ is essential for limiting genome damage

One question is whether the essential functions of REV3L depend directly on the DNA polymerase, or whether only

the structural presence of REV3L is needed. A catalytically deficient REV3L might still interact with its protein partners and other DNA substrates, allowing viability of cells and mice. To determine whether ablation of the catalytic activity of pol ζ is responsible for the severe consequences in pol ζ mutant mouse cells, mice were engineered with specific inactivation of the catalytic activity. Conserved active site aspartate residues necessary for activity (D2773, D2775) were changed to alanine, in two independent studies. In both cases, no viable homozygous mice were produced, and the corresponding embryos died early in embryogenesis (6,100). To investigate cell-autonomous consequences of the specific polymerase alteration, primary MEFs were derived that carried one null *Rev3l* allele, and one active site mutant allele (6). The growth defects, DNA break formation and cisplatin sensitivity of these cells were similar to cells harboring two null alleles (101). These results showed that the DNA polymerase activity of REV3L is essential for the genome protective functions.

Different consequences of *REV7* and *REV3L* knockouts in mammals.

As we noted earlier, REV7 (encoded by *MAD2L2*) functions as a component of pol ζ. A dimer of REV7 binds to two adjacent sites in REV3L by grasping a peptide of REV3L with a 'safety-belt' loop (13,14). REV7 serves as a bridge between REV3L and the REV1 protein (Figure 1B). The REV7 protein is, however, at least an order of magnitude more abundant than REV3L (13) and has additional functions and numerous protein partners. These include chromatin-associated proteins and post-translational modification proteins (102–108). A further function of REV7 has been established as a DNA resection inhibitor in the 'Shieldin' complex (109,110). In view of these many functions, it is fascinating that complete ablation does not always seem to block viability of mice. *Rev7* disruption in mice (using embryonic stem cells derived from 129 strains) gave homozygous progeny in a sub-mendelian ratio, with a smaller overall size and pronounced defects in primary germ cells (105,111). Germline deletion of *Rev7* in C57BL/6 embryonic stem cells resulted in embryonic lethality (112). Specific deletion of Rev7 from the B cell lineage using *Mb1-Cre* yielded normally functioning B cells, unlike the severe defects in *Rev3l*-defective B cells (112). These considerations suggest that *Rev3l* has as yet unappreciated functions that are independent of Rev7, and perhaps independent of TLS. Studies that elucidate the broad consequences of *Rev3l* deletion, such as profiling the changes in the transcriptome and proteome, may help uncover novel functions of pol ζ.

Cancer and disease-associated mutations in *REV3L*

REV3L is a large gene and therefore many *REV3L* mutations are found in the genomes of human cancers. *REV3L* is mutated in more than 10% of the NCI-60 cancer cell line set (113). Analysis via cBioportal of cancers with both mutation and copy number data currently shows that *REV3L* is altered in ∼2% of cancers and that ∼18% of these mutations occur with loss of heterozygosity (114,115). Heterozygous mutations in the human *REV3L* gene have also been suggested to be associated with an inherited neurological disorder, Möbius syndrome (116). It will be interesting to investigate whether the association can be confirmed in further studies of mouse models and the human disorder.

Inhibition of pol ζ as a potential therapeutic target

Genomic instability fuels cancer development and also provides a reservoir of genomic variations that facilitate resistance to therapies. Tumor cells often have altered DNA repair pathways, resulting in genomic instability, one of the defining hallmarks of cancer (117). However, this can lead tumor cells to rely more heavily on remaining intact repair pathways to prevent genomic damage from reaching levels that lead to unescapable cell death. As a result, targeted therapies have emerged that exploit specific DNA repair deficiencies in tumor cells, the most prominent being the rise of PARP inhibitors in the clinic (118).

A potentially exciting prospect for cancer therapy is to inhibit the function of TLS in tumor cells to block a remaining route of survival of cancer cells challenged with DNA damaging chemotherapeutics. Depleting pol ζ function should increase sensitivity of a tumor to some DNA damaging drugs. Concomitantly, reduction of the TLS function would be expected to reduce drug-induced base substitution and frameshift mutagenesis. This could advantageously reduce the pool of drug-induced mutated tumor cells, which could slow down recurrence of a tumor.

This strategy has been explored in mouse tumor models. Mouse lung adenocarcinoma cells were transplanted into syngeneic mice, and the animals exposed to cisplatin (119). Cisplatin toxicity was higher in *shRev3l*-depleted cells, and mice receiving tumors harboring *shRev3l* had a longer survival time following cisplatin treatment (22.5 days) compared to control tumors with normal Rev3l levels (11 days). The mutation frequency of cells expressing *shRev3l* was ∼20% of normal following cisplatin treatment. Other experiments used a model of Burkitt lymphoma, the *Eμ-myc* mouse (120). Pure populations of control and *Rev3l*-deficient lymphoma cells were introduced into recipient mice until palpable tumors formed (∼2 weeks). Following treatment with a single 10 mg/kg dose of cisplatin, *Rev3l*-deficient tumors exhibited a more rapid reduction in size than controls, with tumor regression occurring within 24 h.

Other experiments focused on REV1 (120). To examine whether *Rev1* deficiency similarly could delay the development of chemoresistant tumors in vivo, control and *Rev1*-deficient *Eμ-myc* lymphoma cells were injected into syngeneic recipient mice. After tumors formed, mice were treated with cyclophosphamide. To examine the role of *Rev1* in the evolution of chemoresistance, tumors were harvested from individual mice at relapse and reinjected into mice for a second round of therapy. Even after three cycles of engraftment, treatment and regrowth, three of four *Rev1*-deficient tumors retained a pronounced sensitivity to CTX treatment, whereas all control recipients succumbed to their tumors. These data suggest that TLS inhibition may have dual anticancer effects, sensitizing tumors to therapy as well as preventing the emergence of chemoresistance.

Because pol ζ inactivation is toxic to normal cells, a major practical challenge is the development and delivery

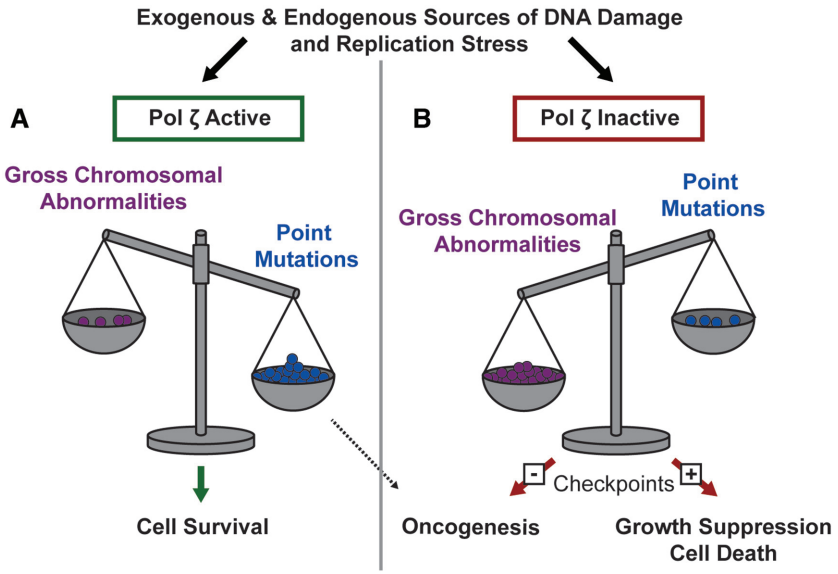


Figure 6. Pol ζ operates as a tumor suppressor by preserving chromosomal stability at the cost of point mutations in mammals. (**A**) In cells with proficient pol ζ, endogenous and exogenous sources of stress or damage require pol ζ synthesis to prevent gross chromosomal abnormalities, but at the cost of introducing point mutations. (**B**) Without pol ζ, cells may accrue fewer point mutations but also acquire increased chromosomal abnormalities when challenged with exogenous or endogenous sources of damage or stress. These chromosomal abnormalities lead to growth suppression and cell death in cells with functional checkpoints. Cells without functional checkpoints may be able to survive and the chromosomal abnormalities can drive oncogenesis.

of a pol ζ inhibitor to tumor cells. In one approach, a nanoparticle was used to deliver a cisplatin prodrug and REV1/ REV3L-specific siRNAs simultaneously. Administering such nanoparticles to a prostate cancer cell mouse model gave a synergistic inhibition of tumor growth (121).

Another approach may be to target small-molecule inhibitors to the pol ζ complex (122). A further strategy is to inhibit specific interactions between proteins in the extended pol ζ complex. Protein-protein interactions between the C-terminal domain of REV1 and the REV1-interacting region of other TLS DNA polymerases play an essential role in TLS. A fluorescence polarization-based assay was used in a pilot screen for small molecule inhibitors of this PPI (123). Small molecule scaffolds that disrupt this interaction were identified. Survival and mutagenesis assays in mammalian cells exposed to cisplatin or UV radiation indicated that these compounds inhibit mutagenic REV1/pol ζ-dependent TLS in cells (123).

There are many other protein–protein interactions within the pol ζ complex (20) (Figure 1). Some of these interactions are candidates for small-molecule targeting. Specific inhibitors of the catalytic activity of REV3L could also inhibit pol ζ function (6,100). Recently an inhibitor that prevents the REV7-REV1 interaction through promoting dimerization of REV1 was identified (39). Importantly, when combined with cisplatin, this inhibitor reduced tumor size and increased survival in a mouse xenograft model. An important caveat is that disruption of *Rev3l* function causes cells to be susceptible to increased chromosome breaks, gaps, and oncogenic chromosome rearrangements (Figure 6). Crippling of REV1 function (in a nucleotide excision repair-defective background) enhances UV radiation-induced carcinogenesis by elevating inflammatory hyperplasia (124). Therefore, pol ζ inhibition could sometimes contribute to generating drug-resistant variants in recurrent cancer, and such possibilities will have to be monitored carefully.

CONCLUSION

Pol ζ is a specialized DNA polymerase that functions in canonical translesion synthesis but also has broader roles in unchallenged replicating cells. Uncovering the roles pol ζ plays in preserving genomic stability in mammalian cells could have broad reaching implications in understanding carcinogenesis and in developing combination therapies for some cancers.

ACKNOWLEDGEMENTS

We appreciate helpful comments and edits from Dr Sylvie Doublié (University of Vermont), Dr Denisse Carvajal-Maldonado, Megan Lowery and Sarita Bhetawal. We thank Dr Junya Tomida for advice and guidance.

FUNDING

Studies of pol ζ in our laboratory are funded by National Institutes of Health [CA193124]; Department of Defense [W81XWH-17-10239]; Grady F. Saunders Ph.D. Distinguished Research Professorship (to R.D.W.); S.K.M. was supported by a Research Training Grant award from the Cancer Prevention and Research Institute of Texas [RP170067]. The open access publication charge for this paper has been waived by Oxford University Press – *NAR*.

Conflict of interest statement. None declared.

REFERENCES

1. Gibbs,P.E., McGregor,W.G., Maher,V.M., Nisson,P. and Lawrence,C.W. (1998) A human homolog of the Saccharomyces cerevisiae REV3 gene, which encodes the catalytic subunit of DNA polymerase ζ. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 6876–6880.
2. Gomez-Llorente,Y., Malik,R., Jain,R., Choudhury,J.R., Johnson,R.E., Prakash,L., Prakash,S., Ubarretxena-Belandia,I. and Aggarwal,A.K. (2013) The architecture of yeast DNA polymerase ζ. *Cell Rep.*, **5**, 79–86.
3. Swan,M.K., Johnson,R.E., Prakash,L., Prakash,S. and Aggarwal,A.K. (2009) Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase δ. *Nat. Struct. Mol. Biol.*, **16**, 979–986.
4. Hogg,M., Osterman,P., Bylund,G.O., Ganai,R.A., Lundstrom,E.B., Sauer-Eriksson,A.E. and Johansson,E. (2014) Structural basis for processive DNA synthesis by yeast DNA polymerase ϵ. *Nat. Struct. Mol. Biol.*, **21**, 49–55.
5. Coloma,J., Johnson,R.E., Prakash,L., Prakash,S. and Aggarwal,A.K. (2016) Human DNA polymerase α in binary complex with a DNA:DNA template-primer. *Sci. Rep.*, **6**, 23784.
6. Lange,S.S., Tomida,J., Boulware,K.S., Bhetawal,S. and Wood,R.D. (2016) The Polymerase Activity of Mammalian DNA Pol ζ Is Specifically Required for Cell and Embryonic Viability. *PLos Genet.*, **12**, e1005759.
7. Lee,Y.S., Gregory,M.T. and Yang,W. (2014) Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 2954–2959.
8. Netz,D.J., Stith,C.M., Stumpfig,M., Kopf,G., Vogel,D., Genau,H.M., Stodola,J.L., Lill,R., Burgers,P.M. and Pierik,A.J. (2012) Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. *Nat. Chem. Biol.*, **8**, 125–132.
9. Baranovskiy,A.G., Lada,A.G., Siebler,H., Zhang,Y., Pavlov,Y.I. and Tahirov,T.H. (2012) DNA polymerases δ and ζ switching by sharing the accessory subunits of DNA polymerase δ. *J. Biol. Chem.*, **287**, 17281–17287.
10. Johnson,R.E., Prakash,L. and Prakash,S. (2012) Pol31 and Pol32 subunits of yeast DNA polymerase δ are also essential subunits of DNA polymerase ζ. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 12455–12460.
11. Murakumo,Y., Roth,T., Ishii,H., Rasio,D., Numata,S., Croce,C.M. and Fishel,R. (2000) A human REV7 homolog that interacts with the polymerase ζ catalytic subunit hREV3 and the spindle assembly checkpoint protein hMAD2. *J. Biol. Chem.*, **275**, 4391–4397.
12. Hanafusa,T., Habu,T., Tomida,J., Ohashi,E., Murakumo,Y. and Ohmori,H. (2010) Overlapping in short motif sequences for binding to human REV7 and MAD2 proteins. *Genes Cells*, **15**, 281–296.
13. Tomida,J., Takata,K., Lange,S.S., Schibler,A.C., Yousefzadeh,M.J., Bhetawal,S., Dent,S.Y. and Wood,R.D. (2015) REV7 is essential for DNA damage tolerance via two REV3L binding sites in mammalian DNA polymerase ζ. *Nucleic Acids Res.*, **43**, 1000–1011.
14. Hara,K., Hashimoto,H., Murakumo,Y., Kobayashi,S., Kogame,T., Unzai,S., Akashi,S., Takeda,S., Shimizu,T. and Sato,M. (2010) Crystal structure of human REV7 in complex with a human REV3 fragment and structural implication of the interaction between DNA polymerase ζ and REV1. *J. Biol. Chem.*, **285**, 12299–12307.
15. Rizzo,A.A., Vassel,F.M., Chatterjee,N., D'Souza,S., Li,Y., Hao,B., Hemann,M.T., Walker,G.C. and Korzhnev,D.M. (2018) Rev7 dimerization is important for assembly and function of the Rev1/Polζ translesion synthesis complex. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, e8191–e8200.
16. Murakumo,Y., Ogura,Y., Ishii,H., Numata,S., Ichihara,M., Croce,C.M., Fishel,R. and Takahashi,M. (2001) Interactions in the error-prone postreplication repair proteins hREV1, hREV3, and hREV7. *J. Biol. Chem.*, **276**, 35644–35651.
17. Nelson,J.R., Lawrence,C.W. and Hinkle,D.C. (1996) Thymine-thymine dimer bypass by yeast DNA polymerase ζ. *Science*, **272**, 1646–1649.
18. Kikuchi,S., Hara,K., Shimizu,T., Sato,M. and Hashimoto,H. (2012) Structural basis of recruitment of DNA polymerase ζ by interaction between REV1 and REV7 proteins. *J. Biol. Chem.*, **287**, 33847–33852.
19. Wojtaszek,J., Liu,J., D'Souza,S., Wang,S., Xue,Y., Walker,G.C. and Zhou,P. (2012) Multifaceted recognition of vertebrate Rev1 by translesion polymerases ζ and κ. *J. Biol. Chem.*, **287**, 26400–26408.
20. Pustovalova,Y., Magalhaes,M.T., D'Souza,S., Rizzo,A.A., Korza,G., Walker,G.C. and Korzhnev,D.M. (2016) Interaction between the Rev1 C-Terminal domain and the PolD3 subunit of Polζ suggests a mechanism of polymerase exchange upon Rev1/Polζ-Dependent translesion synthesis. *Biochemistry*, **55**, 2043–2053.
21. Acharya,N., Johnson,R.E., Pages,V., Prakash,L. and Prakash,S. (2009) Yeast Rev1 protein promotes complex formation of DNA polymerase ζ with Pol32 subunit of DNA polymerase δ. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 9631–9636.
22. Acharya,N., Haracska,L., Johnson,R.E., Unk,I., Prakash,S. and Prakash,L. (2005) Complex formation of yeast Rev1 and Rev7 proteins: a novel role for the polymerase-associated domain. *Mol. Cell Biol.*, **25**, 9734–9740.
23. D'Souza,S. and Walker,G.C. (2006) Novel role for the C terminus of *Saccharomyces cerevisiae* Rev1 in mediating protein-protein interactions. *Mol. Cell Biol.*, **26**, 8173–8182.
24. Morelli,C., Mungall,A.J., Negrini,M., Barbanti-Brodano,G. and Croce,C.M. (1998) Alternative splicing, genomic structure, and fine chromosome localization of REV3L. *Cytogenet. Cell Genet.*, **83**, 18–20.
25. Kawamura,K., O-Wang,J, Bahar,R., Koshikawa,N., Shishikura,T., Nakagawara,A., Sakiyama,S., Kajiwara,K., Kimura,M. and Tagawa,M. (2001) The error-prone DNA polymerase ζ catalytic subunit (Rev3) gene is ubiquitously expressed in normal and malignant human tissues. *Int. J. Oncol.*, **18**, 97–103.
26. Gouw,M., Michael,S., Samano-Sanchez,H., Kumar,M., Zeke,A., Lang,B., Bely,B., Chemes,L.B., Davey,N.E., Deng,Z. *et al.* (2018) The eukaryotic linear motif resource - 2018 update. *Nucleic Acids Res.*, **46**, D428–D434.
27. Singh,B., Li,X., Owens,K.M., Vanniarajan,A., Liang,P. and Singh,K.K. (2015) Human REV3 DNA polymerase ζ localizes to mitochondria and protects the mitochondrial genome. *PLoS One*, **10**, e0140409.
28. Zhang,H., Chatterjee,A. and Singh,K.K. (2006) *Saccharomyces cerevisiae* polymerase ζ functions in mitochondria. *Genetics*, **172**, 2683–2688.
29. Slade,D. (2018) Maneuvers on PCNA Rings during DNA Replication and Repair. *Genes (Basel)*, **9**, E416.
30. Leung,W., Baxley,R.M., Moldovan,G.L. and Bielinsky,A.K. (2018) Mechanisms of DNA damage Tolerance: Post-Translational regulation of PCNA. *Genes (Basel)*, **10**, E10.
31. Kanao,R. and Masutani,C. (2017) Regulation of DNA damage tolerance in mammalian cells by post-translational modifications of PCNA. *Mutat. Res.*, **803–805**, 82–88.
32. Yang,W. (2014) An overview of Y-Family DNA polymerases and a case study of human DNA polymerase η. *Biochemistry*, **53**, 2793–2803.
33. Dieckman,L.M. and Washington,M.T. (2013) PCNA trimer instability inhibits translesion synthesis by DNA polymerase η and by DNA polymerase δ. *DNA Repair (Amst.)*, **12**, 367–376.
34. Guo,C., Sonoda,E., Tang,T.S., Parker,J.L., Bielen,A.B., Takeda,S., Ulrich,H.D. and Friedberg,E.C. (2006) REV1 protein interacts with PCNA: significance of the REV1 BRCT domain in vitro and in vivo. *Mol. Cell*, **23**, 265–271.
35. Sharma,N.M., Kochenova,O.V. and Shcherbakova,P.V. (2011) The non-canonical protein binding site at the monomer-monomer interface of yeast proliferating cell nuclear antigen (PCNA) regulates the Rev1-PCNA interaction and Polzeta/Rev1-dependent translesion DNA synthesis. *J. Biol. Chem.*, **286**, 33557–33566.
36. Guo,C., Fischhaber,P.L., Luk-Paszyc,M.J., Masuda,Y., Zhou,J., Kamiya,K., Kisker,C. and Friedberg,E.C. (2003) Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. *EMBO J.*, **22**, 6621–6630.
37. Tissier,A., Kannouche,P., Reck,M.P., Lehmann,A.R., Fuchs,R.P. and Cordonnier,A. (2004) Co-localization in replication foci and interaction of human Y-family members, DNA polymerase pol η and REV1 protein. *DNA Repair (Amst.)*, **3**, 1503–1514.
38. Ohashi,E., Murakumo,Y., Kanjo,N., Akagi,J., Masutani,C., Hanaoka,F. and Ohmori,H. (2004) Interaction of hREV1 with three human Y-family DNA polymerases. *Genes Cells*, **9**, 523–531.

39. Wojtaszek,J.L., Chatterjee,N., Najeeb,J., Ramos,A., Lee,M., Bian,K., Xue,J.Y., Fenton,B.A., Park,H., Li,D. *et al.* (2019) A small molecule targeting mutagenic translesion synthesis improves chemotherapy. *Cell*, **178**, 152–159.
40. Simpson,L.J. and Sale,J.E. (2003) Rev1 is essential for DNA damage tolerance and non-templated immunoglobulin gene mutation in a vertebrate cell line. *EMBO J.*, **22**, 1654–1664.
41. Kim,N., Mudrak,S.V. and Jinks-Robertson,S. (2011) The dCMP transferase activity of yeast Rev1 is biologically relevant during the bypass of endogenously generated AP sites. *DNA Repair (Amst.)*, **10**, 1262–1271.
42. Waters,L.S., Minesinger,B.K., Wiltrout,M.E., D'Souza,S., Woodruff,R.V. and Walker,G.C. (2009) Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. *Microbiol. Mol. Biol. Rev.*, **73**, 134–154.
43. Makarova,A.V. and Burgers,P.M. (2015) Eukaryotic DNA polymerase ζ. *DNA Repair (Amst.)*, **29**, 47–55.
44. Lin,Y.C., Owen,N., Minko,I.G., Lange,S.S., Li,L., Stone,M.P., Wood,R.D., McCullough,A.K. and Lloyd,R.S. (2016) DNA polymerase zeta limits chromosomal damage and promotes cell survival following aflatoxin exposure. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 13774–13779.
45. Li,Z., Zhang,H., McManus,T.P., McCormick,J.J., Lawrence,C.W. and Maher,V.M. (2002) hREV3 is essential for error-prone translesion synthesis past UV or benzo[a]pyrene diol epoxide-induced DNA lesions in human fibroblasts. *Mutat. Res.*, **510**, 71–80.
46. Li,Z., Zhang,H., McManus,T.P., McCormick,J.J., Neal,J.A., Lawrence,C.W. and Maher,V.M. (2009) Erratum to: "hREV3 is essential for error-prone translesion synthesis past UV or benzo[a]pyrene diol epoxide-induced DNA lesions in human fibroblasts". *Mutat. Res.*, **670**, 107–108.
47. Larsen,N.B., Gao,A.O., Sparks,J.L., Gallina,I., Wu,R.A., Mann,M., Raschle,M., Walter,J.C. and Duxin,J.P. (2019) Replication-Coupled DNA-Protein crosslink repair by SPRTN and the proteasome in xenopus Egg extracts. *Mol. Cell*, **73**, 574–588.
48. Shen,X., Jun,S., O'Neal,L.E., Sonoda,E., Bemark,M., Sale,J.E. and Li,L. (2006) REV3 and REV1 play major roles in recombination-independent repair of DNA interstrand cross-links mediated by monoubiquitinated proliferating cell nuclear antigen (PCNA). *J. Biol. Chem.*, **281**, 13869–13872.
49. Raschle,M., Knipscheer,P., Enoiu,M., Angelov,T., Sun,J., Griffith,J.D., Ellenberger,T.E., Scharer,O.D. and Walter,J.C. (2008) Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell*, **134**, 969–980.
50. Ho,T.V., Guainazzi,A., Derkunt,S.B., Enoiu,M. and Scharer,O.D. (2011) Structure-dependent bypass of DNA interstrand crosslinks by translesion synthesis polymerases. *Nucleic Acids Res.*, **39**, 7455–7464.
51. Johnson,R.E., Washington,M.T., Haracska,L., Prakash,S. and Prakash,L. (2000) Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature*, **406**, 1015–1019.
52. Kochenova,O.V., Daee,D.L., Mertz,T.M. and Shcherbakova,P.V. (2015) DNA polymerase zeta-dependent lesion bypass in *Saccharomyces cerevisiae* is accompanied by error-prone copying of long stretches of adjacent DNA. *PLos Genet.*, **11**, e1005110.
53. Lopes,M., Foiani,M. and Sogo,J.M. (2006) Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Mol. Cell*, **21**, 15–27.
54. Zhong,X., Garg,P., Stith,C.M., Nick McElhinny,S.A., Kissling,G.E., Burgers,P.M. and Kunkel,T.A. (2006) The fidelity of DNA synthesis by yeast DNA polymerase zeta alone and with accessory proteins. *Nucleic Acids Res.*, **34**, 4731–4742.
55. Makarova,A.V., Stodola,J.L. and Burgers,P.M. (2012) A four-subunit DNA polymerase ζ complex containing Pol δ accessory subunits is essential for PCNA-mediated mutagenesis. *Nucleic Acids Res.*, **40**, 11618–11626.
56. Roche,H., Gietz,R.D. and Kunz,B.A. (1994) Specificity of the yeast *rev3Δ*; antimutator and *REV3* dependency of the mutator resulting from a defect (*rad1Δ*;) in nucleotide excision repair. *Genetics*, **137**, 637–646.
57. Quah,S.K., von Borstel,R.C. and Hastings,P.J. (1980) The origin of spontaneous mutation in *Saccharomyces cerevisiae*. *Genetics*, **96**, 819–839.
58. Northam,M.R., Robinson,H.A., Kochenova,O.V. and Shcherbakova,P.V. (2010) Participation of DNA polymerase zeta in replication of undamaged DNA in *Saccharomyces cerevisiae*. *Genetics*, **184**, 27–42.
59. Kochenova,O.V., Bezalel-Buch,R., Tran,P., Makarova,A.V., Chabes,A., Burgers,P.M. and Shcherbakova,P.V. (2017) Yeast DNA polymerase ζ maintains consistent activity and mutagenicity across a wide range of physiological dNTP concentrations. *Nucleic Acids Res.*, **45**, 1200–1218.
60. Gallo,D., Kim,T., Szakal,B., Saayman,X., Narula,A., Park,Y., Branzei,D., Zhang,Z. and Brown,G.W. (2019) Rad5 recruits Error-Prone DNA polymerases for mutagenic repair of ssDNA gaps on undamaged templates. *Mol. Cell*, **73**, 900–914.
61. Kotov,I.N., Siebring-van Olst,E., Knobel,P.A., van der Meulen-Muileman,I.H., Felley-Bosco,E., van Beusechem,V.W., Smit,E.F., Stahel,R.A. and Marti,T.M. (2014) Whole genome RNAi screens reveal a critical role of REV3 in coping with replication stress. *Mol. Oncol.*, **8**, 1747–1759.
62. Northam,M.R., Garg,P., Baitin,D.M., Burgers,P.M. and Shcherbakova,P.V. (2006) A novel function of DNA polymerase zeta regulated by PCNA. *EMBO J.*, **25**, 4316–4325.
63. Szwajczak,E., Fijalkowska,I.J. and Suski,C. (2017) The CysB motif of Rev3p involved in the formation of the four-subunit DNA polymerase ζ is required for defective-replisome-induced mutagenesis. *Mol. Microbiol.*, **106**, 659–672.
64. Lu,S., Wang,G., Bacolla,A., Zhao,J., Spitser,S. and Vasquez,K.M. (2015) Short inverted repeats are hotspots for genetic Instability: Relevance to cancer genomes. *Cell Rep.*, **10**, 1674–1680.
65. Cooper,D.N., Bacolla,A., Ferec,C., Vasquez,K.M., Kehrer-Sawatzki,H. and Chen,J.M. (2011) On the sequence-directed nature of human gene mutation: the role of genomic architecture and the local DNA sequence environment in mediating gene mutations underlying human inherited disease. *Hum. Mutat.*, **32**, 1075–1099.
66. Northam,M.R., Moore,E.A., Mertz,T.M., Binz,S.K., Stith,C.M., Stepchenkova,E.I., Wendt,K.L., Burgers,P.M.J. and Shcherbakova,P.V. (2014) DNA polymerases ζ and Rev1 mediate error-prone bypass of non-B DNA structures. *Nucleic Acids Res.*, **42**, 290–306.
67. Williams,J.S., Lujan,S.A. and Kunkel,T.A. (2016) Processing ribonucleotides incorporated during eukaryotic DNA replication. *Nat. Rev. Mol. Cell Biol.*, **17**, 350–363.
68. Nick McElhinny,S.A., Watts,B.E., Kumar,D., Watt,D.L., Lundstrom,E.B., Burgers,P.M., Johansson,E., Chabes,A. and Kunkel,T.A. (2010) Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 4949–4954.
69. Nick McElhinny,S.A., Kumar,D., Clark,A.B., Watt,D.L., Watts,B.E., Lundstrom,E.B., Johansson,E., Chabes,A. and Kunkel,T.A. (2010) Genome instability due to ribonucleotide incorporation into DNA. *Nat. Chem. Biol.*, **6**, 774–781.
70. Reijns,M.A., Rabe,B., Rigby,R.E., Mill,P., Astell,K.R., Lettice,L.A., Boyle,S., Leitch,A., Keighren,M., Kilanowski,F. *et al.* (2012) Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell*, **149**, 1008–1022.
71. Caldecott,K.W. (2014) Molecular biology. Ribose–an internal threat to DNA. *Science*, **343**, 260–261.
72. Clausen,A.R., Zhang,S., Burgers,P.M., Lee,M.Y. and Kunkel,T.A. (2013) Ribonucleotide incorporation, proofreading and bypass by human DNA polymerase δ. *DNA Repair (Amst.)*, **12**, 121–127.
73. Cho,J.E., Kim,N., Li,Y.C. and Jinks-Robertson,S. (2013) Two distinct mechanisms of Topoisomerase 1-dependent mutagenesis in yeast. *DNA Repair (Amst.)*, **12**, 205–211.
74. Kim,N., Huang,S.N., Williams,J.S., Li,Y.C., Clark,A.B., Cho,J.E., Kunkel,T.A., Pommier,Y. and Jinks-Robertson,S. (2011) Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I. *Science*, **332**, 1561–1564.
75. Lazzaro,F., Novarina,D., Amara,F., Watt,D.L., Stone,J.E., Costanzo,V., Burgers,P.M., Kunkel,T.A., Plevani,P. and Muzi-Falconi,M. (2012) RNase H and postreplication repair protect cells from ribonucleotides incorporated in DNA. *Mol. Cell*, **45**, 99–110.

76. Moore,A., Dominska,M., Greenwell,P., Aksenova,A.Y., Mirkin,S. and Petes,T. (2018) Genetic control of genomic alterations induced in yeast by interstitial telomeric sequences. *Genetics*, **209**, 425–438.
77. Saini,N., Zhang,Y., Nishida,Y., Sheng,Z., Choudhury,S., Mieczkowski,P. and Lobachev,K.S. (2013) Fragile DNA motifs trigger mutagenesis at distant chromosomal loci in *Saccharomyces cerevisiae*. *PLos Genet.*, **9**, e1003551.
78. Tang,W., Dominska,M., Gawel,M., Greenwell,P.W. and Petes,T.D. (2013) Genomic deletions and point mutations induced in *Saccharomyces cerevisiae* by the trinucleotide repeats (GAA·TTC) associated with Friedreich's ataxia. *DNA Repair (Amst.)*, **12**, 10–17.
79. Shah,K.A. and Mirkin,S.M. (2015) The hidden side of unstable DNA repeats: Mutagenesis at a distance. *DNA Repair (Amst.)*, **32**, 106–112.
80. Shah,K.A., Shishkin,A.A., Voineagu,I., Pavlov,Y.I., Shcherbakova,P.V. and Mirkin,S.M. (2012) Role of DNA polymerases in repeat-mediated genome instability. *Cell Rep.*, **2**, 1088–1095.
81. Holbeck,S.L. and Strathern,J.N. (1997) A role for Rev3 in mutagenesis during Double-Strand break repair in saccharomyces cerevisiae. *Genetics*, **147**, 1017–1024.
82. Rattray,A.J., Shafer,B.K., McGill,C.B. and Strathern,J.N. (2002) The roles of REV3 and RAD57 in double-strand-break-repair-induced mutagenesis of *Saccharomyces cerevisiae*. *Genetics*, **162**, 1063–1077.
83. McVey,M., Khodaverdian,V.Y., Meyer,D., Cerqueira,P.G. and Heyer,W.D. (2016) Eukaryotic DNA polymerases in homologous recombination. *Annu. Rev. Genet.*, **50**, 393–421.
84. Sakofsky,C.J., Ayyar,S., Deem,A.K., Chung,W.H., Ira,G. and Malkova,A. (2015) Translesion polymerases drive Microhomology-Mediated Break-Induced replication leading to complex chromosomal rearrangements. *Mol. Cell*, **60**, 860–872.
85. Lemacon,D., Jackson,J., Quinet,A., Brickner,J.R., Li,S., Yazinski,S., You,Z., Ira,G., Zou,L., Mosammaparast,N. *et al.* (2017) MRE11 and EXO1 nucleases degrade reversed forks and elicit MUS81-dependent fork rescue in BRCA2-deficient cells. *Nat. Commun.*, **8**, 860.
86. Putnam,C.D., Srivatsan,A., Nene,R.V., Martinez,S.L., Clotfelter,S.P., Bell,S.N., Somach,S.B., de Souza,J.E., Fonseca,A.F., de Souza,S.J. *et al.* (2016) A genetic network that suppresses genome rearrangements in *Saccharomyces cerevisiae* and contains defects in cancers. *Nat. Commun.*, **7**, 11256.
87. Lange,S.S., Bedford,E., Reh,S., Wittschieben,J.P., Carbajal,S., Kusewitt,D.F., DiGiovanni,J. and Wood,R.D. (2013) Dual role for mammalian DNA polymerase ζ in maintaining genome stability and proliferative responses. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, E687–E696.
88. Suzuki,T., Gruz,P., Honma,M., Adachi,N. and Nohmi,T. (2016) The role of DNA polymerase ζ in translesion synthesis across bulky DNA adducts and cross-links in human cells. *Mutat. Res.*, **791–792**, 35–41.
89. Saha,L.K., Kim,S., Kang,H., Akter,S., Choi,K., Sakuma,T., Yamamoto,T., Sasanuma,H., Hirota,K., Nakamura,J. *et al.* (2018) Differential micronucleus frequency in isogenic human cells deficient in DNA repair pathways is a valuable indicator for evaluating genotoxic agents and their genotoxic mechanisms. *Environ. Mol. Mutagen.*, **59**, 529–538.
90. Raykov,V., Marvin,M.E., Louis,E.J. and Maringele,L. (2016) Telomere dysfunction triggers palindrome formation independently of Double-Strand break repair mechanisms. *Genetics*, **203**, 1659–1668.
91. Wittschieben,J.P., Patil,V., Glushets,V., Robinson,L.J., Kusewitt,D.F. and Wood,R.D. (2010) Loss of DNA polymerase zeta enhances spontaneous tumorigenesis. *Cancer Res.*, **70**, 2770–2778.
92. Schenten,D., Kracker,S., Esposito,G., Franco,S., Klein,U., Murphy,M., Alt,F.W. and Rajewsky,K. (2009) Polζ ablation in B cells impairs the germinal center reaction, class switch recombination, DNA break repair, and genome stability. *J. Exp. Med.*, **206**, 477–490.
93. Daly,J., Bebenek,K., Watt,D.L., Richter,K., Jiang,C., Zhao,M.L., Ray,M., McGregor,W.G., Kunkel,T.A. and Diaz,M. (2012) Altered Ig hypermutation pattern and frequency in complementary mouse models of DNA polymerase ζ activity. *J. Immunol.*, **188**, 5528–5537.
94. Lange,S.S., Bhetawal,S., Reh,S., Powell,K.L., Kusewitt,D.F. and Wood,R.D. (2018) DNA polymerase ζ deficiency causes impaired wound healing and stress-induced skin pigmentation. *Life Sci. Alliance*, **1**, e201800048.
95. McGowan,K.A., Li,J.Z., Park,C.Y., Beaudry,V., Tabor,H.K., Sabnis,A.J., Zhang,W., Fuchs,H., de Angelis,M.H., Myers,R.M. *et al.* (2008) Ribosomal mutations cause p53-mediated dark skin and pleiotropic effects. *Nat. Genet.*, **40**, 963–970.
96. Cui,R., Widlund,H.R., Feige,E., Lin,J.Y., Wilensky,D.L., Igras,V.E., D'Orazio,J., Fung,C.Y., Schanbacher,C.F., Granter,S.R. *et al.* (2007) Central role of p53 in the suntan response and pathologic hyperpigmentation. *Cell*, **128**, 853–864.
97. Hyter,S., Coleman,D.J., Ganguli-Indra,G., Merrill,G.F., Ma,S., Yanagisawa,M. and Indra,A.K. (2013) Endothelin-1 is a transcriptional target of p53 in epidermal keratinocytes and regulates ultraviolet-induced melanocyte homeostasis. *Pigment Cell Melanoma Res.*, **26**, 247–258.
98. Chen,S., Zhu,B., Yin,C., Liu,W., Han,C., Chen,B., Liu,T., Li,X., Chen,X., Li,C. *et al.* (2017) Palmitoylation-dependent activation of MC1R prevents melanomagenesis. *Nature*, **549**, 399–403.
99. Liu,H., Murphy,C.J., Karreth,F.A., Emdal,K.B., White,F.M., Elemento,O., Toker,A., Wulf,G.M. and Cantley,L.C. (2018) Identifying and targeting sporadic oncogenic genetic aberrations in mouse models of Triple-Negative breast cancer. *Cancer Discov.*, **8**, 354–369.
100. Fritzen,R., Delbos,F., De Smet,A., Palancade,B., Canman,C.E., Aoufouchi,S., Weill,J.C., Reynaud,C.A. and Storck,S. (2016) A single aspartate mutation in the conserved catalytic site of Rev3L generates a hypomorphic phenotype in vivo and in vitro. *DNA Repair (Amst.)*, **46**, 37–46.
101. Lange,S.S., Wittschieben,J.P. and Wood,R.D. (2012) DNA polymerase ζ is required for proliferation of normal mammalian cells. *Nucleic Acids Res.*, **40**, 4473–4482.
102. Listovsky,T. and Sale,J.E. (2013) Sequestration of CDH1 by MAD2L2 prevents premature APC/C activation prior to anaphase onset. *J. Cell Biol.*, **203**, 87–100.
103. Medendorp,K., van Groningen,J.J., Vreede,L., Hetterschijt,L., van den Hurk,W.H., de Bruijn,D.R., Brugmans,L. and van Kessel,A.G. (2009) The mitotic arrest deficient protein MAD2B interacts with the small GTPase RAN throughout the cell cycle. *PLoS One*, **4**, e7020.
104. Hara,K., Taharazako,S., Ikeda,M., Fujita,H., Mikami,Y., Kikuchi,S., Hishiki,A., Yokoyama,H., Ishikawa,Y., Kanno,S.I. *et al.* (2017) Dynamic feature of mitotic arrest deficient 2-like protein 2 (MAD2L2) and structural basis for its interaction with chromosome alignment-maintaining phosphoprotein (CAMP). *J. Biol. Chem.*, **292**, 17658–17667.
105. Pirouz,M., Pilarski,S. and Kessel,M. (2013) A critical function of Mad2l2 in primordial germ cell development of mice. *PLos Genet.*, **9**, e1003712.
106. Vermeulen,M., Eberl,H.C., Matarese,F., Marks,H., Denissov,S., Butter,F., Lee,K.K., Olsen,J.V., Hyman,A.A., Stunnenberg,H.G. *et al.* (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell*, **142**, 967–980.
107. Itoh,G., Kanno,S., Uchida,K.S., Chiba,S., Sugino,S., Watanabe,K., Mizuno,K., Yasui,A., Hirota,T. and Tanaka,K. (2011) CAMP (C13orf8, ZNF828) is a novel regulator of kinetochore-microtubule attachment. *EMBO J.*, **30**, 130–144.
108. Pfleger,C.M., Salic,A., Lee,E. and Kirschner,M.W. (2001) Inhibition of Cdh1-APC by the MAD2-related protein MAD2L2: a novel mechanism for regulating Cdh1. *Genes Dev.*, **15**, 1759–1764.
109. Boersma,V., Moatti,N., Segura-Bayona,S., Peuscher,M.H., van der Torre,J., Wevers,B.A., Orthwein,A., Durocher,D. and Jacobs,J.J. (2015) MAD2L2 controls DNA repair at telomeres and DNA breaks by inhibiting 5' end resection. *Nature*, **521**, 537–540.
110. Xu,G., Chapman,J.R., Brandsma,I., Yuan,J., Mistrik,M., Bouwman,P., Bartkova,J., Gogola,E., Warmerdam,D., Barazas,M. *et al.* (2015) REV7 counteracts DNA double-strand break resection and affects PARP inhibition. *Nature*, **521**, 541–544.
111. Watanabe,N., Mii,S., Asai,N., Asai,M., Niimi,K., Ushida,K., Kato,T., Enomoto,A., Ishii,H., Takahashi,M. *et al.* (2013) The REV7 subunit of DNA polymerase ζ is essential for primordial germ cell maintenance in the mouse. *J. Biol. Chem.*, **288**, 10459–10471.

112. Ghezraoui,H., Oliveira,C., Becker,J.R., Bilham,K., Moralli,D., Anzilotti,C., Fischer,R., Deobagkar-Lele,M., Sanchiz-Calvo,M., Fueyo-Marcos,E. *et al.* (2018) 53BP1 cooperation with the REV7-shieldin complex underpins DNA structure-specific NHEJ. *Nature*, **560**, 122–127.
113. Sousa,F.G., Matuo,R., Tang,S.W., Rajapakse,V.N., Luna,A., Sander,C., Varma,S., Simon,P.H., Doroshow,J.H., Reinhold,W.C. *et al.* (2015) Alterations of DNA repair genes in the NCI-60 cell lines and their predictive value for anticancer drug activity. *DNA Repair (Amst.)*, **28**, 107–115.
114. Cerami,E., Gao,J., Dogrusoz,U., Gross,B.E., Sumer,S.O., Aksoy,B.A., Jacobsen,A., Byrne,C.J., Heuer,M.L., Larsson,E. *et al.* (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.*, **2**, 401–404.
115. Gao,J., Aksoy,B.A., Dogrusoz,U., Dresdner,G., Gross,B., Sumer,S.O., Sun,Y., Jacobsen,A., Sinha,R., Larsson,E. *et al.* (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.*,**6**, pl1.
116. Tomas-Roca,L., Tsaalbi-Shtylik,A., Jansen,J.G., Singh,M.K., Epstein,J.A., Altunoglu,U., Verzijl,H., Soria,L., van Beusekom,E., Roscioli,T. *et al.* (2015) De novo mutations in PLXND1 and REV3L cause Mobius syndrome. *Nat. Commun.*, **6**, 7199.
117. Hanahan,D. and Weinberg,R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646–674.
118. Brown,J.S., O'Carrigan,B., Jackson,S.P. and Yap,T.A. (2017) Targeting DNA repair in Cancer: Beyond PARP inhibitors. *Cancer Discov.*, **7**, 20–37.
119. Doles,J., Oliver,T.G., Cameron,E.R., Hsu,G., Jacks,T., Walker,G.C. and Hemann,M.T. (2010) Suppression of Rev3, the catalytic subunit of Pol ζ, sensitizes drug-resistant lung tumors to chemotherapy. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 20786–20791.
120. Xie,K., Doles,J., Hemann,M.T. and Walker,G.C. (2010) Error-prone translesion synthesis mediates acquired chemoresistance. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 20792–20797.
121. Xu,X., Xie,K., Zhang,X.Q., Pridgen,E.M., Park,G.Y., Cui,D.S., Shi,J., Wu,J., Kantoff,P.W., Lippard,S.J. *et al.* (2013) Enhancing tumor cell response to chemotherapy through nanoparticle-mediated codelivery of siRNA and cisplatin prodrug. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 18638–18643.
122. Yamanaka,K., Chatterjee,N., Hemann,M.T. and Walker,G.C. (2017) Inhibition of mutagenic translesion synthesis: A possible strategy for improving chemotherapy? *PLos Genet.*, **13**, e1006842.
123. Sail,V., Rizzo,A.A., Chatterjee,N., Dash,R.C., Ozen,Z., Walker,G.C., Korzhnev,D.M. and Hadden,M.K. (2017) Identification of small molecule translesion synthesis inhibitors that target the Rev1-CT/RIR Protein-Protein interaction. *ACS Chem. Biol.*, **12**, 1903–1912.
124. Tsaalbi-Shtylik,A., Verspuy,J.W., Jansen,J.G., Rebel,H., Carlee,L.M., van der Valk,M.A., Jonkers,J., de Gruijl,F.R. and de Wind,N. (2009) Error-prone translesion replication of damaged DNA suppresses skin carcinogenesis by controlling inflammatory hyperplasia. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 21836–21841.
125. Kelley,L.A., Mezulis,S., Yates,C.M., Wass,M.N. and Sternberg,M.J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **10**, 845–858.

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-019-12234-1 OPEN

Genetic determinants of cellular addiction to DNA polymerase theta

Wanjuan Feng[1], Dennis A. Simpson[1], Juan Carvajal-Garcia[1], Brandon A. Price[1], Rashmi J. Kumar[1], Lisle E. Mose[1], Richard D. Wood[2], Naim Rashid[1,3], Jeremy E. Purvis[4], Joel S. Parker[1,4], Dale A. Ramsden[1,5] & Gaorav P. Gupta[1,5,6]

Polymerase theta (Pol θ, gene name *Polq*) is a widely conserved DNA polymerase that mediates a microhomology-mediated, error-prone, double strand break (DSB) repair pathway, referred to as Theta Mediated End Joining (TMEJ). Cells with homologous recombination deficiency are reliant on TMEJ for DSB repair. It is unknown whether deficiencies in other components of the DNA damage response (DDR) also result in Pol θ addiction. Here we use a CRISPR genetic screen to uncover 140 *Polq* synthetic lethal (PolqSL) genes, the majority of which were previously unknown. Functional analyses indicate that Pol θ/TMEJ addiction is associated with increased levels of replication-associated DSBs, regardless of the initial source of damage. We further demonstrate that approximately 30% of TCGA breast cancers have genetic alterations in PolqSL genes and exhibit genomic scars of Pol θ/TMEJ hyperactivity, thereby substantially expanding the subset of human cancers for which Pol θ inhibition represents a promising therapeutic strategy.

[1] Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [2] Department of Epigenetics and Molecular Carcinogenesis, University of Texas MD Anderson Cancer Center, Smithville, TX 78957, USA. [3] Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [4] Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [5] Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [6] Department of Radiation Oncology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. Correspondence and requests for materials should be addressed to G.P.G. (email: gaorav@med.unc.edu)

DNA double strand breaks (DSBs) arise spontaneously during DNA replication or upon exposure to exogenous clastogens and threaten both genome integrity and cellular viability[1–3]. Efficient and accurate DSB repair is thus vital for cancer prevention and organismal survival. DSB repair pathways are broadly classified into two categories: homology-directed repair (HDR) and non-homologous end joining (NHEJ). HDR requires 5′ to 3′ end resection, Rad51 loading, strand invasion, and DNA synthesis using an intact homologous template[4]. In contrast, classical NHEJ (c-NHEJ) does not require a homologous template and is dependent on the Ku complex, DNA-PK, and XRCC4/Ligase 4[5]. An alternative end joining (alt-EJ) pathway has also been described, but unlike c-NHEJ, alt-EJ acts on the same 5′ to 3′ resected DSBs that are intermediates in HR. Alt-EJ employs a synthesis-dependent mechanism that is directed by short tracts of flanking microhomology (MH)[1,6], giving rise to a characteristic pattern of MH-flanked deletions and/or templated insertions. Several genes have been implicated in alt-EJ, including 5′ to 3′ resection factors (e.g., Mre11, Rad50, Nbn, CtiP, and Exo1), PARP1, and LIG3. However, the gene that is most specifically linked to Alt-EJ is the A-family DNA Polymerase θ (Pol θ, gene name *Polq*)[1,7]. Alt-EJ signatures at chromosomal breaks are substantially reduced in *Polq*$^{-/-}$ cells from diverse metazoan and plant organisms[8–10]. Thus, Pol θ has emerged as the predominant mediator of alt-EJ, and this alternative DSB repair pathway has been designated Theta Mediated End Joining (TMEJ)[9,11].

TMEJ is intrinsically an error-prone pathway, yet its evolutionary conservation in metazoans and plants suggests that it likely has a physiological role in promoting genome integrity[1]. Indeed, *Polq*$^{-/-}$ cells demonstrate elevated levels of spontaneous DNA damage[12]. A prior study suggested that TMEJ competes with HDR for DSB repair[13], but this model does not explain how TMEJ may promote genome stability. In *C. elegans*, TMEJ has an important role in the repair of replication-associated DSBs, particularly at G-quadruplex (G4) structures[14,15]. In that study, Pol θ deficiency resulted in large-scale deletions at chromosomal G4 sites. However, the physiological role of TMEJ in promoting genome integrity in mammals remains unclear.

In normal cells, TMEJ accounts for a small minority of DSB repair[10]. Consistent with a limited role in global DSB repair, *Polq* deficiency has a relatively minor impact on organismal development in flies[16], worms[17], and mice[12]. However, recent studies have demonstrated that *Polq* nevertheless becomes essential in cells with deficiency in canonical DSB repair pathway genes (*Brca1*, *Brca2*, and *Ku70*), indicating synthetic lethal genetic interactions that are consistent with an essential role for Polθ/TMEJ as a backup to repair by either HR or NHEJ[10,13,18]. This observation has resulted in enthusiasm for Pol θ as a therapeutic target in breast and ovarian cancers with *BRCA1/2* deficiency[19]. However, it remains unknown whether *Polq* is also synthetic lethal with other genes in the HR and NHEJ pathways, and more broadly, with other genetic mediators of the DNA damage response (DDR) pathway. Here, we report findings from a synthetic lethal CRISPR screen to identify DDR gene mutations that induce cellular addiction to Pol θ. We uncover a broad landscape of synthetic lethality with *Polq*, and provide evidence that this reflects a critical role for Pol θ in protecting cells from accumulation of non-productive HR intermediates at sites of DNA replication-associated DSBs. Finally, we find that human breast cancers with mutations in *Polq* synthetic lethal (PolqSL) genes identified in our CRISPR screen may be addicted to Pol θ, based on increased expression of TMEJ-associated genomic scars.

Results

CRISPR synthetic lethal screens. To gain broader insight into the contexts where Pol θ-mediated genome maintenance is essential for cellular viability, we performed a CRISPR loss of function screen in *WT*, *Polq*$^{-/-}$, and *Polq*^{hPOLQ} (*Polq*$^{-/-}$ reconstituted with human *POLQ*) MEF cell lines, which were described previously and functionally validated[10,12,20]. *Polq*$^{-/-}$ MEFs have a normal cell cycle profile[10], yet exhibit elevated levels of spontaneous chromosomal aberrations that are reversed after complementation with human *POLQ* (Supplementary Fig. 1). The goal of the CRISPR screen was to identify gene mutations that are tolerated in *WT* and *Polq*^{hPOLQ} MEFs yet lethal in *Polq*$^{-/-}$ MEFs, thereby indicative of a synthetic lethal genetic interaction. A custom synthesized "DDR-CRISPR" lentiviral library was used for the screen, which targets 309 murine DDR genes with 10 small guide RNAs (sgRNAs) per gene and also includes 834 non-targeting sgRNA controls (Supplementary Data 1). For each biological replicate, 2×10^6 MEFs were transduced with the DDR-CRISPR lentiviral library at low multiplicity of infection (<1), and passaged for 8 population doublings prior to genomic DNA isolation (Fig. 1a). High-throughput sequencing (average 250× read depth) was used to quantify the abundance of each sgRNA sequence relative to all mapped reads, similar to previously described methods[21] (Fig. 1a). A "Gene Abundance Change Score" was calculated as described in the methods. Thresholds for statistical significance were established by using the set of control sgRNAs as an internal control for abundance changes that are due to off-target effects (see "Methods").

Plotting the Gene Abundance Change Scores, we observed a striking depletion of many DDR gene-targeting sgRNAs in *Polq*$^{-/-}$ relative to *WT* MEFs (Fig. 1b). In contrast, control sgRNAs were not depleted in *Polq*$^{-/-}$ cells and, in fact, were enriched relative to their abundance in *WT* cells due to the depletion of a large number of DDR gene-targeting sgRNAs. Moreover, the vast majority of sgRNA abundance changes in *Polq*$^{-/-}$ MEFs could be definitively attributed to Pol θ deficiency, as they were not observed when *Polq*$^{-/-}$ cells were reconstituted with WT human *POLQ* (*Polq*^{hPOLQ}) (Fig. 1b). To mitigate any clone-specific genetic interactions we directly compared Gene Abundance Change Scores in *Polq*$^{-/-}$ MEFs relative to *Polq*^{hPOLQ} MEFs, and identified 142 significant genetic interactions using two complementary statistical tests (Fig. 1c). All but two of these genes (140 total) had corresponding sgRNAs that were depleted in *Polq*$^{-/-}$ MEFs relative to reconstituted *Polq*^{hPOLQ} MEFs, and thus classified as *Polq* synthetic lethal (PolqSL) genes.

Due to the large proportion of *Polq* synthetic lethal gene interactions identified in our screen (45% of 309 genes evaluated), we performed two additional control experiments. First, we conducted the same DDR-CRISPR screen in an immortalized MEF line that is deficient in another DNA repair polymerase, Pol μ (*Polm*), that participates in NHEJ repair[22]. We did not identify any statistically significant synthetic lethal gene interactions with *Polm* deficiency (Supplementary Fig. 2a, b and Supplementary Data 3), indicating the broad landscape of DDR gene synthetic lethality is not observed for all DNA repair-associated polymerases. To address whether *Polq*$^{-/-}$ cells are prone to synthetic sickness with Cas9-mediated gene editing events, we utilized a separate CRISPR library targeting genes that encode membrane proteins. Only 19 out of 951 genes (2%) targeted in this library exhibited synthetic lethality with *Polq*$^{-/-}$, which is below the 3% false discovery rate threshold used during statistical analyses (Supplementary Fig. 2c, d and Supplementary Data 3). Thus, the large number of PolqSL genes identified in our screen is due to a broad landscape of DDR gene mutations that render cells dependent on *Polq* for viability.

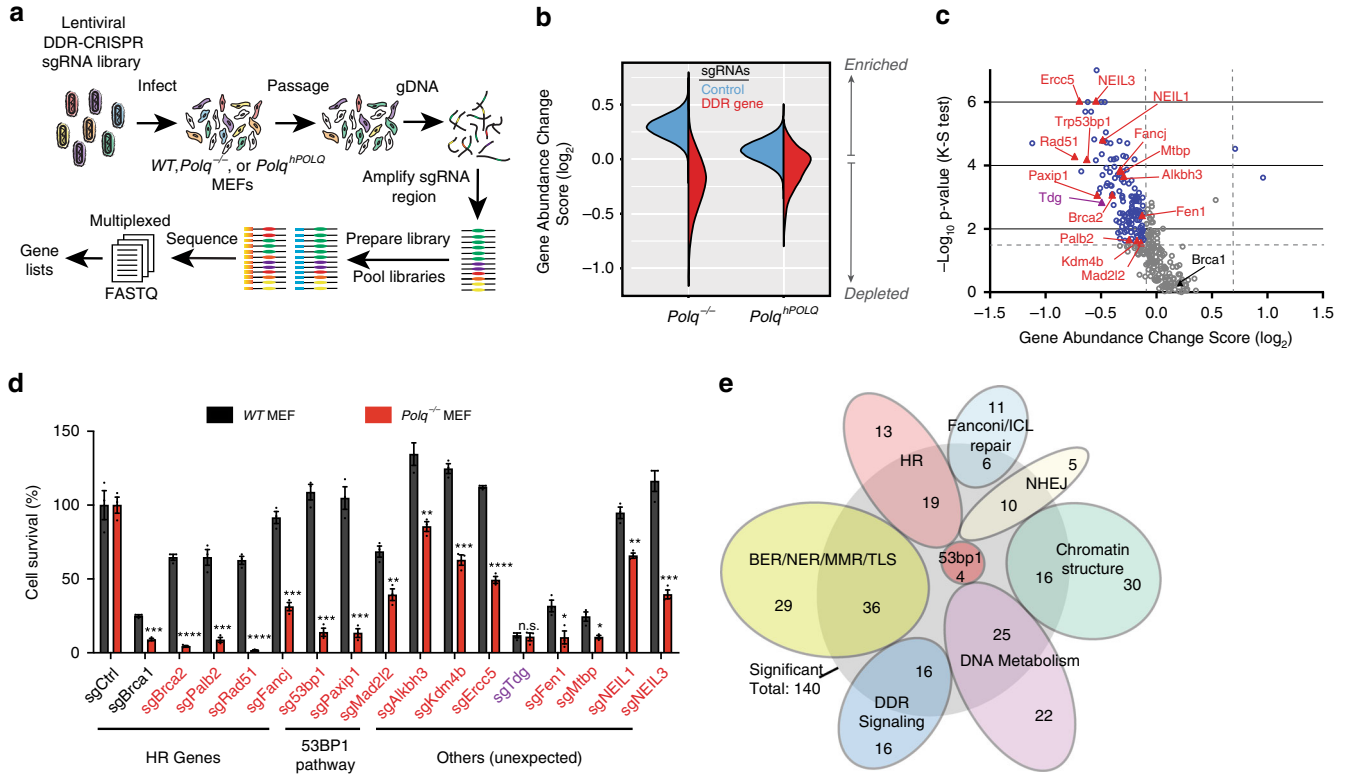


Fig. 1 Identification of *Polq* synthetic lethal (PolqSL) genes by CRISPR screening. **a** Schematic of the CRISPR genetic screen to identify PolqSL genes. **b** Violin plot of Gene Abundance Change Scores (Log_2) for DDR gene-targeting sgRNAs (red) and non-targeting control sgRNAs (blue) in $Polq^{-/-}$ and $Polq^{hPOLQ}$ MEFs, relative to *WT* MEFs. **c** Volcano plot of Gene Abundance Change Scores ($Polq^{-/-}$ versus $Polq^{hPOLQ}$) and -Log_{10} p-value of the Kolmogorov-Smirnov test for DDR gene-targeting sgRNAs relative to non-targeting control sgRNAs. Thresholds for statistical significance are indicated by dashed lines (see "Methods" for details). Genes with statistically significant (Blue dots) and non-significant (Gray dots) Gene Abundance Changes Scores are indicated. Genes with red/purple triangles are further validated in Fig. 1d. **d** Relative cell survival measured by colony forming efficiency of *WT* or $Polq^{-/-}$ cells transduced with a lentivirus containing Cas9 and control sgRNA (sgCtrl) or DDR gene-targeting sgRNAs. Data shown are the mean ± SEM ($n = 3$ biological replicates). Significance determined using an unpaired, two-tailed t-test (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$, ****$p < 0.0001$). **e** Functional classification of PolqSL genes identified in our CRISPR screen depicted as a Euler diagram

We validated 15 of the candidate PolqSL genes using standard colony forming assays after transduction with gene-targeting or control sgRNAs (Fig. 1d). For 14 out of 15 genes, we observed significantly reduced viability in $Polq^{-/-}$ MEFs relative to *WT* MEFs (i.e., 93% hit validation rate). We also tested sgRNAs against *Brca1* in *WT* and $Polq^{-/-}$ MEFs, due to previously published reports of a synthetic lethal interaction[13,18], although *Brca1* did not emerge as a significant genetic interaction in our CRISPR screen. We observed a modest yet statistically significant reduction in cell viability when sgBrca1 was introduced in $Polq^{-/-}$ cells relative to *WT* cells. The relatively small magnitude of viability difference between $Polq^{-/-}$ and *WT* MEFs transduced with sgBrca1 may explain why it was a false negative result in our screen.

Previous work identified two members of the HR pathway (*Brca1* and *Brca2*)[13,18], and 1 member of the NHEJ pathway (*Ku70*)[10], as synthetic lethal with *Polq* deficiency. Our work considerably expands the list of DSB repair genes that are synthetic sick or lethal with *Polq*, such that it now includes 13 additional HR mediators, as well as 4 additional genes specific to NHEJ (Fig. 1e, Supplementary Data 2). We also observed highly significant synthetic sickness between *Polq* deficiency and all four components of the 53BP1 anti-resection pathway included in our screen (*53bp1*, *Paxip1*, *Mad2l2*, and *Rif1*). Surprisingly, many of the remaining PolqSL genes have no direct role in canonical DSB repair. These include genes involved in base/nucleotide excision repair, translesion synthesis, mismatch repair, DNA metabolism, DDR signaling, chromatin structure, and the Fanconi Anemia repair pathway (Fig. 1e). We postulated that a common feature of these gene mutations may be an increase in endogenously generated replication-associated DSBs. To directly test whether Pol θ is essential for repair of collapsed replication forks, we quantified chromosomal aberrations after Aphidicolin treatment. $Polq^{-/-}$ MEFs accumulated significantly more metaphase aberrations and had reduced viability after aphidicolin treatment relative to *WT* or $Polq^{hPOLQ}$ MEFs (Supplementary Fig. 3a-c). Loss of Neil3 has previously been shown to increase replication-associated DSBs[23]. We identified *Neil3* as a PolqSL gene, and observed that CRISPR-mediated knockout of *Neil3* increased nuclear 53BP1 foci more significantly in $Polq^{-/-}$ MEFs relative to *WT* MEFs, which is consistent with an accumulation of unrepaired replication-associated DSBs. Collectively, these findings argue an essential role for *Polq* is not limited to cells deficient in *BRCA1/2*, or even cells deficient in DSB repair—we show Pol θ is an important compensatory repair mechanism in the background of deficiency in many genes implicated in DDR.

Synthetic lethality of *Polq*/*53bp1* DKO cells despite HR and NHEJ proficiency. Synthetic lethality between *Polq* and *53bp1* has previously been reported[10], and had been presumed to be due to deficiency in NHEJ. To evaluate this possibility, we measured NHEJ, HR, and TMEJ repair at a CRISPR/Cas9 induced break at the murine Rosa26 locus using digital PCR (dPCR) assays designed based on previously published high throughput

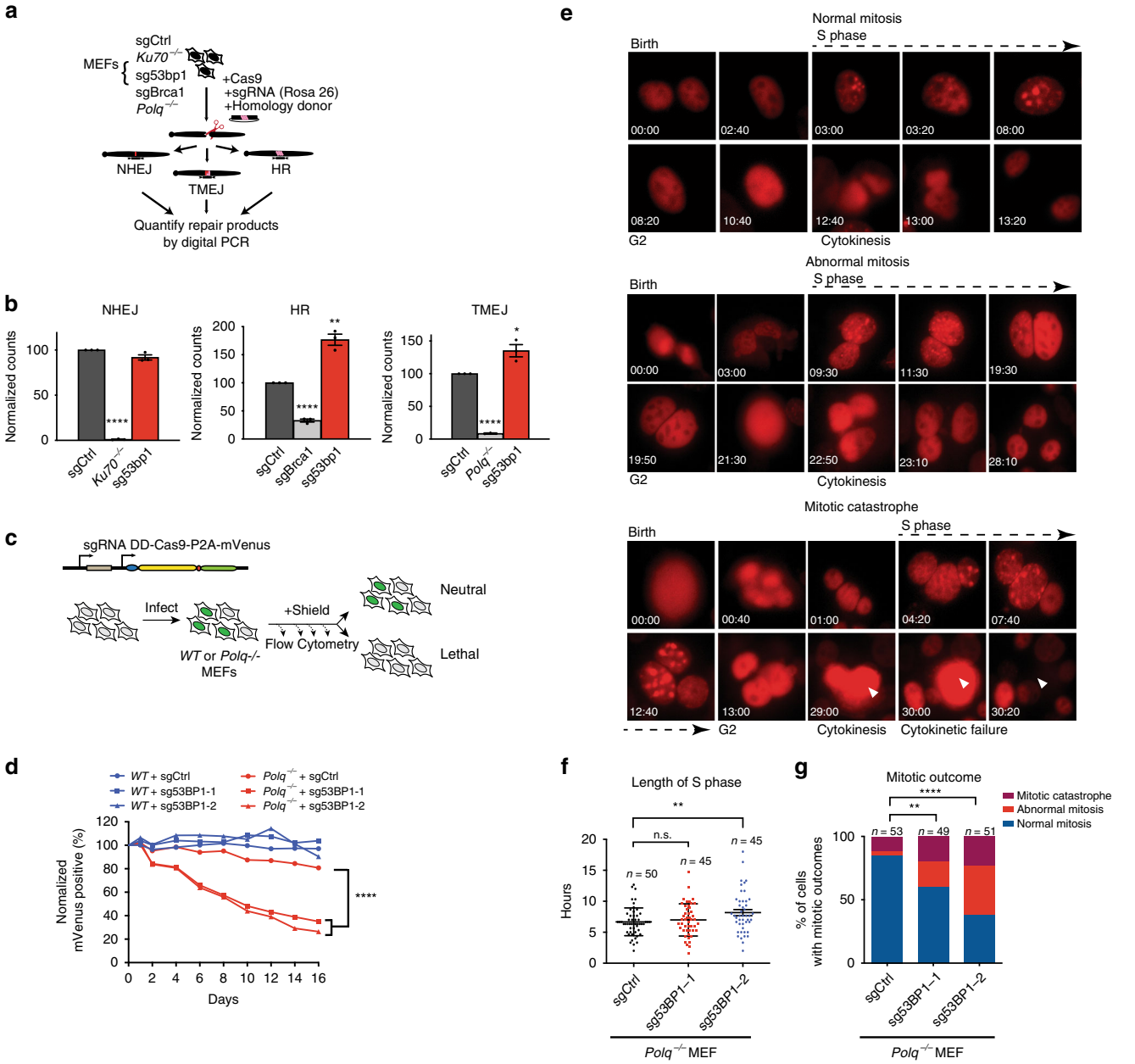


Fig. 2 Synthetic lethality of *Polq/53bp1* double knockout cells despite HR and NHEJ proficiency. **a** Schema for marker-free quantification of DSB repair pathway choice at a CRISPR-induced chromosomal break using dPCR. **b** Quantification of HR, NHEJ, and TMEJ repair at the *Rosa26* locus, relative to *WT* + sgCtrl. Data shown are the mean ± SEM ($n = 3$ biological replicates). Significance determined using an unpaired, two-tailed *t*-test (*$p < 0.05$; **$p < 0.01$; ****$p < 0.0001$). **c** Diagram of growth competition assay to assess kinetics of synthetic lethality between *Polq* and *53bp1*. DD-Cas9 is protected from degradation upon exposure to a synthetic ligand (Shield1). Depletion of mVenus-positive cells over time after Shield1 exposure is indicative of sgRNA lethality. **d** Normalized percentage of mVenus-positive cells over time after Shield1 treatment, measured by flow cytometry. The mean fraction of mVenus-positive cells ± SEM ($n = 3$) is shown for the various genotypes, normalized to day 0 (no shield). *Polq*$^{-/-}$ + sg53bp1-1 or -2 ****, $p < 0.0001$ using two-tailed nonparametric Spearman correlation in GraphPad Prisim v7.04. **e**–**g** Time lapse microscopy of PCNA-mCherry to assess cell cycle phase transitions in individual *Polq*$^{-/-}$ cells 48 h after Shield1 treatment. Cell lines used in this experiment are described in (**c**) after mVenus sorting. **e** Image descriptions for the three mitotic outcomes. Mitotic catastrophe is a terminal event with cells undergoing nuclear degradation, and abnormal mitosis refers to cytokinesis failure or chromosomal mis-segregation events resulting in abnormal nuclear structure in subsequent daughter cells. **f** Distribution of S phase lengths in *Polq*$^{-/-}$ MEFs transduced with sgCtrl, sg53bp1-1, and sg53bp1-2. **$p < 0.01$ using a two-tailed *t*-test. **g** Analysis of mitotic outcome of individual cells with the indicated genotypes. Abnormal mitosis refers to cytokinetic failure or mis-segregation events resulting in abnormal nuclear structure in subsequent daughter cells. Mitotic catastrophe refers to abnormal mitoses resulting in cell death or disappearance. Significance determined using a Chi-square test (**$p < 0.01$ and ****$p < 0.0001$)

sequencing analyses at this locus[10] (Fig. 2a). *53bp1* deficiency resulted in a nonsignificant reduction in NHEJ and increased frequencies of both HR and TMEJ repair (Fig. 2b), consistent with previously described roles for 53bp1 in DSB repair pathway choice[24]. Based on these observations, NHEJ deficiency cannot explain synthetic lethality between *53bp1* and *Polq*. However, it was also surprising that HR could not compensate for TMEJ deficiency in *53bp1*/*Polq* DKO cells, given that they both act on resected DSBs. We therefore investigated the synthetic lethal phenotype in greater detail.

To assess kinetics of synthetic lethality between *53bp1* and *Polq*, we utilized an inducible Cas9 expression system (DD-Cas9[25]), and monitored the relative growth rate of transduced cells by flow cytometry over time (Fig. 2c). Interestingly, although 53BP1 expression was already diminished by 48 h after Shield1 treatment (Supplementary Fig. 4), the growth disadvantage of $Polq^{-/-}$ + sg53bp1 (i.e., *Polq*/*53bp1* double knockout, DKO) cells persisted over at least 14 days (Fig. 2d). Time lapse microscopy using PCNA-mCherry as a fluorescent cell cycle reporter[26] revealed a statistically significant prolongation of S phase duration with one out of two *53bp1*-targeting sgRNAs in $Polq^{-/-}$ cells (Fig. 2f), although G1 and G2/M duration did not differ significantly for any of the genotypes (Supplementary Fig. 5). More strikingly, there was a significantly higher rate of aberrant mitoses (improper chromosomal segregation or abnormal cytokinesis) and mitotic catastrophe when either of the *53bp1* targeting sgRNAs was expressed in $Polq^{-/-}$ cells (Fig. 2e, g).

***Polq*/*53bp1* DKO cells accumulate aberrant HR intermediates in S phase**. Because mitotic aberrations can arise from unresolved DNA damage in the preceding S phase[27–29], we performed co-immunofluorescence for Rad51 and γH2AX to assess levels of HR intermediates and DNA damage-associated chromatin, respectively. Notably, we observed large Rad51 aggregates selectively in *Polq*/*53bp1* DKO cells, which frequently were also positive for γH2AX (Fig. 3a, b). The Rad51 foci observed in *Polq*/*53bp1* DKO cells were substantially larger than spontaneous Rad51 foci that occur in a normal S phase in *WT* cells (Supplementary Fig. 6a–b). In addition, we analyzed EdU incorporation to distinguish non-S phase cells from cells in early, middle, or late S phase (Fig. 3c). The most significant increase in abnormal Rad51 aggregates was observed in middle and late S phase cells (Fig. 3d). We hypothesized that these Rad51 foci arose from spontaneous replication fork collapse. Indeed, aphidicolin treatment increased the percentage of nuclei with large Rad51 foci (Fig. 3e). Furthermore, Rad51 foci that formed in *Polq*/*53bp1* DKO cells persisted even after 12 h, a timepoint when a significant fraction of Rad51 foci had resolved in *WT*, *WT* + sg53bp1, and $Polq^{-/-}$ cells (Fig. 3e and Supplementary Fig. 6c). Collectively, these observations indicate that synthetic lethality between *53bp1* and *Polq* deficiency is due to unsuccessful HR-mediated repair of a subset of replication-associated DSBs.

***Polq* is required for Mitomycin C (MMC) induced DNA damage repair**. We next evaluated potential roles for Pol θ after exposure to agents known to cause stalled replication forks, including MMC, which introduces interstrand crosslinks (ICL), and pyridostatin (PDS), which stabilizes G quadruplex (G4) DNA. Prior findings in Drosophila[8,12,30] have implicated *Polq* in ICL repair. In contrast, MEFs expressing a hypomorphic *Polq* allele, $Polq^{chaos1}$, were not hypersensitive to MMC[12]. We find that $Polq^{-/-}$ MEFs are hypersensitive to MMC, which can be restored by reconstitution with human *POLQ* (Supplementary Fig. 7). The discrepancy between these findings may be due to residual activity of the $Polq^{chaos1}$ allele in mediating ICL repair. $Polq^{-/-}$ MEFs exposed to a low dose of MMC (20 ng/mL) had a significantly higher frequency of mitotic crossovers in a sister chromatid exchange (SCE) assay and unrepaired chromosomal aberrations than was observed in wild type cells. Increases in both classes of aberrations were reversed upon exogenous expression of human *POLQ* (Fig. 4a–c). These observations indicate that TMEJ is a major pathway for ICL repair in mammals that prevents accumulation of mitotic crossovers. Notably, $Polq^{-/-}$ cells treated with MMC also accumulated large Rad51 foci (Fig. 4d–f), similar in character to those observed in *Polq* and *53bp1* DKO cells.

***Polq* is required for pyridostatin induced DNA damage repair**. Pol θ has been implicated in repair of replication-dependent DNA damage at G4 DNA in *C. elegans*[14,15]. Our CRISPR screen identified a synthetic sickness genetic interaction between *Polq* and *Fancj*, which was validated by performing a colony forming assay (Fig. 1d). *Fancj* is a conserved helicase that unfolds G4 DNA[31] and mutations in its *C. elegans* ortholog, *dog-1*, result in high levels of TMEJ signature repair at G4 sites in the genome[32]. We found that $Polq^{-/-}$ MEFs are hypersensitive to the G4 stabilizer pyridostatin (PDS)[33] relative to *WT* cells (Fig. 5a). Similarly, *WT* MEFs transduced with sgRNA targeting the *Polq* polymerase domain induced sensitivity to PDS relative to a control sgRNA (Fig. 5b). $Polq^{-/-}$ cells treated with PDS accumulate a significantly greater number of Rad51 and 53BP1 foci (Fig. 5c–e). Interestingly, Rad51 foci in $Polq^{-/-}$ cells were larger and more frequently adjacent to 53BP1 foci than in *WT* cells (Supplementary Fig. 8). Altogether, these observations demonstrate an essential role for Pol θ in protection against accumulation of non-productive HR intermediates at sites of replication-associated DNA damage.

Elevated TMEJ repair signatures in cells with PolqSL gene mutations. We next investigated whether there was more frequent utilization of TMEJ for DSB repair in cells deficient in genes represented in the PolqSL list. We first induced chromosomal breaks in a wild type MEF line, as well as stable variants of this line deficient in *53bp1* ($53bp1^{-/-}$) or *Brca2* ($Brca2^{Mut/-}$) (Supplementary Fig. 9), and characterized repair of these breaks by high throughput sequencing (Fig. 6a). TMEJ events were defined as deletions >5 bp with >2 bp flanking microhomology (MHD), which is a signature pattern of repair product that has previously been shown to be Pol θ-dependent in this cell line[10]. Both of the PolqSL list gene mutants (*Brca2* and *53bp1*) showed increased use of the TMEJ signature (Fig. 6b, c), although $53bp1^{-/-}$ had longer MHD compared to $Brca2^{Mut/-}$, likely due to increased DSB resection in *53bp1* deficient cells. Similar results were observed using a dPCR assay specific for a Pol θ-dependent MHD in cells that were CRISPR-targeted for two additional HR genes in the PolqSL list, *Palb2* and *Rad51* (Fig. 6d, e). Frequent synthetic lethality with *Polq* deficiency thus tightly correlates with the importance of TMEJ as a commonly used compensatory, or backup mechanism for repair of replication-associated DSBs.

Elevated TMEJ repair signatures in human breast cancers with PolqSL gene alterations. The association between PolqSL gene mutations and increased utilization of TMEJ repair in MEFs led us to hypothesize that human cancers with PolqSL gene mutations may also contain higher levels of TMEJ-associated genomic scars. Towards this end, we identified 275 out of 926 (29.7%) breast cancers in the TCGA cohort[34] as likely deficient in one or more of the 140 PolqSL genes identified in our CRISPR screen (Supplementary Data 4–5), due to a truncating mutation or a deep copy number deletion. Notably, this is a much larger

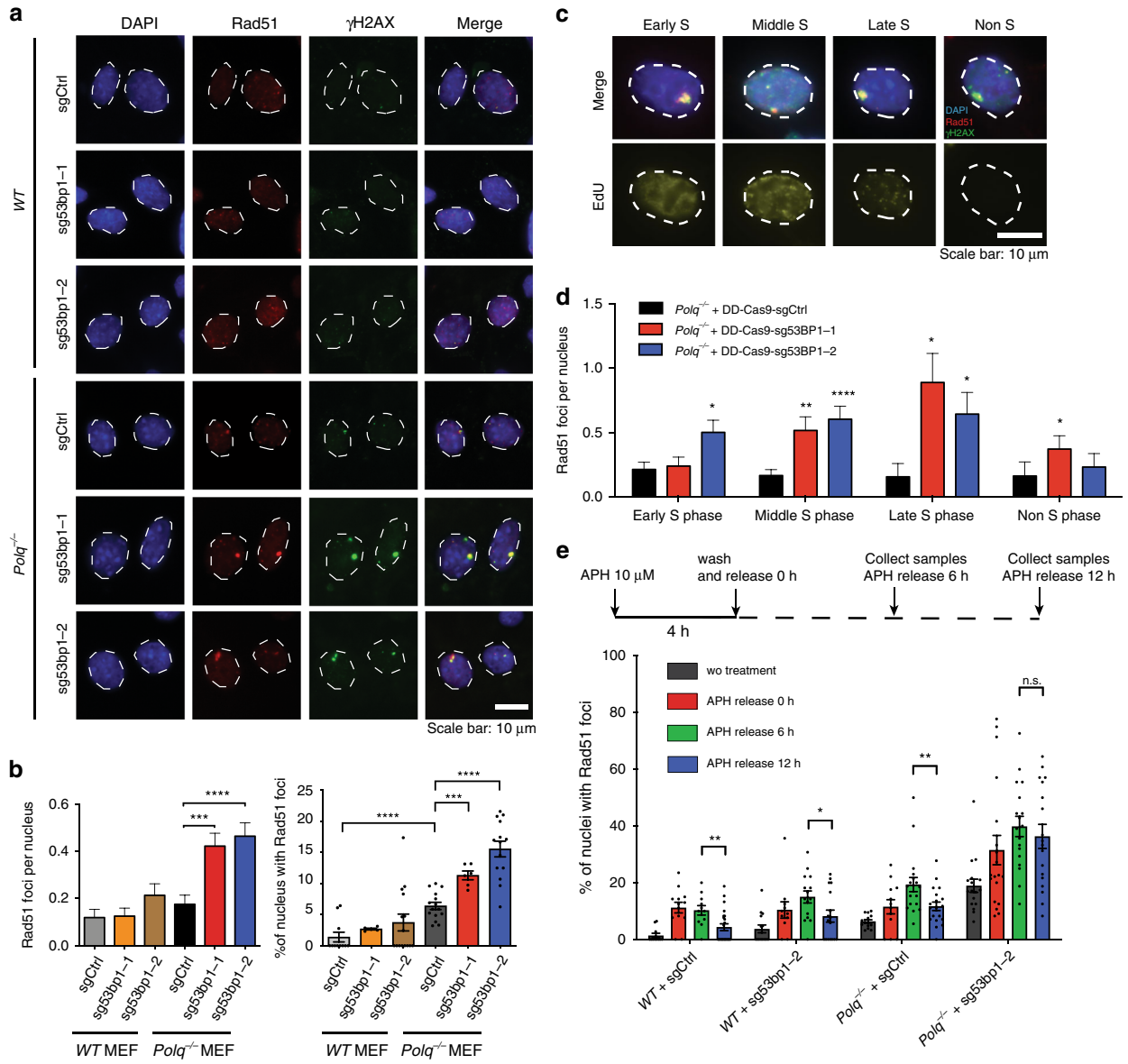


Fig. 3 *Polq/53bp1* double knockout cells accumulate non-productive HR intermediates in S phase. **a** Representative immunofluorescence (IF) staining of Rad51 (red) and γH2AX (green) foci formation in cells with the indicated genotypes, 48 h after Shield1 treatment. ($n = 3$ biologically independent experiments). **b** Quantification of large Rad51 foci per nucleus and percentage of nucleus with Rad51 foci. Data shown are mean ± SEM, and consistent across three independent biological replicates. ***$p < 0.001$; ****$p < 0.0001$ using a Mann–Whitney test. **c** Representative co-IF images for Rad51 (red), γH2AX (green), and EdU (yellow,10 min EdU pulse) in $Polq^{-/-}$ + sg53BP1-2 MEFs to distinguish cell cycle stages as indicated. **d** Quantification of large Rad51 foci per nucleus stratified by cell cycle stage. *$p < 0.05$; **$p < 0.01$; ****$p < 0.0001$ by a Mann–Whitney test. **e** Cells were treated with 10 μM Aphidicolin for 4 h followed by release for 0, 6, and 12 h, fixed cells and stained cells with Rad51. Quantification of the percentage of nuclei with Rad51 foci was performed. Data shown are mean ± SEM, and consistent across two independent biological replicates. *$p < 0.05$; **$p < 0.01$ using a Mann–Whitney test

fraction of cancers than previously considered as having addiction to Pol θ—only 21 of these 275 cases were *BRCA* mutated. We observed significantly higher levels of *POLQ* mRNA (Fig. 6f) in breast cancers with PolqSL gene alteration. We also investigated correlation with COSMIC mutation signature 3, which is upregulated in cancers with *BRCA1/2* deficiency and also in BRCA non-mutant cancers with suspected homologous recombination deficiency (HRD)[35,36]. We observed highly significant enrichment of COSMIC signature 3 in breast cancers with PolqSL gene alterations (Fig. 6g), relative to breast cancers without PolqSL gene alteration. These observations are consistent with excessive employment of Pol θ in PolqSL deficient cancers. We further explored this possibility by implementing a validated algorithm for indel detection[37] to quantify the signature readout of TMEJ repair—microhomology-flanked deletions (MHD), defined as deletion size of 5 bp or greater and 2 bp or more of flanking microhomology. Breast cancers with PolqSL gene alterations were significantly more likely to have a detectable TMEJ signature MHD identified from whole exome sequencing (WES) analyses (Fig. 6h). As expected, whole genome sequencing (WGS) identified a 20-fold higher rate of TMEJ signature MHD than WES in a subset of 94 TCGA breast cancers for which both WES and WGS were performed (Fig. 6i, Supplementary Data 6). Forty one out of 94 (43.6%) breast cancers with WGS data available in TCGA had PolqSL gene alterations, and this subset of cancers had significantly higher levels of TMEJ signature MHD than cancers

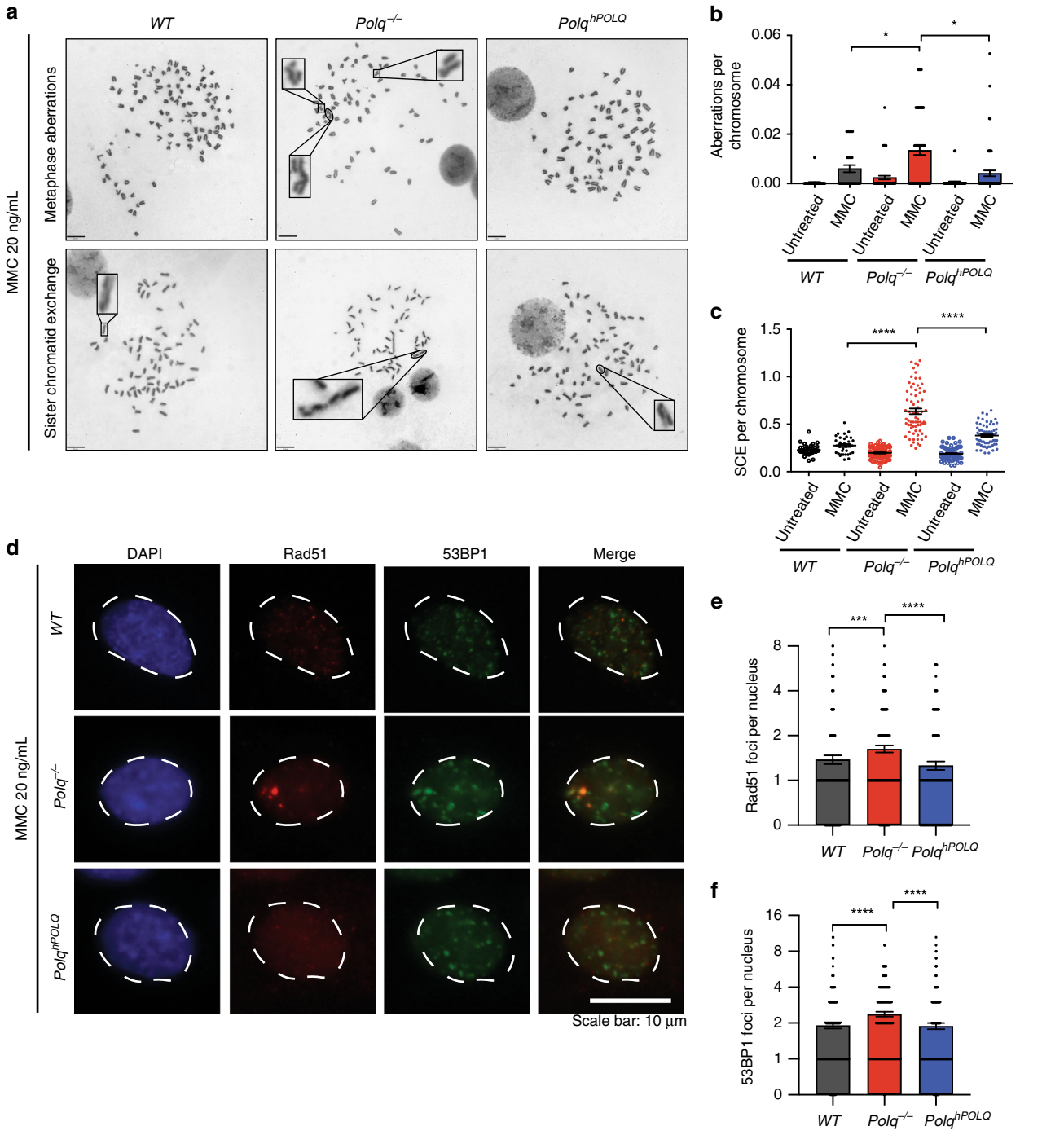


Fig. 4 *Polq* is required for Mitomycin C induced DNA damage repair. **a** Metaphase aberrations and sister chromatid exchanges are shown in *WT*, *Polq*$^{-/-}$ and *Polq*^{hPOLQ} cells 12 h after treatment with 20 ng/mL Mitomycin C (MMC). Scale bar = 10 µm. **b**, **c** Quantification of (**a**), 35 metaphase spreads for each condition was scored, and shown are mean ± SEM. Significance determined using an unpaired, two-tailed *t*-test (*$p < 0.05$; ****$p < 0.0001$). **d** IF analysis of *WT*, *Polq*$^{-/-}$, and *Polq*^{hPOLQ} cells six hours after treatment with 20 ng/mL MMC, stained with DAPI (blue) and antibodies specific for Rad51 (red) and 53BP1 (green) ($n = 3$ biologically independent experiments). (**e**, **f**) are quantification of (**d**). Statistical significance was assessed by unpaired, two-tailed Mann–Whitney tests (***$p < 0.001$; ****$p < 0.0001$)

without PolqSL gene alterations (Fig. 6j). Thus, mimicking our findings in genetically engineered MEFs, we find that human breast cancers with deficiency in PolqSL genes have multiple indices of a hyperactive TMEJ repair pathway.

Discussion

We have defined a surprisingly diverse landscape of DDR gene mutations that renders cells addicted to TMEJ for survival. The functional diversity of PolqSL genes suggests that Pol θ becomes essential upon increased levels of endogenous, unrepaired DNA damage, regardless of the precise nature of that damage. The lack of specificity for a specific type of DNA damage argues against a translesion synthesis function for Pol θ, and is consistent with Pol θ-mediated repair of replication-associated DSBs via TMEJ. Indeed, we found Pol θ is essential for repair of DSBs arising from aphidicolin-induced replication fork collapse (Supplementary Fig. 3).

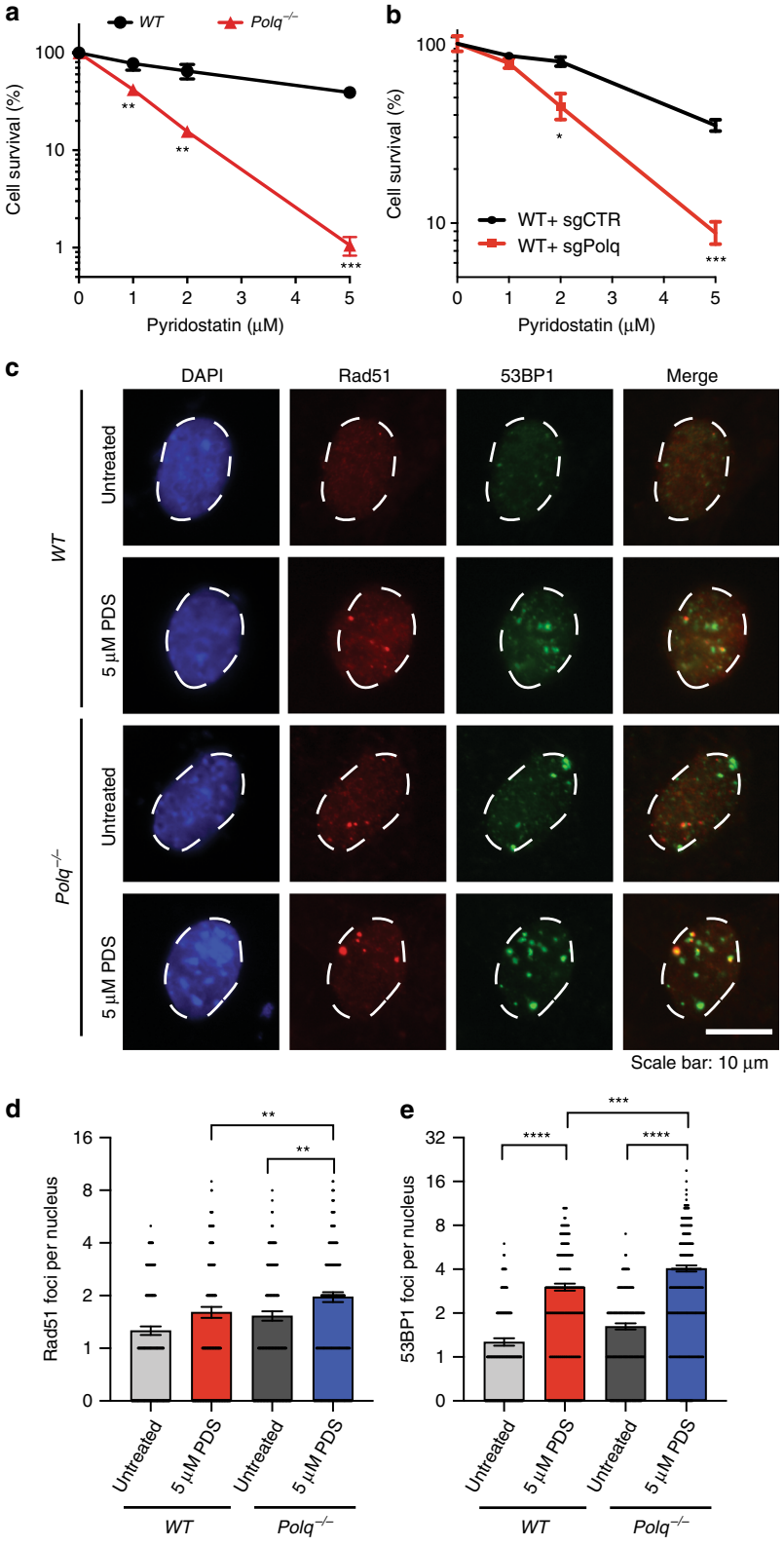


Fig. 5 *Polq* is required for Pyridostatin induced DNA damage repair. **a** Colony forming efficiency after treatment with Pyridostatin (PDS, 1, 2, 5 μM) in *WT* and *Polq*$^{-/-}$ MEFs. **b** Colony forming efficiency after treatment with Pyridostatin (PDS, 1, 2, 5 μM) in *WT* MEFs transduced with Cas9 and either sgCtrl or sgPolq (targeting the polymerase domain). **a**, **b** Data shown are the mean ± SEM ($n = 3$). Statistical significance was assessed by two-tailed *t*-tests. *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$. **c** IF images for DAPI (blue), Rad51 (red) and 53BP1 (green) in *WT* and *Polq*$^{-/-}$ MEFs six hours after treatment with 5 μM PDS ($n = 3$ biologically independent experiments). **d**, **e** are quantification of large Rad51 and 53BP1 foci, as observed in (**c**). Shown are mean ± SEM, and representative of three independent experimental replicates. Statistical significance was assessed by Mann–Whitney tests. **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$

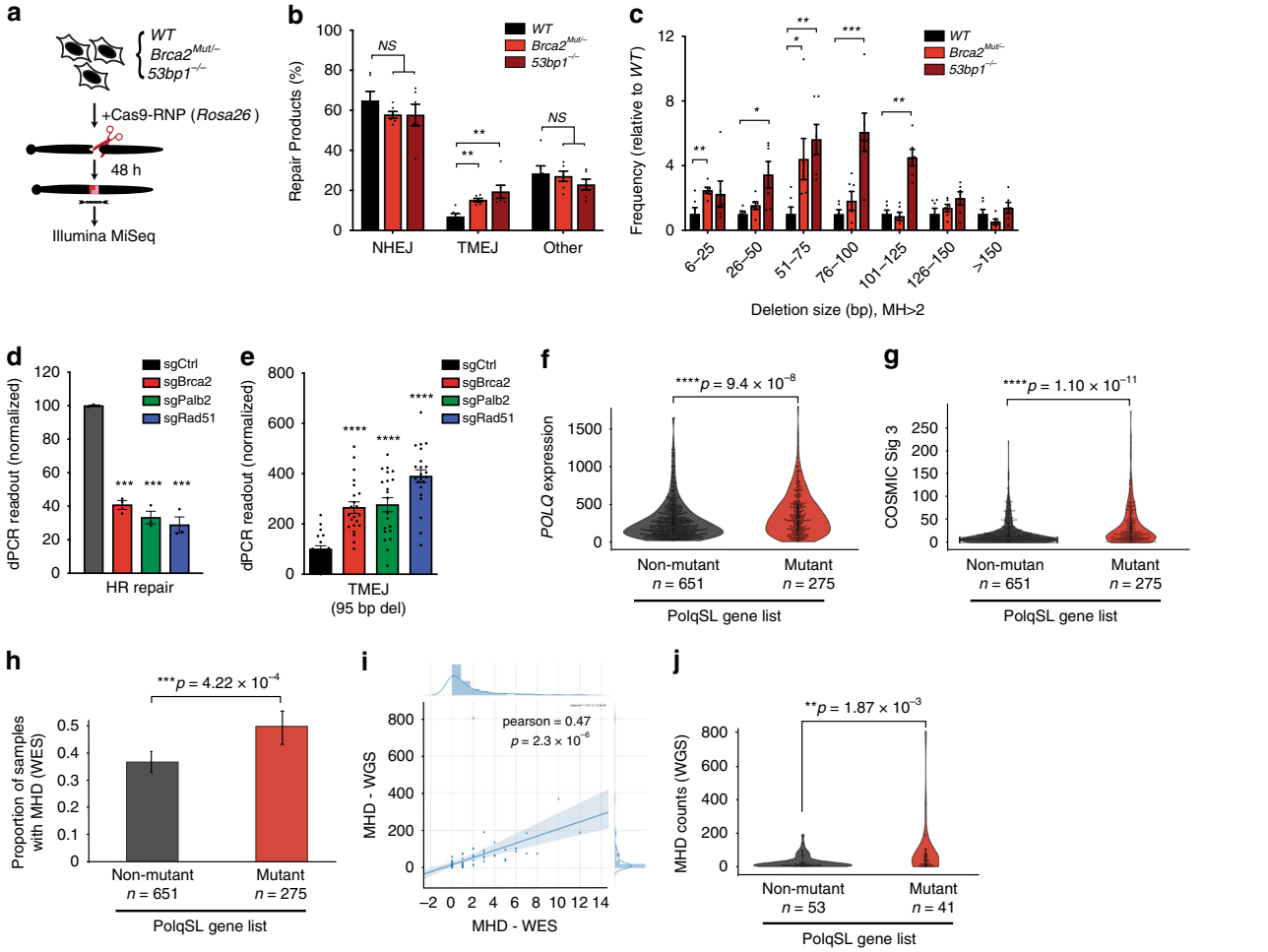


Fig. 6 Elevated TMEJ repair signatures in cells and cancers with PolqSL gene mutations. **a** *WT*, *53bp1*$^{-/-}$, and *Brca2*$^{Mut/-}$ MEFs were transfected with Cas9 ribonucleoprotein (Cas9-RNP) targeting the Rosa26 locus. Forty-eight hours later the targeted region was amplified and analyzed by high throughput sequencing (Illumina MiSeq). **b** Percentage of repair products classified as NHEJ (≤5 bp del, or 1–3 bp insertion), TMEJ (>5 bp deletion and >2 bp MH), and "Other" (>5 bp deletion and 0–2 bp MH), for the indicated MEF genotypes. **c** The relative frequency (normalized to *WT*) of end joining products with >2 bp MH and deletion size within the indicated ranges. *53bp1*$^{-/-}$ MEFs are associated with larger-sized TMEJ signature deletions. **b**, **c** Mean values ± SEM ($n = 6$) are shown. Significance determined using an unpaired, two-tailed *t*-test (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). **d**, **e** DNA repair products are detected by digital PCR using WT MEFs transduced with the indicated PolqSL gene-targeting sgRNA. HR is detected after co-transfection of a homology donor. Relative rates (normalized to *WT* + sgCtrl) of (**d**) homologous recombination (HR) and (**e**) TMEJ (95 bp deletion with 5 bp MH) are indicated as mean values ± SEM ($n = 3$). Significance determined using an unpaired, two-tailed *t*-test (***$p < 0.001$; ****$p < 0.0001$). **f**–**j** Analysis of breast cancers in TCGA. PolqSL mutant breast cancers are identified using cBioPortal as having a truncating mutation or deep copy number deletion. *POLQ* mRNA expression (**f**) and COSMIC Signature 3 (**g**) are elevated in cancers with PolqSL gene alterations. **h**–**j** Whole exome sequencing (WES) and whole genome sequencing (WGS) analyses in TCGA breast cancers to detect MH-flanked deletions (MHD). **h** Breast cancers with PolqSL gene alterations are more likely to have a detectable MHD by WES than breast cancers without alteration in PolqSL genes. The error bar is a bootstrapped 95% confidence interval. WGS data is available for 94 TCGA breast cancers. **i** High correlation between MHD detected by WGS or WES. **j** Significantly higher MHD count, detected by WGS, among breast cancers with PolqSL gene alterations, relative to the non-altered breast cancers. **f**–**g**, **i**–**j** Statistical significance was assessed by two-tailed Mann–Whitney tests. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$

Whilst prior studies suggested that TMEJ primarily functions as a backup pathway to HR and NHEJ, our study identifies numerous examples of TMEJ essentiality when canonical DSB repair pathways are unperturbed. Analysis of synthetic lethality in *53bp1*/*Polq* DKO cells reveals an accumulation of unrepaired HR intermediates in S phase that is further exacerbated by aphidicolin-induced replication fork collapse. We also find TMEJ essentiality upon G quadruplex stabilization and after exposure to interstrand crosslinking agents, both of which promote replication fork stalling and/or collapse. These observations suggest that TMEJ is required for repair of a subset of replication-associated DSBs that is not amenable to repair by HR (Fig. 7). An example of such a break could be one where the template contains a replication-blocking lesion. Recent evidence supports a model wherein unresolved replication-blocking lesions can be inherited as tracts of single-stranded gaps surrounding the lesion[15,38,39]. Replication of these lesions in daughter cells would give rise to two-ended DSBs that are not amenable to HR due to persistence of the replication-blocking lesion in the template DNA strand (Fig. 7, left panel). TMEJ may be a preferred repair mechanism at these sites due to its ability to re-join resected breaks without requiring a homologous template (Fig. 7, right panel). Thus, we postulate that DDR gene mutations that induce a higher prevalence of unresolved replication-blocking lesions may induce TMEJ essentiality. Alternative activities of Pol θ may also be operative at replication-associated DSBs. For example, prior

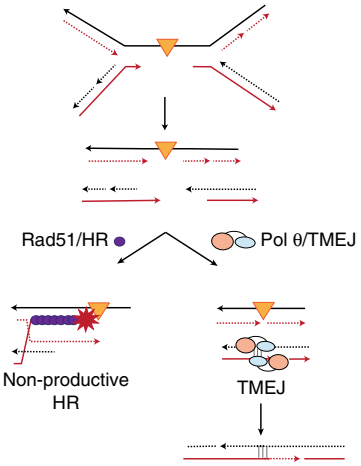


Fig. 7 Model for TMEJ in suppressing non-productive HR at replication-associated DNA damage. Orange triangle indicates a replication-obstructing lesion, such as ICL, G4, or base damage, with an associated region of under-replicated DNA. Converging replication forks will generate a two-ended DSB that can undergo end resection to expose 3′ overhangs. Rad51 loading and attempted HR may result in unsuccessful repair due to persistence of the replication blocking lesion in the homologous template DNA. Alternatively, TMEJ is able to perform microhomology-mediated end joining of the exposed 3′ overhangs, without requiring a homologous template

studies have shown that Pol θ can promote microhomology-mediated integration of plasmid DNA, implying an ability to invade donor templates that lack overt DSBs[40,41]. Future studies will be necessary to unravel the mechanism by which Pol θ resolves stalled or collapsed replication forks, and its relationship to HR-mediated repair.

Our findings also demonstrate that a hallmark feature of cells and cancers with gene mutations that induce Pol θ addiction is an increased prevalence of TMEJ pattern genomic scars. The striking similarity and overlap between genomic scar signatures previously ascribed to HR deficiency (HRD) or "BRCA-ness"[42,43] and TMEJ repair raises the distinct possibility that hyperactive TMEJ is the etiologic driver of this genomic scar pattern in human cancer. While we have shown that HRD induces hyperactive TMEJ, our study also demonstrates that non-HR gene alterations (such as 53BP1) are also sufficient to induce TMEJ hyperactivity and addiction. An important clinical implication of these findings is that only a subset of cancers identified by HRD or BRCA-ness genomic scar signatures may be functionally HR deficient. This may explain the incomplete correlation between HRD signatures and functional HR assays[44], and the recent finding that HRD signatures are unable to accurately predict platinum chemotherapy sensitivity in metastatic breast cancer patients[45]. We propose that genomic classifiers that incorporate COSMIC signature 3, and especially MHD burden, may be more precisely described as "hyper-TMEJ" signatures, rather than signatures specific for HRD.

Our study suggests that cancers with "hyper-TMEJ" signatures may be dependent on Pol θ for their survival. Recent pan-cancer genomic analyses[46,47] suggest that hyper-TMEJ signatures and PolqSL gene deficiency may account for as many as 20% of all human cancers, thus greatly expanding the number of cancers for which Pol θ represents an attractive therapeutic target. Refining the optimal "hyper-TMEJ" genomic scar signature that predicts Pol θ addiction in human cancers will be a clinically relevant area of future investigation.

Methods

Cell culture. *WT* (*Polq*$^{+/+}$), *Polq*$^{-/-}$, *Polm*$^{-/-}$, and *Polm*$^{+/+}$ cells were SV-40 large T antigen immortalized MEFs[12,20,22], while *Polq*^{hPOLQ} MEFs were generated by complemented *Polq*$^{-/-}$ MEFs by human *POLQ* cDNA expression[20]. 293T cells were purchased from ATCC (CRL11268). All cells were maintained in Dulbecco's modified Eagle's medium (DMEM), with 10% Bovine Calf Serum (Hyclone BCS) and 2 mM L-glutamine (ThermoFisher). *Polq*^{hPOLQ} cells were maintained in the same media supplemented with 2 μg/ml puromycin. All cells were maintained at 37 °C in an atmosphere of 5% CO_2. Cells in culture were routinely monitored for mycoplasma contamination using the Plasmo Test™ (Invivogen).

Oligo synthesis and pooled library cloning. DNA oligonucleotide sgGuide library was synthesized by LC Sciences (Supplementary Data 1). A subset of this library was then amplified by PCR using AmpliTaq Gold® 360 DNA Polymerase (ThermoFisher) with forward primer ArrayF and reverse primer ArrayR (Supplementary Data 7) followed by purification with MinElute PCR Purification Kit (Qiagen) to produce a double strand product suitable for Gibson cloning[48]. The CRISPR library cassette was cloned into lentiCRISPR v2 (a gift from Feng Zhang, Addgene plasmid # 52961) followed by transformation into Endura™ Electro-Competent Cells (Lucigen) according to the manufacturer's protocol using BTX Gemini system (ThermoFisher). To ensure no loss of representation, six parallel transformations were performed using the same Gibson reaction and plated into twelve, 10 cm petri dishes (VWR) containing LB agar (ThermoFisher) with 100 μg/ml carbenicillin (ThermoFisher). Colonies were scraped off plates and combined for DNA extraction (Qiagen).

Lentivirus generation. Lentiviruses were generated by 293T cells in 150 mm dish with transfection of 3 μg pMD2.G (a gift from Didier Trono, Addgene # 12259), 4.5 μg psPAX2 (a gift from Didier Trono, Addgene # 12260) and 6 μg custom DDR-CRISPR pooled lentiviral library, transfection was performed using Polyethylenimine (Linear, MW 25,000, Polysciences, Inc)[49,50]. Supernatant from the packaging reaction was collected at 48 h and 72 h. This was pooled and then filtered through 0.45 μm filter. The virus was then concentrated by pelleting at 113,000 × *g* in a SW28 ultracentrifuge rotor for two hours at 4 °C. The pellet was allowed to dissolve overnight in desired volume of PBS at 4 °C and then aliquoted and frozen at −80 °C.

CRISPR library screening. Our custom DDR-CRISPR pooled lentiviral plasmid library containing 3908 sgRNAs targeting 309 murine DNA damage response (DDR) genes with an average of 10 sgRNAs per gene, as well as 834 non-targeting sgRNA controls was used to infect cells at a MOI ~0.8. Twenty-four hours after addition of virus, the media was removed and replaced with fresh media. Forty-eight hours after adding the virus the cells were split. One million cells from each infection was seeded into a 15 cm dish in media containing 2 μg/ml puromycin. Cells were passaged once every two to three days, 1×10^6 cells were reseeded into a 15 cm plate each time. After 8 population doublings, cells were harvested and genomic DNA isolated using QIAamp DNA Blood Kit (Qiagen). To amplify lentiCRISPRv2 sgRNAs, PCR was performed in two steps: For the first PCR, the amount of input genomic DNA (gDNA) for each sample was calculated to achieve 120 X coverage over the DDR-CRISPR library, which resulted in 5 μg DNA per sample (assuming 6.6 μg of gDNA for 10^6 cells). For each sample 5 separate 100 μl PCR reactions with 1 μg genomic DNA in each reaction using Herculase II Fusion DNA Polymerase (Agilent) were carried out using DDR_CRISPR_Ion_1st_FWD and DDR_CRISPR_Ion_1st_REV and then combined. A second PCR was performed to attach Ion adaptors and to barcode samples. The second PCR was done in two 50 μl reactions using around 80 ng product from the first PCR. Primer sequences for the first and second PCR are attached in Supplementary Data 7. Amplification was carried out with 30 cycles for the first PCR and 10 cycles for the second PCR. Twenty-four to twenty-eight libraries were then pooled and sequenced on an Ion S5 (ThermoFisher) using the 530v1 chip.

The "membrane CRISPR library" was obtained from Addgene (ID: 1000000124), and targets 951 genes that encode membrane associated proteins[51]. Pooled CRISPR library lentivirus was transduced into *WT* and *Polq*$^{-/-}$ cells expressing Shield1-inducible DD-Cas9-mVenus. We sorted at least 2×10^6 cells expressing both DD-Cas9 (mVenus) and membrane CRISPR library (mCherry) before adding Shield1 treatment. 24 h after Shield1, cells were passaged once every two to three days, by reseeding 1×10^6 cells into a 15 cm plate each time. We performed the same genomic DNA extraction and library prep and Ion sequencing as our DDR CRISPR library, except the first round primers for membrane CRISPR library amplification were: MMB_CRISPR_Ion_1st_FWD and MMB_CRISPR_Ion_1st_REV (Supplementary Data 7).

CRISPR library analysis. The number and significance of guides present for each library in the multiplexed FASTQ file is determined using our custom algorithm (Völundr). Völundr identifies and counts the sgRNA sequence in the FASTQ reads allowing for a single mismatch in the sequence. It then writes a count file for each library in the pool and a summary file describing the FASTQ file and the libraries. The count files are used by the Völundr target analysis module as the input data for determining which genes are significantly different than the biological control

sample. To accomplish this the counts in each file are first normalized to the total counts for its library (Supplementary Note 1 see Eq. 1). The normalized data from the replicate samples are then combined on a per guide basis by determining the geometric mean for each guide across replicates (Supplementary Note 1 see Eq. 2). At this step the sgRNA TD_{Norm} value and sgRNA Abundance Change Scores are determined as shown in Supplementary Note 1 see Eq. 3. Any guides with no counts in the Plasmid sample are masked from all analysis steps at this point. The "Gene Abundance Change Scores" are determined as in Supplementary Note 1 see Eq. 4. For each sgRNA targeting a gene of interest, "ABC", the experimental sample sgGuide TD_{Norm} value is first subtracted from the corresponding biological sample control sgGuide TD_{Norm} value. The $\log_2$ transformed, geometric mean of this set is the "Gene Abundance Change Scores". These scores are also computed in the next section on the control guides to empirically estimate the distribution of the Gene Abundance Change Scores in the absence of real biological change using a resampling-based scheme.

The Völundr pipeline takes two different approaches to determine if a targeted gene is significantly different than the biological control. The first is to estimate an empirical null distribution for the Gene Abundance Change Scores by randomly sampling ten non-targeting guides (sgControl) sgRNA TD_{Norm} values, from the 834 sgControl guides a total of 100,000 times (Python NumPy, random choice). For each random set of control guides sampled, we repeat the procedure of the prior paragraph to calculate an empirical Gene Abundance Change Score. The 99.99 percentile and 0.01 percentile values are used as the boundaries for the null set. The second method uses the Kolmogorov-Smirnov test to determine if the sgRNA Abundance Change Scores for the genes were drawn from the same population sgRNA Abundance Change Scores for the sgControls. The p-values of the Kolmogorov-Smirnov test (Python Scipy Stats, ks_2samp) are corrected for multiple tests using a false discovery of 3% with the two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli found in GraphPad Prism v7.04. The first test evaluates the observed score distribution across the 10 guides for a gene relative to the empirical null distribution across randomly sampled sets of 10 control guides. The second score evaluates the overall gene sgRNA Abundance Change Score relative to the overall sgControl sgRNA Abundance Change Score. For stringency, we require genes to pass both tests to be reported as significant.

Establishment of mammalian expression constructs and stable cell lines. DNA corresponding to sgRNAs was cloned into LentiCRISPRv2 (a gift from Feng Zhang, Addgene # 52961), or DD-Cas9 (a gift from Raffaella Sordella, Addgene plasmid # 90085), or pGL3-U6-sgRNA-PGK-puromycin (A gift of Xingxu Huang, Addgene # 51133), using the same protocol described above. Cells were incubated with fresh lentivirus for 24 h and then were recovered for another 24 h. Infected cells are selected by 2 µg/µl puromycin or mVenus by flow cytometry.

For 53BP1 and Brca2 mutant cell lines, we used the Alt-R CRISPR-Cas9 system (IDT). We performed transfection using the Neon transfection kit (Invitrogen) according to manufacturer's protocol. Alt-R HiFi Cas9 nuclease, crRNA and tracrRNA were purchased from IDT and were used at the manufacturer's recommended concentration; crRNA is designed using MIT CRISPR (http://crispr.mit.edu). Forty-eight hours after transfection, cells were seeded for single clone outgrowth, PCR screening and Sanger sequencing to confirm gene targeting, and subsequent functional tests.

Synthetic lethal and colony formation assay. Two individual sgRNAs were chosen to target each gene. Forty-eight hours after infection, infected cells were selected by puromycin and counted for colony formation efficiency. Cells were incubated for 7–10 days at 37 °C to allow colony formation. Colonies were stained by Coomassie blue. sgRNA sequences are attached in Supplementary Data 7.

Competitive growth assay. DD-Cas9-sgRNA was transduced into the indicated cell lines by lentivirus infection. Forty-eight hours later, mVenus positive cells were quantified by flow cytometry. Normally, more than 50% cells are mVenus positive cells. Cells were treated with 200 nM Shield1 (Takara) and were collected for flow cytometry at the indicated time points.

Time-Lapse microscopy. Cells stably expressing Proliferating Cell Nuclear Antigen (PCNA)-mCherry were transduced with DD-Cas9-sgRNA. PCNA-mCherry fusion reporter is a gift from Dr. Jeremy Purvis and Hui Chao Xiao. Cells were plated on Cell-Tak (Corning) coated glass-bottom 12-well plates (Cellvis) with Phenol-free DMEM (Invitrogen) supplemented with 10% FBS, and L-glutamine with or without Shield1. Forty-eight hours post plating, cells were image captured every 20 min for 72 h in the mCherry and mVenus fluorescence channels. Fluorescence images were obtained using a Nikon Ti Eclipse inverted microscope with a 40x objective and Nikon Perfect Focus (PFS) system to maintain focus during acquisition period. Cells were maintained at constant temperature (37 °C) and atmosphere (5% CO_2). Image analysis was performed on ImageJ – Fiji.

DNA repair assay. Cell lines used in the assay are indicated in the figure. 2×10^6 cells were transfected with 5 µg pGL3-U6-sgRNA-PGK-puromycin (A gift of Xingxu Huang, Addgene # 51133), 5 µg Flag-Cas9 (A gift of Xingxu Huang, Addgene # 44758), with or without 10 µg HR long donor[10] and 1 µg pEGFP-N2 (Takara) by Neon transfection kit (Invitrogen) using a 1350 V, 30 ms pulse in a 100 µL chamber. Forty-eight hours post transfection, a portion of the cells were analyzed by flow cytometry to quantify the transfection efficiency, and the remaining cells were harvested for genomic DNA extraction (Qiagen). Digital PCR (QX-200, Bio-Rad) was performed to quantify the frequency of gene conversion events using the primers and Taqman probes listed in Supplementary Data 7. The repair signal was normalized to 5000 copies of genomic DNA, measured using a Chromosome 6 control dPCR assay, using primers/probes sequences listed in Supplementary Data 7. Analysis of dPCR data was performed using QuantaSoft (Bio-Rad).

Immunofluorescence. Cells were permeabilized by CSK buffer (10 mM Hepes, 300 mM Sucrose, 100 mM NaCl, 3 mM MgCl2, and 0.5% Triton X-100, pH = 7.4) for 2 min followed by fixation for 15 min in 3% paraformaldehyde. Cells were subsequently processed for immunostaining experiments using the indicated antibodies. Nuclei were visualized by staining with DAPI. The primary antibodies used were: Rad51 (1:500, Novus Biologicals, NB100-148), γH2AX (1:500, Trevigen, 4418-APC-100), and 53BP1 (1:500 for immunofluorescence, 1:5000 for western blot, Bethyl, A300-272A). The secondary antibodies were: Rhodamine Goat Anti-Mouse IgG (H + L) (1:500, Jackson ImmunoResearch, 115-025-146) and FITC Goat Anti Rabbit IgG (H + L) (1:500, Jackson ImmunoResearch, 111-095-144). For S phase stain, we incubated cells with 10 µM EdU for 10 min, EdU was detected according to the EdU-Click 647 kit protocol (baseclick). Images were acquired using an Olympus BX61 fluorescence microscope or Zeiss 880 with Airyscan processing for the super-resolution images.

Metaphase and sister chromatin exchange assay. Metaphases were prepared by a previously published method[52] with the noted changes. Cells were treated with 100 ng/ml of Colcemid (KaryoMAX® 15210-040 from Gibco) for 1 h prior to harvest and swelling in 75 mM potassium chloride for 20 min at 37 °C. Once the metaphases were dropped onto slides, the slides were stored at room temperature for at least two days prior to staining with Giemsa (KaryoMAX® 10092-013 from Gibco) for 2–3 min. After staining, the slides are rinsed with distilled water and allowed to air dry completely before mounting the coverslips with DPX Mountant (Millipore Sigma). Spreads were imaged under a 100× objective using an Olympus BX61 Light Microscope with QImaging RETIGA 4000R camera.

The SCE assay was performed as previously described[53]. Briefly, 24 h after cells were plated, 10 µM bromodeoxyuridine (BrdU) (Millipore-Sigma) was added to the plates for 24 h. MMC (20 ng/ml, Millipore-Sigma) was added for the final 12 h in BrdU. For the final hour 100 ng/ml of Colcemid (KaryoMAX® 15210-040 from Gibco) was added to the media. The cells were harvested by trypsinization and processed for metaphase spreads as described above. After 2–3 days the metaphases were stained for 30 min by placing the slides in a Coplin jar containing 10 µg/ml Hoechst 33342 (Thermo Fisher) in PBS. The slides were then removed from the Hoechst solution and placed in a tray of 2 × SSC (20 × SSC Stock: 3 M sodium chloride with 300 mM sodium citrate) on a 45 °C heat block. While in the warm SSC the metaphases where exposed to UVB radiation from a Danmar UVB compact fluorescent bulb (peak emission 365 nM) at a distance of 5 cm for 20 min. After exposure to UVB place the slides in a Coplin jar of 2 × SSC for ≥10 min to let the Hoechst and degraded DNA wash away. The slides were then stained with Giemsa (KaryoMAX® 10092-013 from Gibco) for 5 min. The slides were cover slipped and imaged as described above for metaphases.

High throughput sequencing. Two hundred nanograms of genomic DNA were amplified using a two-step PCR that added unique library bar-codes, heterogeneity spacers and Illumina MiSeq adapters (as in[54]). Two-step PCR primers are attached in Supplementary Data 7. Samples were sequenced using a 2 × 300 MiSeq kit[55]. Quantification and classification of the sequences was done in R and excel.

Analysis of MHD in human breast cancers. TCGA WES reads were aligned using bwa-mem (Li, 2013, arXiv:1303.3997 [q-bio.GN]) and realigned using ABRA2 (https://github.com/mozack/abra2)[37]. Read duplicates were marked by biobambam2 (https://github.com/gt1/biobambam2[56],). Variants were called with Strelka[57], UNCeqR[58] and Cadabra[37] (https://cadabra.science). Variant calls were annotated with the Variant Effect Predictor (VEP)[59]. For TCGA cases where WGS data were available, reads were trimmed using SeqPurge[60], aligned using bwa-mem and realigned using ABRA2. Read duplicates were marked by biobambam2 and indels were called using Cadabra. Variant calls were annotated with VEP. TCGA mRNA reads were aligned to human reference genome hg38 using STAR[61] and quantified with Salmon[62] which was run against STAR's transcriptome alignments. Quantification values were upper quantile normalized.

MHD were defined as deletions ≥ 5 bp in length with ≥ 2 bp of flanking microhomology. To determine flanking microhomology, the 3′ end of the deletion was matched to the sequence directly upstream of the deletion junction. Deletions located in regions enriched in short repeats were ignored. MHD in TCGA breast cancer samples were summed per sample with a sample being classified as having MHD if at least one occurrence of MHD was observed.

PolqSL mutant TCGA samples were identified as having at least one deep copy number deletion or truncating mutation in a PolqSL gene using cBioPortal. POLQ

mRNA expression, COSMIC mutation signature 3 scores[63], and proportion of samples with MHD were compared between PolqSL mutant and non-mutant groups using a two-tailed Mann–Whitney test.

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

Data availability

Sequencing data is available at [https://www.ncbi.nlm.nih.gov/sra/PRJNA556352]. Unanalyzed raw data is available at [https://figshare.com/projects/Genetic_Determinants_of_Cellular_Addiction_to_DNA_Polymerase_Theta/67331]. All data is available from the corresponding author upon request. The source data underlying Figs. 1d, 2b, d, f, g, 3b, d, e, 4b, c, e, f, 5a, b, d, e, 6b–e and Supplementary Figs. 1, 3, 5, 6, 7a and 8 are provided as a source data file.

Code availability

Code for Völundr has been made publicly available at [https://github.com/pkMyt1/Volundr]. Other software for statistical analysis is publicly available and referenced as noted.

Received: 20 December 2018 Accepted: 22 August 2019
Published online: 19 September 2019

References

1. Wood, R. D. & Doublié, S. DNA polymerase theta (POLQ), double-strand break repair, and cancer. *DNA Repair* **44**, 22–32 (2016).
2. Chapman, J. R., Taylor, M. R. & Boulton, S. J. Playing the end game: DNA double-strand break repair pathway choice. *Mol. Cell* **47**, 497–510 (2012).
3. Kass, E. M., Moynahan, M. E. & Jasin, M. When genome maintenance goes badly awry. *Mol. Cell* **62**, 777–787 (2016).
4. Mehta, A. & Haber, J. E. Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb. Perspect. Biol.* **6**, a016428 (2014).
5. Chang, H. H. Y., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* **18**, 495–506 (2017).
6. Sallmyr, A. & Tomkinson, A. E. Repair of DNA double-strand breaks by mammalian alternative end-joining pathways. *J. Biol. Chem.* **293**, 10536–10546 (2018).
7. Sfeir, A. & Symington, L. S. Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway? *Trends Biochem Sci.* **40**, 701–714 (2015).
8. Chan, S. H., Yu, A. M. & McVey, M. Dual roles for DNA polymerase theta in alternative end-joining repair of double-strand breaks in *Drosophila*. *PLoS Genet* **6**, e1001005 (2010).
9. van Schendel, R., Roerink, S. F., Portegijs, V., van den Heuvel, S. & Tijsterman, M. Polymerase theta is a key driver of genome evolution and of CRISPR/Cas9-mediated mutagenesis. *Nat. Commun.* **6**, 7394 (2015).
10. Wyatt, D. W. et al. Essential roles for polymerase theta-mediated end joining in the repair of chromosome breaks. *Mol. Cell* **63**, 662–673 (2016).
11. Roerink, S. F., van Schendel, R. & Tijsterman, M. Polymerase theta-mediated end joining of replication-associated DNA breaks in *C. elegans*. *Genome Res.* **24**, 954–962 (2014).
12. Shima, N., Munroe, R. J. & Schimenti, J. C. The mouse genomic instability mutation chaos1 is an allele of Polq that exhibits genetic interaction with Atm. *Mol. Cell Biol.* **24**, 10381–10389 (2004).
13. Mateos-Gomez, P. A. et al. Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination. *Nature* **518**, 254–257 (2015).
14. Koole, W. et al. A polymerase theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. *Nat. Commun.* **5**, 3216 (2014).
15. Lemmens, B., van Schendel, R. & Tijsterman, M. Mutagenic consequences of a single G-quadruplex demonstrate mitotic inheritance of DNA replication fork barriers. *Nat. Commun.* **6**, 8909 (2015).
16. Harris, P. V. et al. Molecular cloning of *Drosophila* mus308, a gene involved in DNA cross-link repair with homology to prokaryotic DNA polymerase I genes. *Mol. Cell Biol.* **16**, 5764–5771 (1996).
17. Muzzini, D. M., Plevani, P., Boulton, S. J., Cassata, G. & Marini, F. Caenorhabditis elegans POLQ-1 and HEL-308 function in two distinct DNA interstrand cross-link repair pathways. *DNA Repair* **7**, 941–950 (2008).
18. Ceccaldi, R. et al. Homologous-recombination-deficient tumours are dependent on Poltheta-mediated repair. *Nature* **518**, 258–262 (2015).
19. Higgins, G. S. & Boulton, S. J. Beyond PARP-POLtheta as an anticancer target. *Science* **359**, 1217–1218 (2018).
20. Yousefzadeh, M. J. et al. Mechanism of suppression of chromosomal instability by DNA polymerase POLQ. *PLoS Genet.* **10**, e1004654 (2014).
21. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
22. Pryor, J. M. et al. Essential role for polymerase specialization in cellular nonhomologous end joining. *Proc. Natl Acad. Sci. USA* **112**, E4537–E4545 (2015).
23. Klattenhoff, A. W. et al. Loss of NEIL3 DNA glycosylase markedly increases replication associated double strand breaks and enhances sensitivity to ATR inhibitor in glioblastoma cells. *Oncotarget* **8**, 112942–112958 (2017).
24. Bunting, S. F. et al. 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell* **141**, 243–254 (2010).
25. Senturk, S. et al. Rapid and tunable method to temporally control gene editing based on conditional Cas9 stabilization. *Nat. Commun.* **8**, 14370 (2017).
26. Chao, H. X. et al. Orchestration of DNA damage checkpoint dynamics across the human cell cycle. *Cell Syst.* **5**, 445–459, e445 (2017).
27. Feng, W. & Jasin, M. BRCA2 suppresses replication stress-induced mitotic and G1 abnormalities through homologous recombination. *Nat. Commun.* **8**, 525 (2017).
28. Lukas, C. et al. 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. *Nat. Cell Biol.* **13**, 243–253 (2011).
29. Mankouri, H. W., Huttner, D. & Hickson, I. D. How unfinished business from S-phase affects mitosis and beyond. *EMBO J.* **32**, 2661–2671 (2013).
30. Beagan, K. et al. *Drosophila* DNA polymerase theta utilizes both helicase-like and polymerase domains during microhomology-mediated end joining and interstrand crosslink repair. *PLoS Genet.* **13**, e1006813 (2017).
31. Wu, C. G. & Spies, M. G-quadruplex recognition and remodeling by the FANCJ helicase. *Nucleic Acids Res.* **44**, 8742–8753 (2016).
32. Castillo Bosch, P. et al. FANCJ promotes DNA synthesis through G-quadruplex structures. *EMBO J.* **33**, 2521–2533 (2014).
33. Rodriguez, R. et al. Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol.* **8**, 301–310 (2012).
34. Cancer Genome Atlas, N. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
35. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
36. Nik-Zainal, S. et al. Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).
37. Mose, L. E., Wilkerson, M. D., Hayes, D. N., Perou, C. M. & Parker, J. S. ABRA: improved coding indel detection via assembly-based realignment. *Bioinformatics* **30**, 2813–2815 (2014).
38. Moreno, A. et al. Unreplicated DNA remaining from unperturbed S phases passes through mitosis for resolution in daughter cells. *Proc Natl. Acad. Sci. USA* **113**, E5757–E5764 (2016).
39. Spies, J. et al. 53BP1 nuclear bodies enforce replication timing at under-replicated DNA to limit heritable DNA damage. *Nat. Cell Biol.* **21**, 487–497 (2019).
40. Saito, S., Maeda, R. & Adachi, N. Dual loss of human POLQ and LIG4 abolishes random integration. *Nat. Commun.* **8**, 16112 (2017).
41. Zelensky, A. N., Schimmel, J., Kool, H., Kanaar, R. & Tijsterman, M. Inactivation of Pol theta and C-NHEJ eliminates off-target integration of exogenous DNA. *Nat. Commun.* **8**, 66 (2017).
42. Davies, H. et al. HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. *Nat. Med* **23**, 517–525 (2017).
43. Lord, C. J. & Ashworth, A. BRCAness revisited. *Nat. Rev. Cancer* **16**, 110–120 (2016).
44. Tumiati, M. et al. A functional homologous recombination assay predicts primary chemotherapy response and long-term survival in ovarian cancer patients. *Clin. Cancer Res.* **24**, 4482–4493 (2018).
45. Tutt, A. et al. Carboplatin in BRCA1/2-mutated and triple-negative breast cancer BRCAness subgroups: the TNT Trial. *Nat. Med.* **24**, 628–637 (2018).
46. Knijnenburg, T. A. et al. Genomic and molecular landscape of DNA damage repair deficiency across The Cancer Genome Atlas. *Cell Rep.* **23**, 239–254, e236 (2018).
47. Riaz, N. et al. Pan-cancer analysis of bi-allelic alterations in homologous recombination DNA repair genes. *Nat. Commun.* **8**, 857 (2017).
48. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
49. Gupta, G. P. et al. The Mre11 complex suppresses oncogene-driven breast tumorigenesis and metastasis. *Mol. Cell* **52**, 353–365 (2013).
50. Zufferey, R. et al. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J. Virol.* **72**, 9873–9880 (1998).
51. Morgens, D. W. et al. Genome-scale measurement of off-target activity using Cas9 toxicity in high-throughput screens. *Nat. Commun.* **8**, 15178 (2017).

52. Filatov, L. et al. Chromosomal instability is correlated with telomere erosion and inactivation of G2 checkpoint function in human fibroblasts expressing human papillomavirus type 16 E6 oncoprotein. *Oncogene* **16**, 1825–1838 (1998).
53. Wolff, S. & Perry, P. Differential Giemsa staining of sister chromatids and the study of chromatid exchanges without autoradiography. *Chromosoma* **48**, 341–353 (1974).
54. de Muinck, E. J., Trosvik, P., Gilfillan, G. D., Hov, J. R. & Sundaram, A. Y. M. A novel ultra high-throughput 16S rRNA gene amplicon sequencing library preparation method for the Illumina HiSeq platform. *Microbiome* **5**, 68 (2017).
55. Waters, C. A. et al. The fidelity of the ligation step determines how ends are resolved during nonhomologous end joining. *Nat. Commun.* **5**, 4286 (2014).
56. Tischler, G. & Leonard, S. biobambam: tools for read pair collation based algorithms on BAM files. *Source Code Biol. Med.* **9**, 13 (2014).
57. Saunders, C. T. et al. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* **28**, 1811–1817 (2012).
58. Wilkerson, M. D. et al. Integrated RNA and DNA sequencing improves mutation detection in low purity tumors. *Nucleic Acids Res.* **42**, e107 (2014).
59. McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
60. Sturm, M., Schroeder, C. & Bauer, P. SeqPurge: highly-sensitive adapter trimming for paired-end NGS data. *BMC Bioinforma.* **17**, 208 (2016).
61. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
62. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
63. Polak, P. et al. A mutational signature reveals alterations underlying deficient homologous recombination repair in breast cancer. *Nat. Genet* **49**, 1476–1486 (2017).

Acknowledgements

This work is dedicated to the memory of Alexander Kenan, who was instrumental in the design of the DDR-CRISPR library. We thank S. Kumar, H.X. Chao, M. Kapustina, V. Roberts, and S. Connor for technical assistance and/or data analysis support. We are grateful to Gupta and Ramsden Lab members for helpful discussions. UNC Core labs (Microscopy Services Laboratory, Hooker Imaging Core, Flow Cytometry Core Facility, and High Throughput Sequencing Facility) used in this study are supported in part by P30 CA016086 Cancer Center Core Support Grant to the UNC Lineberger Comprehensive Cancer Center. G.P.G. holds a Career Award for Medical Scientists from the Burroughs Wellcome Fund. Additional funding support was provided by the University Cancer Research Fund (G.P.G., D.A.R.), NCI/NIH (CA222092, D.A.R. and G.P.G., CA193124, R.D.W.), Dept of Defense (W81XWH-18-1-0047, G.P.G. and D.A.R.), the Grady F. Saunders, PhD Distinguished Research Professorship (R.D.W.) and NCI P30 CA016086 (J.S.P.).

Author contributions

G.P.G. conceived and supervised the study. G.P.G., D.A.R. and W.F. designed the experiments. R.D.W. provided critical reagents and project guidance. W.F. performed the CRISPR screens, D.A.S. developed and implemented the bioinformatics analysis pipeline, and N.R. provided statistical input and review. Unless otherwise stated, W.F. performed all additional experiments and data analyses, with statistical review by N.R. R.J.K. performed live cell imaging experiments with guidance from J.E.P. D.A.S. performed mitotic aberration and SCE experiments. J.C.G. performed Rosa26 high throughput sequencing and data analysis with supervision from D.A.R. B.A.P. and L.E.M. performed analyses of human breast cancer genomic datasets with supervision from J.S.P. G.P.G., D.A.S., W.F. and D.A.R. wrote the manuscript, with contributions from all authors. All authors read and accepted the manuscript.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-12234-1.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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