



Published in final edited form as:

Cancer Res. 2023 April 14; 83(8): 1345–1360. doi:10.1158/0008-5472.CAN-22-3667.

TONSL is an immortalizing oncogene and a therapeutic target in breast cancer

Aditi S Khatpe^{1,2,\$}, **Rebecca Dirks**^{1,\$}, **Poornima Bhat-Nakshatri**¹, **Henry Mang**¹, **Katie Batic**¹, **Sarah Swiezy**¹, **Jacob Olson**³, **Xi Rao**⁴, **Yue Wang**⁴, **Hiroshi Tanaka**⁴, **Sheng Liu**⁴, **Jun Wan**^{4,5}, **Duoqiao Chen**^{4,5}, **Yunlong Liu**^{4,5}, **Fang Fang**⁶, **Sandra Althouse**⁷, **Emily Hulsey**⁸, **Maggie M Granatir**⁸, **Rebekah Addison**⁸, **Constance J. Temm**⁸, **George Sandusky**⁸, **Audrey Lee-Gosselin**⁹, **Kenneth Nephew**⁶, **Kathy D. Miller**¹⁰, **Harikrishna Nakshatri**^{1,2,5,11,*}

¹Department of Surgery, Indiana University School of Medicine, Indianapolis, IN 46202, USA

²Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

³Decatur Central High School, Indianapolis, IN 46221, USA

⁴Department of Medical and Molecular Genetics, Indiana University School of Medicine, IN 46202, USA

⁵Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, IN 46202, USA

⁶Medical Science Program, Indiana University School of Medicine, Bloomington, IN 47405, USA

⁷Department of Biostatistics and Health Data Science, Indiana University School of Medicine, IN 46202, USA

⁸Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, IN 46202, USA

⁹Stark Neurosciences Research Institute, Indiana University School of Medicine, IN 46202, USA

¹⁰Department of Medicine, Indiana University School of Medicine, Indianapolis, IN 46202, USA

¹¹VA Roudebush Medical Center, Indianapolis, IN 46202, USA

Abstract

Study of genomic aberrations leading to immortalization of epithelial cells has been technically challenging due to the lack of isogenic models. To address this, we utilized healthy primary breast luminal epithelial cells of different genetic ancestry and their hTERT-immortalized counterparts to identify transcriptomic changes associated with immortalization. Elevated expression of TONSL (Tonsoku Like, DNA Repair Protein) was identified as one of the earliest events during immortalization. TONSL, which is located on chromosome 8q24.3, was found to be

* **Corresponding author:** Harikrishna Nakshatri, BVSc., PhD, C218C, 980 West Walnut Street, Indianapolis, IN 46202, USA, hnakshat@iupui.edu.

^{\$}Contributed equally to this work

Competing interests: The authors declare that they have no competing interests.

amplified in ~20% of breast cancers. TONSL alone immortalized primary breast epithelial cells and increased telomerase activity, but overexpression was insufficient for neoplastic transformation. However, TONSL-immortalized primary cells overexpressing defined oncogenes generated estrogen receptor-positive adenocarcinomas in mice. Analysis of a breast tumor microarray with ~600 tumors revealed poor overall and progression free survival of patients with TONSL overexpressing tumors. TONSL increased chromatin accessibility to pro-oncogenic transcription factors including NF- κ B and limited access to the tumor suppressor p53. TONSL overexpression resulted in significant changes in the expression of genes associated with DNA repair hubs, including upregulation of several genes in the homologous recombination (HR) and Fanconi Anemia pathways. Consistent with these results, TONSL overexpressing primary cells exhibited upregulated DNA repair via HR. Moreover, TONSL was essential for growth of TONSL-amplified breast cancer cell lines *in vivo*, and these cells were sensitive to TONSL-FACT complex inhibitor CBL0137. Together, these findings identify TONSL as a regulator of epithelial cell immortalization to facilitate cancer initiation and as a target for breast cancer therapy.

Keywords

Chromosome 8q24.3 amplicon; TONSL; immortalization; curaxin; breast cancer

Introduction

Around 80–90% of all cancers are carcinomas, malignancies of epithelial tissue, and one of the most widespread human cancers in females arise from the normal breast epithelium (1). Normal epithelial cells have limited replicative potential and the first step in tumor initiation is to overcome this limitation (2). *In vitro*, the cells acquire the unlimited replicative potential through activation of telomerase enzyme by the process called immortalization (2). *In vivo*, a single catastrophic genomic event called chromothripsis may initiate tumorigenesis (3). Error-prone DNA repair pathways activated as a consequence of chromothripsis could lead to inactivation of tumor suppressors, activation of oncogenes, immortalization, transformation and clinical manifestation of the disease as either locally advanced or metastatic cancers (3).

Recent advances in genomics have enabled identification of cancer-enriched genome aberrations and molecular classification of cancers. For example, based on copy number variations (CNVs), breast cancers have been classified into ten integrative clusters, each with distinct outcome profiles (4). Several breast cancer-enriched mutations have been identified, including a limited number of driver mutations in genes such as TP53 and PIK3CA (5). However, breast cancer is predominantly a CNV-driven disease (5). Although mechanistic studies on breast cancer-enriched mutations and CNVs have been successful in identifying downstream signaling pathways and in elucidating the role of signaling pathways in cancer progression and metastasis, the role of these CNVs in cancer initiation is unknown (6,7). Limited progress in this direction is primarily due to lack of an isogenic model system that would allow comparison of primary cells to their immortalized and transformed counterparts.

We recently developed an assay to propagate primary breast epithelial cells with luminal characteristics from core breast biopsies of healthy donors and an isogenic model permitting dissection of molecular events that occur during immortalization, primary tumor growth, and metastasis (8,9). Since genetic ancestry has been shown to influence cancer initiation and progression process (10), our model system included cells from donors of different genetic ancestry, allowing us to identify molecular events during immortalization/transformation of cells from diverse group. By using this approach, we identified upregulation of components of TONSL (Tonsoku Like, DNA Repair Protein)-FACT (Facilitates Chromatin Transcription) complex during immortalization. As a component of the MMS22L-TONSL complex, TONSL interacts with minichromosome maintenance (MCM), FACT, and replication protein A (RPA), binds histones, and controls homologous recombination (HR) during replication-associated DNA damage (11). MMS22L and TONSL participate in the recovery from replication stress by identifying post-replicative chromatin (12,13). TONSL ankyrin repeat domain (ARD) identifies unmethylated lysine-20 residue on histone H4 (H4K20me0) and binds to the histone as a post-replicative chromatin mark during replication(13). TONSL is part of cell cycle-dependent HR (14) and maintains genomic stability during S phase (15). Since innate cellular mechanisms that regulate replicative potential serve as guardians against malignancy, deregulation of these cellular mechanisms could be the initial event in tumorigenesis (16). We show that TONSL, located on chromosome 8q24.3, is amplified in ~20% of breast cancer. We further report that TONSL is an immortalizing oncogene, and, upon upregulation, TONSL manipulates the cells to increase DNA repair via HR. We demonstrate that FACT targeting drugs such as Curaxins (17) inhibit TONSL amplified breast cancers, identifying TONSL as a new therapeutic target in breast cancer.

Materials and Methods

Primary cell culture and immortalization:

Fresh or cryopreserved, deidentified normal breast tissues from healthy women of European, African or Latina ancestry, donated to Komen Tissue Bank (KTB) at Indiana University, were processed to generate primary breast epithelial cells as described previously (8,9). All tissue samples were collected following a detailed IRB approved protocol, with written informed consent from donors, and HIPAA compliance protocol. hTERT, TONSL, HRAS^{G12V}, SV40 Large + small T antigen, or TP53^{R273C}-GFP overexpression was achieved using lentiviral transduction (9). Plasmids used are described in supplementary methods.

TONSL Knockdown using shRNA:

shTONSL viral particles used are described in supplementary methods. User manual was followed to achieve knockdown.

RNA isolation and quantitative real-time polymerase chain reaction (qRT-PCR):

Total RNA was isolated using total RNA isolation kit followed by cDNA synthesis and later qRT-PCR as described previously (9). Further details are provided in supplementary methods.

RNA sequencing:

RNA-seq data to determine genes differentially expressed between primary and hTERT immortalized cells were described previously (9) and are available publicly (GEO number: GSE108541). Genes differentially expressed in immortalized cells compared to primary cells were identified (Table S1). Details of RNA-seq of 1) primary CD49f-/EpCAM+ mature luminal cells - CD49f+/EpCAM+ luminal progenitor cells, 2) KTB103 primary and TONSL overexpressing cells, and 3) TMD436 shControl (pLKO), TMD436 shTONSL Clones 1, 2 and 3 are described in supplementary methods with accession numbers.

Selection of putative immortalization-associated genes for further analysis:

A schematic view for genes specifically deregulated during immortalization is shown in Fig. 1A. Genes that are differentially expressed in mature luminal cells compared to luminal progenitor cells (Table S2) were excluded in the analysis. CRISPR essentiality screen data, described in Table S3, were used for refinement. More details are provided in supplementary methods. Table 1 provides a list of top 15 genes that met all section criteria.

ATAC sequencing:

KTB103 primary and TONSL overexpressing KTB103 cells were subjected to ATAC-seq using the previously established protocol (18). Assays were done in biologic triplicates with ~50,000 cells. Integration of RNA-seq data with ATAC-seq data and motif enrichment analyses were performed as described previously (18) (Accession number GSE216237).

Antibodies and western blotting analysis:

Cell lysates prepared in radioimmunoassay buffer were analyzed by western blotting as described previously (9). Antibodies used are listed in supplementary methods.

Flow cytometry analysis:

Flow cytometry analysis primary and TONSL overexpressing cells was performed as described previously (9). Data were acquired using a BD LSR II flow cytometer and analyzed using FlowJo software. Detail description of antibodies used is in supplementary methods.

Cell Proliferation assay:

2000 cells/well were plated in 96 well plate. Cells were treated with CBL0137 for 48h. Bromodeoxyuridine (BrDU) incorporation-ELISA was done using kit (description in supplementary methods) as per the manufacturer's instruction.

TRAP Assay:

The assay was performed as described in user manual of the kit and additional details are provided in supplementary methods.

Breast tumor tissue microarray (TMA):

A TMA with breast tumor samples from ~600 patients with ~15 years of follow up has been described recently (19). All tissue samples were collected following a detailed IRB

approved protocol, with written informed patient consent, and HIPAA compliance protocol. TMA staining and quantification have been described previously (19).

Immunohistochemistry (IHC):

H&E, ER α , PR, GATA3, and FOXA1 immunostaining was performed at the CLIA-certified Indiana University Health Pathology Laboratory and the whole-slide digital imaging system of Aperio (ScanScope CS) was used for imaging.

Immunofluorescence and microscopy:

For immunofluorescence, 10,000 cells were plated on 35mm glass bottom plates overnight and treated with 10 mM Hydroxyurea for 0h and 6h. Cells were fixed with 4% paraformaldehyde for 20 min at room temperature (RT). Cells were permeabilized and blocked with 22.52 mg/ml glycine in PBS+ 2% FBS + 1% Triton \times for 20 min at RT. Primary and secondary antibodies were diluted in 1% serum in PBST (PBS + 1% Triton \times). Cells were incubated in primary antibody at 4 $^{\circ}$ C overnight followed by incubation with secondary antibody with Hoechst for 1h at RT. Cells were washed with PBS thrice after every incubation, imaged using Olympus FLUOVIEW FV1000, 63X water objective. Background was subtracted from every image. Foci were quantitated with protocol described by Duke University <https://microscopy.duke.edu/guides/count-nuclear-foci-ImageJ> using ImageJ.

Comet Assay:

Cells were treated with 10 mM Hydroxyurea for 0h and 6h followed by trypsinization. Comet Assay was performed as per user manual (Kit description is in supplementary methods). Slides were imaged using Keyence BZ-X800, 10X. Comets were analyzed using CaspLab - Comet Assay Software Project. Olive movement was used to quantify the tail lengths.

Statistical analysis of data derived from TMA:

T-tests, Chi-square tests, Fisher's Exact tests, or log-rank tests were used to compare patient and tumor variables between those with TONSL H-scores vs. those without. Details of the analysis are provided in supplementary methods.

Animal studies:

Indiana University Animal Care and Use Committee has approved all animal studies and all studies were conducted as per NIH guidelines. For tumor development studies, TONSL-immortalized KTB103 cells and/or transformed cells were injected into the mammary fat pad female NSG mice. All mice were implanted with estradiol pellets and tumor progression was assessed every week. For drug treatment studies, TMD-436 or TMD-231 cells were injected into the mammary fat pad of female nude mice. Treatment was initiated upon formation of palpable tumors. Animals were treated with 30 mg/kg of CBL0137, a previously reported dose (20), and the control group was treated with water via oral gavage for six weeks. To study TONSL dependency for *in vivo* growth of TONSL-amplified breast cancer cell lines - TMD-436 and TMCF-7 shControl and shTONSL cells were implanted

into the mammary fat pad of nude mice and tumor growth were measured for six weeks. Additional details are provided in supplementary methods.

Statistical analysis of *in vitro* and *in vivo* data:

In vitro and *in vivo* data were analyzed using GraphPad Prism. Data was analyzed using Mann-Whitney test and ANOVA. Details are provided in supplementary methods.

Data and Material availability:

All data needed to evaluate the conclusions in the paper are present in the paper and/or in the supplementary materials. Sequence data have been submitted to publicly available databases and accession numbers are indicated in the Materials and Methods section or supplementary methods. Requests for reagents including cell lines should be submitted to HN.

Results

TONSL-FACT complex components are upregulated in immortalized breast epithelial cells compared to primary cells

To study genomic changes in immortalized and transformed cells compared to isogenic primary cells, we developed a model system using breast core biopsies from seven healthy women and analyzed gene expression profile of primary breast epithelial cells and their human telomerase (hTERT) overexpressing counterparts (9). To identify functionally important genes transcriptionally deregulated during immortalization, we applied various filters depicted in Fig. 1A. In our previous study, we had demonstrated that primary and immortalized cells are composed of cells at distinct differentiation stages (9). For example, while primary cells contained variable levels of luminal progenitor (CD49f⁺/EpCAM⁺) and differentiated (CD49f⁻/EpCAM⁺) cells depending on the donor, immortalized cells consisted of only luminal progenitors (9). Previous studies have shown differential expression of ~2000 genes between luminal progenitor and differentiated cells (21). To exclude those genes whose expression is altered in immortalized cells compared to primary cells simply due to differences in differentiation status, we performed RNA-seq analyses of flow cytometrically sorted luminal progenitor and differentiated cells from breast tissues of genetic-ancestry mapped European-ancestry, African-ancestry, and Latina women (n=5 per group). Genes differentially expressed between these two populations (Table S2) were excluded from our analyses that compared primary versus immortalized cells. The remaining genes were then subjected to Ingenuity Pathway Analyses (IPA) to determine cancer progression relevance. The next filter was the recently developed CRISPR-Cas9 fitness and essentiality screens (22), and genes considered essential for cancer cell survival were further analyzed. Specific relevance of selected genes to breast cancer was determined using cBioportal and UALCAN databases (23,24). Gene listed in Table 1 were considered for further evaluation.

To confirm immortalization-associated changes, we quantitated the expression of several of these genes in isogenic primary, hTERT immortalized, and cells transformed with *H-Ras*^{G12V}+SV40-T/t antigens. We observed that TONSL, FACT components *SSRP1* and *SPT16*, and BARD1 were upregulated upon immortalization and further increased

upon transformation (Fig. 1B). Functionally, previous studies have shown TONSL forms a complex with FACT as well as BARD1 (12,25). By contrast, putative tumor suppressors *NDN* and *SEPP1* were downregulated in immortalized cells compared to primary cells (26,27). Since all three components of TONSL-FACT as well as its other interacting partner BARD1 were upregulated in immortalized cells compared to primary cells and TONSL is the only gene among four frequently amplified in cancer (see below), we focused on the role of TONSL in immortalization and transformation.

TONSL-amplified breast cancers are enriched in Integrative cluster 9

We first utilized various publicly available databases to test our hypothesis that genomic aberrations involving *TONSL* are implicated in breast tumorigenesis. Aberrant expression of TONSL in breast tumors was subsequently verified by genomics- and IHC-based analysis of breast tumor tissue microarray (TMA). Using UALCAN (24), we first confirmed that TONSL expression was elevated in breast cancers, irrespective of subtypes (Fig. S1A). *TONSL* is located in chr8q24.3, one of the amplified regions in breast cancer (28). *TONSL* was amplified in ~15–40% of all breast cancers and ~40% of patients display gain in TONSL expression (Fig. S1B, C). Since *cMyc* is well studied oncogene in chromosome 8q amplified region (28), we examined whether *TONSL* and *cMyc* amplification are mutually exclusive or co-occurrence. Interestingly, ~50% of breast cancers with *cMyc* amplifications also harbored *TONSL* amplification (Fig. S1D).

Since *TONSL* is amplified in breast cancer and CNVs primarily drive breast cancers (5), to further delineate the relationship between *TONSL* amplification and integrative clusters, we used METABRIC dataset, which classified breast cancers into 10 integrative clusters based on CNVs (4). Integrative cluster 9 contained the highest level of *TONSL* amplification followed by clusters 10, 1 and 5 (Fig. 1C). Integrative clusters 9 and 5 contain both estrogen receptor positive (ER⁺) and ER⁻ tumors, whereas clusters 1 and 10 comprise of ER⁺ and ER⁻ tumors, respectively (4). In clusters 1, 5, 9 and 10, dominant PAM50 subtypes were luminal B, luminal B and HER2, luminal B (mixed) and basal like, respectively (4). Thus, *TONSL* amplification is not unique to specific intrinsic subtype of breast cancer but shows some degree of correlation with CNV-driven integrative cluster classification. *TONSL* amplification/duplication was observed in breast tumors of stage I to stage IV (Fig. S1E) (29), further suggesting that genomic aberration involving *TONSL* is an early event in at least a subgroup of breast cancers. We also observed race specific differences of *TONSL* expression in Caucasian, African American and Asian breast cancer patients in TCGA data (Fig. S1F). Tumors in African American patients had significantly higher *TONSL* expression than Caucasian ($p=3.21E-14$) and Asian patients ($p=4.89E-04$). *TONSL* amplification was associated with overexpression of its mRNA (Fig. 1D). *TONSL*-amplification was associated with shorter recurrence-free and overall survival (Fig. 1E, 1F). Median recurrence free survival was ~160 months in *TONSL*-amplified group compared to ~260 months in the non-amplified group. *TONSL*-amplified tumors overrepresented gene sets corresponding to E2F targets, G2/M checkpoint, mitotic spindle, and mTORC1 pathways (Table S4).

To independently confirm genomic aberrations involving *TONSL*, we designed multiplex custom CodeSet and examined CNV using a NanoString technologies nCounter platform. We targeted chromosome 8q24.3 region and designed probes for *TONSL* along with the neighboring genes such as *SLC93A4*, *VPS28*, *CYHR1*, *KIFC2*, and *FOXH1* as depicted in (Fig. S2). DNA from immortalized breast epithelial cell lines was used as a negative control and DNA from MDA-MB-436 cell line with known *TONSL* amplification (see below) as a positive control. CNV scores are assigned as follows: deletions (0 to 0.4), normal copy number (0.4 to 1.4), duplication (1.4 to 2.4) and amplification (above 2.4). Of the 33 patient samples analyzed, amplification was seen in two samples and gene duplication in 12 samples, and *TONSL* deletion was not observed in any of these samples (Fig. 1G). In both samples with *TONSL* amplification, genes neighboring *TONSL* were also amplified (Fig. 1G; except *SLC93A4* in patient #2). Tumors with *TONSL* amplification were mostly TNBCs.

Breast tumor tissue microarray (TMA) analysis reveals prognostic significance of *TONSL* in breast cancer

To investigate the prognostic utility of *TONSL* in breast cancer, we evaluated *TONSL* expression in a TMA with breast tumors from 597 patients. *TONSL* expression was measurable in 472 tumors (79%). Table S5 describes the demographics and Fig. 2A shows staining patterns of *TONSL* in primary tumors. We compared *TONSL* H-score expression with ER, Progesterone Receptor (PR), human epidermal growth factor receptor 2 (HER-2/neu), Nodal stage, Tumor Stage and Grade. *TONSL* levels were correlated with ER (higher values within ER-) and Tumor Grade (higher values with higher grade) (Table S6). In univariable analyses, variables significantly related to disease free survival in the Cox proportional hazards regression models were HER2 status, tumor grade, tumor stage, and nodal stage (Table S7). HER2/neu+, higher tumor grade, higher tumor Stage, and nodal stage-positive tumors were correlated with lower disease-free survival. *TONSL* H-score was related to disease free survival with higher scores correlated to lower disease-free survival (log rank test p-value 0.0033).

In the multivariable analysis, tumor grade, and nodal stage were found to be significant. In the model without HER2 status, *TONSL* score was also significant. Higher tumor grade, nodal stage-positive and higher *TONSL* H-score were correlated with lower disease-free survival (Table S8). In multivariable models treating the H-score as dichotomous, H-score category was significant for ER+, patients on endocrine therapy, and patients ER+ and on endocrine therapy (Table S8). Kaplan-Meier plots derived from the univariable analyses using the categorical *TONSL* H-score for overall and the ER subgroup analyses showed specific impact of *TONSL* expression on outcome. There was a significant difference for the categorical *TONSL* H-score (p=0.0022) with the higher H-scores having worse progression free survival (all cases). There was also a significant difference for the categorical *TONSL* H-score within the ER-positive subgroup (p=0.0026) with the higher H-scores having worse progression free survival (ER+ patients) (Fig. 2B–E). In particular, in ER+ group treated with endocrine therapy, *TONSL* overexpression was associated with worst outcome (Fig. 2D, E).

TONSL is an immortalizing oncogene.—TONSL expression was elevated in immortalized cells compared to primary cells (Fig. 1B) and amplification/duplication of *TONSL* was observed in Stage I breast cancers (Fig. S1C, E). These results raised the possibility that TONSL itself possesses an immortalizing function. To test this hypothesis, we overexpressed TONSL in primary breast epithelial cells and transferred cells to regular tissue culture dishes instead of plates pre-coated with conditioned media from 804G cells, which is required for the growth of primary cells (8). After one month in culture, immortalized clones appeared, and these cells expressed higher levels of TONSL compared to parental cells (Fig. 3A). Phase contrast images of primary and TONSL-immortalized cells are shown in Fig. S3A. We also observed elevated telomerase activity in TONSL-overexpressing cells compared to primary cells (Fig. 3B), although TONSL increased TERT mRNA levels only modestly (Fig. 3C). shRNA against TONSL reduced levels of TERT in TONSL amplified cell line TMD-436 (Fig. 3C, see below for further details of these cells). KTB103 TONSL-immortalized cells were enriched for luminal progenitor properties (CD49f+/EpCAM+) compared to primary cells (Fig. 3D), a property of TONSL immortalized cells similar to hTERT immortalized cells (9). Composition of primary cells varied between samples, consistent with our previous report (30); KTB103 and KTB109181 cells are from African ancestry donors and KTB103 contained both luminal progenitor and basal cells.

Similar to previous observations by others (31), we recently reported that the combination of *H-Ras*^{G12V} with SV40-T/t antigens reproducibly transform hTERT-immortalized breast luminal epithelial cells of healthy donors of different genetic ancestry and was the most effective oncogene combination for transformation (9). Furthermore, depending on the donor cell type, the resulting tumors are adenocarcinomas or squamous carcinomas. In our transformation model system, as mutant p53 (TP53^{R273C}) was less efficient than SV40-T/t antigens as a cooperating oncogene, we examined whether overexpression of *H-Ras*^{G12V} + SV40-T/t antigens, *H-Ras*^{G12V} + TP53^{R273C}, and *cMyc* in TONSL immortalized cells could lead to transformation. Overexpression of oncogenes in immortalized cells was confirmed by western blotting or immunofluorescence (in case of TP53^{R273C}) (Fig. S3B–D). Five million cells were injected into the mammary fat pad of NSG mice and tumor development was examined for ~10 weeks. No tumor developed from cells that overexpressed TONSL, TONSL + H-Ras^{G12V} + TP53^{R273C}, TONSL + TP53^{R273C} or TONSL + cMyc (Fig. 3E). However, three out of five mice harboring TONSL-overexpressing cells with H-Ras^{G12V} + SV40-T/t antigens developed invasive ductal carcinoma (Fig. 3E). Interestingly, H-Ras^{G12V} + SV40-T/t antigens derived tumors displayed expression of ER α , PR, and GATA3 (Fig. 3F). Tumor cell lines generated from resulting tumors expressed ER α (Fig. 3G). To our knowledge, this is the first model system where ER α + adenocarcinoma can be generated from primary breast epithelial cells using H-Ras oncogene without the need for unique propagation methods or 3D cultures to enrich for ER α + cells (32).

TONSL overexpression in primary breast epithelial cells leads to chromatin reorganization

Since TONSL-FACT or TONSL-MMS22L complexes bind chromatin during DNA replication, repair and/or transcription (12,14), we next examined whether TONSL overexpression in primary cells leads to chromatin reorganization and gene expression

changes, which consequently leads to immortalization. For this purpose, we performed ATAC-seq and RNA-seq of primary and TONSL-immortalized cells. TONSL caused significant changes in chromatin accessibility including selective opening and closing of chromatin (Fig. 4A, B). An example of TONSL-induced closing of chromatin near the transcription start site of *SMARCA2* gene is shown in Fig. 4C. Similar to our previous study in MCF-7 cells with and without estradiol treatment (18), TONSL-induced chromatin accessibility changes correlated with both increased and reduced expression of downstream genes (Fig. 4D). Chromatin accessibility changes and TONSL-mediated gene expression changes are listed in Tables S9 and S10. To ensure that few of the gene expression changes in TONSL-immortalized cells compared to primary cells was not due to differences in differentiation status, we compared the gene expression changes noted in Table S10 with that of gene expression differences between luminal mature and luminal progenitor cells listed in Table S2. Only 10% of genes showed an overlap, suggesting that the majority of gene expression changes in TONSL overexpressing cells compared to primary cells was due to TONSL overexpression.

IPA of genes differentially upregulated due to TONSL overexpression were related to cell survival and proliferation such as cell cycle control of chromosomal replication ($p=2.05E-08$), mismatch DNA repair ($p=3.6E-06$), and kinetochore metaphase signaling pathway ($p=1.4E-11$). TONSL overexpression led to inhibition of pathways related to cell cycle control including G2/M DNA damage checkpoint pathway ($p=3.7E-04$) and role of CHK proteins in cell cycle checkpoint control ($p=1.54E-07$) (Fig. 4E). Importantly, TONSL overexpression resulted in significant changes in expression of genes associated with specific DNA repair hubs, particularly HR pathway and base excision repair hubs (BER) (Fig. 4F), and elevated expression of TONSL binding partner *MMS22L* with accompanying changes in chromatin accessibility of *MMS22L* gene (Table S9 and Table S10) (11). TONSL overexpression was sufficient to increase the expression of its known other interactors *SSRP1*, *SPT16* (components of FACT) and *BARD1* (25) (Fig. S4). Thus, TONSL may control the expression of its binding partners *SSRP1*, *SPT16*, and *BARD1* in immortalized compared to primary cells (Table 1).

DNA repair hub impairment is observed exceedingly early in breast tumorigenesis (33). Because impaired non-homologous end joining (NHEJ) and nucleotide excision repair (NER) are associated with genomic instability and increased mutation frequencies (33), TONSL overexpression due to chr8q24.3 amplification is likely a trigger for gain of both replicative capacity and genomic instability in tumor initiating cells through an imbalance in DNA repair hubs. To further determine the role of TONSL in genomic instability, we examined TONSL-immortalized cells for the expression of 70 genes associated with chromosomal instability (34). Sixty-seven of these genes were overexpressed in TONSL-immortalized cells compared to parental cells (Fig. S4C). We also examined the expression levels of 11-gene breast cancer proliferation signature (35) and found TONSL upregulating the expression of all of these genes (Table S11), further solidifying the role of TONSL in overcoming replication block and enhancing survival.

Several interactors and antagonists of BRCA1-BRAD1 complex have been described and these interactors/antagonists modulate antitumorigenic roles of BRCA1-BARD1 (36)

and are involved in DNA repair machinery, DNA damage signaling, transcription/R-loop metabolism, cell growth, centromere regulation, chromosome segregation, chromatin modelling, and E3 ligase substrates. Interestingly, TONSL overexpression altered the expression levels of 93/133 interactors/antagonists significantly (p value < 0.01 ; Table S12, genes with two-fold change are described in Fig. 4G). Collectively, these results indicate the profound impact of TONSL overexpression on genome integrity and BRCA1-BARD1 mediated tumor suppressor pathways.

Transcription factor binding site enrichment analysis of genomic regions that became inaccessible upon TONSL overexpression revealed enrichment for binding sites for transcription factors such as AP1 family, Bach2, p53, and p63 (Fig. 4H). Accessible regions upon TONSL overexpression were enriched for binding sites for ATF3, NF-E2, NF- κ B, and BATF (Fig. 4I). Genes with enrichment of NF- κ B binding sites included transcription regulators SMAD3, KDM2A, TWIST1, IL-1 β , CDC7, and TGM2 (Table S13). Genes enriched for p53 binding sites included DLK1, JAK2, CDH2, TRAF6, TET2, CDK6, and GLI3 (Table S13).

TONSL is required for growth of TONSL-amplified cell lines *in vivo*.—To independently identify TONSL-regulated genes and their requirement for growth *in vivo*, we utilized breast cancer cell line models. Based on [Depmap.org](https://depmap.org) database, MDA-MB-436, HCC1937, BT483, and MCF7 have chromosome 8q24.3 amplification and HCC1419 has 8q24.3 amplification but the TONSL gene is disrupted by translocation. TONSL protein is expressed at a higher level in cell lines with chromosome 8q24.3 amplification (Fig. 5A). We knocked down TONSL in TMD-436 cell line, a cell line generated from xenografts derived from parental MDA-MB-436 cells (37). Three independent clones were generated, and each carried different shRNA targeting different regions of TONSL (Fig. 5B). RNA-seq analysis of control (TMD-436pLKO) and three shRNA clones followed by IPA revealed a role for TONSL in growth, proliferation, and metastasis such as tumor microenvironment pathway ($p=2.36E-09$), notch signaling ($p=9.01E-03$), and cancer metastasis signaling ($1.5E-04$) (Fig. 5C and Table S14 for RNA-seq data). We next compared genes differentially expressed upon TONSL overexpression in primary cells with genes differentially expressed in TMD-436 upon TONSL knockdown. This analysis identified 280 genes whose expression was elevated upon TONSL overexpression in primary cells but reduced upon TONSL knockdown in TONSL-amplified cancer cells. By contrast, expression of 283 genes were repressed upon TONSL overexpression in primary cells but elevated upon knockdown of TONSL in TMD-436 cells (Fig. 5D and Table S15). We confirmed TONSL-dependent changes in expression of select genes by qRT-PCR in primary cells as well as in TMD-436 cells (Fig. 5E). Genes downregulated by TONSL included those associated with luminal cell identity (FOXA1, GATA3 (38)) and those elevated included genes such as TWIST and ZEB1, which are associated with mesenchymal phenotype (38).

Since cell growth pathway activation by TONSL was apparent in the above analysis, we studied the impact of TONSL manipulation on tumor growth *in vivo*. For this purpose, we implanted parental control TMD-436pLKO cells and three TONSL shRNA clones into the mammary fat pad of nude mice and monitored tumor growth (7–9 animals per group). Tumors generated from TMD-436pLKO cells grew at a significantly faster rate

than tumors generated by TONSL shRNA expressing clones, (two of three clones, Fig. 5F), confirming growth promoting properties of TONSL. We also generated MCF-7 cells, which have TONSL amplification, with TONSL knockdown (Fig. 5G) and injected in 8–9 animals per group. MCF-7 cells with TONSL shRNA were less efficient in generating tumors than control MCF-7pLKO cells (Fig. 5H). Collectively, these data strongly support a role for TONSL in promoting initiation and progression of breast cancer.

TONSL overexpressing primary cells exhibit upregulated DNA repair via HR: TONSL

along with MMS22L repairs DNA upon replication fork collapse and regulates the replication process (12) (15), whereas loss of TONSL and/or mutated TONSL gene leads to increased replication stress and spontaneous DNA double strand breaks (12,39). Based on these observations, we sought to study the effect of replication stress in TONSL overexpressing cells. We treated primary and TONSL overexpressing primary cells with 10mM Hydroxyurea (HU) for 0 and 6 hours. HU causes replication stress by prolonging the replication initiation and elongation by inhibiting the nucleotide synthesis process causing DNA double strand breaks (DDSB) and cell cycle arrest at S-phase (40). Untreated and treated cells were analyzed for DNA damage with comet assay. Damaged DNA migrates faster upon electrophoresis, and the length of the tail calculated as an olive movement (with CaspLab software) is directly proportional to the level of DNA damage within cells (41). As expected, HU treatment led to significant damage within primary cells detected as a longer tail of damaged DNA, while TONSL overexpressing cells were resistant to treatment and displayed shorter comet tails (Fig. 6A, B). With 6h of HU treatment, TONSL overexpressing cells had reduced DNA damage compared to 0h control (Fig. 6B). These results suggest that TONSL overexpression hyperactivates the DNA repair pathways upon HU treatment.

To repair DDSB, cells either activate HR or NHEJ depending on the phase of cell cycle(33). To study the TONSL mediated DNA repair pathway, the same set of cells was treated with 10mM HU for 0 and 6 hours followed by detection of RAD51 and 53BP1 foci formation by immunofluorescence to quantitate HR and NHEJ (33), respectively. Damaged DNA was assessed with p- γ H2AX staining. Changes in p- γ H2AX along with RAD51 in primary and TONSL+ cells were detected as a nuclear-foci. With DNA damage, we observed accumulation of RAD51 at the damaged sites (Fig. 6C–E), demonstrating active repair via HR within primary and TONSL overexpressing cells. The TONSL overexpressing cells showed notably active HR pathway compared to primary cells with and without DNA damage. By contrast, no dynamic changes in the 53BP1 foci formation were observed without treatment in either cell type (Fig. 6F). With HU treatment, significant increase in p- γ H2AX and 53BP1 foci was observed in primary as well as TONSL overexpressing cells (Fig. 6F–H). Hence, no effective change in NHEJ activity upon TONSL overexpression. These results indicate that TONSL specifically activates HR pathway with no direct impact on NHEJ pathway.

TONSL overexpressing breast cancer cell lines are sensitive to CBL0137 *in vivo* and *in vitro*: TONSL is known to interact with FACT complex to modulate DNA repair and overcome replication stress (42). In addition, we observed that overexpression of TONSL in primary cells caused upregulation of FACT components SSRP1 and SPT16 (Fig. S4A).

Furthermore, TONSL knockdown in TMD-436 reduced the levels of both SSRP1 and SPT16 (Fig. S4B). These results suggested that the expression and activity of TONSL and FACT are interconnected and cells that overexpress TONSL would be sensitive to FACT targeting drugs such as curaxins. Curaxins induce FACT trapping on chromatin by redistributing it from actively transcribed regions to other genomic regions (17). The selective toxicity of curaxins on cancer cells compared to normal cells may lie with FACT's role as histone chaperone in three key processes - transcription, replication, and DNA repair. Among various curaxins, CBL0137 has demonstrated clinical activity and good safety profile in a phase I clinical trial (43). Since TONSL expression is elevated in immortalized and transformed cells compared to primary cells in our isogenic system, we first evaluated the effect of CBL0137 on cell proliferation in the isogenic model. IC50 for primary cells was 941 nM, whereas IC50 of immortalized cells was 336 nM and transformed cells was 375 nM (Fig. 7A). Immortalized breast epithelial cells with BRCA1 mutation carrier were also sensitive with IC50 of 320 nM (Fig. 7A). Thus, immortalization increases dependency on TONSL-FACT activity.

Next, breast cancer cell lines with chr8q24.3 amplification (MDA-MB-436, HCC1937, BT483, MCF7), and cell lines without chr8q24.3 amplification [(HCC1419* has 8q24.3 amplification but TONSL gene is disrupted by translocation within the gene DepMap data), MDA-MB-231, tumor derived MDA-MB-231 and MDA-MB-468] were treated with increasing concentrations of CBL0137. Chr8q24.3 amplified cell lines were sensitive to CBL0137 and showed systemic decline in cell proliferation with increasing drug concentration (Fig. 7B). However, Chr8q24.3 non-amplified cells demonstrated no difference in proliferation (Fig. 7C). Mechanistically, CBL0137 treated cells showed disrupted cell cycle as treated cells showed lower number of cells at G1 but elevated number of cells at S or G2/M phase depending on concentration of drug (Fig. 7D).

CBL0137 through FACT trapping can indirectly cause DNA damage (17). If the sensitivity of TONSL-amplified cells to CBL0137 is simply due to DNA damage induced by the drug, TONSL-amplified cell lines should show higher sensitivity to other DNA damaging agents. To test this possibility, we analyzed the Genomics of Drug Sensitivity in Cancer (GDSC) (44) dataset for sensitivity of various breast cancer cell lines to cisplatin, doxorubicin and bleomycin and correlated sensitivity to these DNA damaging agents with TONSL amplification. We did not observe any correlation between sensitivity to these drugs and TONSL amplification, suggesting the CBL0137 sensitivity of the TONSL amplified tumor is not due to DNA damaging effect but due specific targeting of TONSL-FACT complex.

The effect of CBL0137 treatment on cell proliferation was validated *in vivo*. Tumor derived MDA-MB-436 (TMD-436) and MDA-MB-231 (TMD-231) cells (500,000 cells) were injected into the mammary fat pad of nude mice (TMD-436: N=11 per group, TMD-231: N=10 per group) and treatment was initiated upon formation of palpable tumors. Animals were treated with drug (test group, 30 mg/kg, five days a week by oral gavage) and water (control group) for six weeks. CBL0137 inhibited the growth of TMD-436 cells-derived tumors (Fig. 7E) but TMD-231 cell-derived tumors were resistant to the treatment. (Fig. 7F). Both cell lines correspond to mesenchymal stem-like subtype of TNBC and, thus, difference in sensitivity is less likely due to differences in TNBC

subtypes (29,45). CBL0137 treated TMD-436 cells-derived tumors contained lower levels of Ki67+ cells compared to vehicle treated tumors suggesting the effects of CBL0137 on cell proliferation (Fig. 7G). Moreover, metastatic MDA-MB-436 cells remain dormant till the microenvironment promotes angiogenic switch (46) and CBL0137 likely blocks this switch as lungs of untreated but not treated mice showed inflammatory changes required for metastasis dormancy. Thus, chr8q24.3 amplified tumors are sensitive to CBL0137 *in vivo*.

Discussion

Cancer progression including mechanisms associated with uncontrolled cell proliferation, loss of contact inhibition, inhibition of apoptosis/senescence pathway, gain of stem cell properties, mutations leading to activation of oncogenes and inactivation of tumor suppressor genes and metastasis cascades have been studied extensively, culminating in the development of hallmarks of cancer (2). However, the ability to detect the earliest changes during cancer initiation has been limited due to deficiencies in isogenic model system where every step of cancer progression can be evaluated. To address this key limitation, we created a system by first developing a resource for primary cells from healthy donors and then establishing a culturing method that allows propagation of primary epithelial cells with luminal epithelial characteristics sufficient for immortalization and transformation using cancer-relevant oncogenes such as RAS and SV40 T/t antigen. Though RAS mutations are rare in breast cancers, the majority of established breast cancer cell lines have mutations in RAS effector pathways e.g., MDA-MB-231 and/or patients with certain RAS abnormalities are at higher risk of developing breast cancer highlighting the significance of model used (47,48). Ras pathway activation is linked to endocrine resistance in breast cancer, further supporting the relevance of Ras pathway in breast cancer (49). SV40 T/t antigens inhibit two of the breast cancer relevant pRB and p53 tumor suppressor pathways, mimicking the disease (50). By using this model system, we identified TONSL as an immortalizing oncogene. Previous studies have demonstrated that cancer initiates due a single catastrophic genomic event (51) and a genomic event that leads to chromosome 8q24.3 amplification could cause tumor initiation through TONSL.

Most of the *in vitro* model systems to achieve immortalization utilize hTERT and our studies provide an alternative method to achieve immortalization. It is important to note that *TERT* gene is rarely activated through mutations or amplified in cancer. Cancer-specific upregulation in few cases is due to mutations in the promoter regions and promoter duplications (52). Amplification of TONSL could be an alternative mechanism to achieve immortalization. It is interesting that in pan-cancer studies available through cBioportal (23), 13% of all cancers show *TONSL* amplification. *TONSL* amplification is observed in 35% of ovarian cancers, 35% of pancreatic cancers, 15% of esophageal/gastric cancers, 9% of urothelial cancers, 11% of head and neck cancers, and 16% of hepatocellular carcinomas. In addition, 60% of small cell lung cancers have *TONSL* mutations. Therefore, results obtained in our breast cancer models could be relevant to multiple cancer types.

TONSL is a multifunctional protein and has been studied in the context of its association with MMS22L and FACT (11,12,14). TONSL has been co-implicated with FACT and BARD1/BRCA1 in resolving UV-induced DNA damage. As a component of the MMS22L-

TONSL complex, it interacts with MCM, FACT and RPA, specifically identifies and binds to H4K20me0, and controls HR during replication-associated DNA damage. TONSL is part of the cell cycle-dependent HR and maintains genomic stability during S phase (15). However, the specific functions of TONSL that contribute to tumor initiation have not been identified. We observed that TONSL overexpression results in increase telomerase activity. However, whether TONSL directly increases telomerase activity to cause immortalization or other functions of TONSL indirectly leads to increase in telomerase activity remains to be fully understood. In this regard, we did not observe specific effects of TONSL overexpression on chromatin accessibility around the *Telomerase* gene. Because telomerase activity is dependent on multiple proteins (53), it is possible that other components required for telomerase activity could be the targets of TONSL. It is also possible that TONSL causes immortalization through recombination-based mechanisms such as ALT (Alternative Lengthening of Telomeres) (54) although this is less likely as TONSL overexpression increased telomerase activity. ALT is dependent on DNA repair and HR-associated proteins such as BRCA1, BLM, PALB2 (55). TONSL influences the expression of many HR-associated genes as well as interacts with several of them (Fig. 4F). It is possible that HR-associated function of TONSL is responsible for immortalization as this function may allow cells to overcome replication checkpoint. Although HR was thought to be error-free DNA repair pathway for a long time, recent studies suggest that HR is a error-prone repair process in the context of large amounts of DNA synthesis and contributes to translocations and complex chromosome rearrangements (56). Thus, TONSL's ability to overcome replication checkpoint combined with error-prone HR while overcoming replication checkpoint could lead to cancer initiation. Consistent with this possibility, TONSL increased the expression of genes associated with chromosome instability (Fig. S4C). Further dissection of HR-associated functions of TONSL may provide mechanistic insights into TONSL-mediated immortalization and cancer initiation.

We show that *TONSL* amplification and/or overexpression sensitize cancer cells to CBL0137, but additional mechanistic studies are needed to decipher how *TONSL*-amplified cancer cells are susceptible to CBL0137. Based on gene expression analysis, it is apparent that TONSL causes dramatic imbalance in various components of DNA repair pathways including components of BRCA1-BARD1 tumor suppressor network, which could alter the cell cycle checkpoints and cell senescence pathways while promoting chromosomal replication, chromosomal instability, and cell cycle progression despite chromosomal abnormalities. These functions of TONSL could still be dependent on FACT complex and trapping of FACT by CBL0137 on irrelevant regions of the chromatin may render TONSL ineffective in performing these functions.

Can additional drugs targeting TONSL or Chr8q24.3 amplicon be developed? Drugs that disrupt TONSL-MMS22L1 complex are potential therapeutic agents for TONSL amplified cancers as TONSL overexpression increased MMS22L1 level suggesting elevated activity of this complex in TONSL amplified cancers. Additionally, recognition and binding of TONSL-MMS22L1 to H4K20me0 at DNA lesions is an essential step to resolve stalled replication forks in rapidly growing cells (13). Although lacking enzymatic activity on its own, disruptors of interactions between TONSL and its several interacting partners can be developed as therapies for cancers with *TONSL* amplification (12). Other genes

within amplified Chr.8q24.3 locus remain uncharacterized for oncogenic functions. With our studies demonstrating immortalizing function of TONSL, the oncogenic role of other genes of the locus can be investigated and if proven necessary for tumorigenesis, those genes become the targets for drug discovery. Chr.8q24.3 amplification is an established marker of early relapse and drug resistance in breast cancer patients (57). This knowledge can be extended to identify new agents that perturb the Chr.8q24.3 dependence of cancer cells and perhaps improve response to chemotherapy. Collectively, our studies reveal a new targetable molecule/pathway potentially involved in initiation of 13% of cancers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We thank Connie J. Eaves (University of British Columbia, Vancouver, Canada) and Dr. Michael D. Cole (Princeton University) for mutant TP53 and cMyc expression constructs, respectively. We also thank IUSCCC tissue procurement core, Ms. Casey Bales for breast tumor TMA, Ms. Anjanappa for preparing luminal mature and luminal progenitor cells for RNA-seq, Komen Tissue Bank for normal breast tissue, IUSCCC flow cytometry core, Genomics and Bioinformatics and C3B cores at IUSM, Stark Neuroscience Research Institute Biomarker Core and Genomics core at IU-Bloomington for their services. BioRender software was used to generate graphical abstract and Figure 3E.

Funding support:

Susan G Komen for the Cure (SAC110025), Becky Sobel Breast Cancer Research Fund, Precision Health Initiative of Indiana University, Breast Cancer Research Foundation (all to H.N.); Breast Cancer Research Foundation (KDM); Marvella Bayh Memorial Scholarship (to S.S.) and Shirley Dawn Wyss Breast Cancer Fellowship at the Indiana University Foundation (to R.D.). Susan G. Komen for the Cure, Breast Cancer research Foundation, Vera Bradley Foundation for Breast Cancer Research and the Catherine Peachy Foundation provide funding support for the Komen Normal Tissue Bank

References

1. Surveillance Research program, National Cancer Institute SEER*Stat software.
2. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*;144(5):646–74. [PubMed: 21376230]
3. Jones MJ, Jallepalli PV. Chromothripsis: chromosomes in crisis. *Dev Cell* 2012;23(5):908–17. [PubMed: 23153487]
4. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 2012;486(7403):346–52. [PubMed: 22522925]
5. Cancer Genome Atlas N Comprehensive molecular portraits of human breast tumours. *Nature* 2012;490(7418):61–70. [PubMed: 23000897]
6. Luond F, Tiede S, Christofori G. Breast cancer as an example of tumour heterogeneity and tumour cell plasticity during malignant progression. *Br J Cancer* 2021;125(2):164–75. [PubMed: 33824479]
7. Marcotte R, Sayad A, Brown KR, Sanchez-Garcia F, Reimand J, Haider M, et al. Functional Genomic Landscape of Human Breast Cancer Drivers, Vulnerabilities, and Resistance. *Cell* 2016;164(1–2):293–309. [PubMed: 26771497]
8. Prasad M, Kumar B, Bhat-Nakshatri P, Anjanappa M, Sandusky G, Miller KD, et al. Dual TGFbeta/BMP Pathway Inhibition Enables Expansion and Characterization of Multiple Epithelial Cell Types of the Normal and Cancerous Breast. *Mol Cancer Res* 2019;17(7):1556–70. [PubMed: 30992305]

9. Kumar B, Prasad MS, Bhat-Nakshatri P, Anjanappa M, Kalra M, Marino N, et al. Normal breast-derived epithelial cells with luminal and intrinsic subtype-enriched gene expression document inter-individual differences in their differentiation cascade. *Cancer Res* 2018;78(17):5107–23. [PubMed: 29997232]
10. Carrot-Zhang J, Chambwe N, Damrauer JS, Knijnenburg TA, Robertson AG, Yau C, et al. Comprehensive Analysis of Genetic Ancestry and Its Molecular Correlates in Cancer. *Cancer Cell* 2020;37(5):639–54 e6. [PubMed: 32396860]
11. O’Connell BC, Adamson B, Lydeard JR, Sowa ME, Ciccio A, Bredemeyer AL, et al. A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability. *Mol Cell* 2010;40(4):645–57. [PubMed: 21055985]
12. O’Donnell L, Panier S, Wildenhain J, Tkach JM, Al-Hakim A, Landry MC, et al. The MMS22L-TONSL complex mediates recovery from replication stress and homologous recombination. *Mol Cell* 2010;40(4):619–31. [PubMed: 21055983]
13. Saredi G, Huang H, Hammond CM, Alabert C, Bekker-Jensen S, Forne I, et al. H4K20me0 marks post-replicative chromatin and recruits the TONSL-MMS22L DNA repair complex. *Nature* 2016;534(7609):714–18. [PubMed: 27338793]
14. Hustedt N, Durocher D. The control of DNA repair by the cell cycle. *Nat Cell Biol* 2016;19(1):1–9. [PubMed: 28008184]
15. Piwko W, Olma MH, Held M, Bianco JN, Pedrioli PG, Hofmann K, et al. RNAi-based screening identifies the Mms22L-Nfkbil2 complex as a novel regulator of DNA replication in human cells. *EMBO J* 2010;29(24):4210–22. [PubMed: 21113133]
16. Yaswen P, MacKenzie KL, Keith WN, Hentosh P, Rodier F, Zhu J, et al. Therapeutic targeting of replicative immortality. *Semin Cancer Biol* 2015;35 Suppl:S104–S28. [PubMed: 25869441]
17. Chang HW, Valieva ME, Safina A, Chereji RV, Wang J, Kulaeva OI, et al. Mechanism of FACT removal from transcribed genes by anticancer drugs curaxins. *Sci Adv* 2018;4(11):eaav2131. [PubMed: 30417101]
18. Chen DJ, Parker TM, Bhat-Nakshatri P, Chu XN, Liu YL, Wang Y, et al. Nonlinear relationship between chromatin accessibility and estradiol-regulated gene expression. *Oncogene* 2021;40(7):1332–46. [PubMed: 33420376]
19. Bhat-Nakshatri P, Gao H, Sheng L, McGuire PC, Xuei X, Wan J, et al. A single-cell atlas of the healthy breast tissues reveals clinically relevant clusters of breast epithelial cells. *Cell Rep Med* 2021;2(3):100219. [PubMed: 33763657]
20. Gasparian AV, Burkhart CA, Purmal AA, Brodsky L, Pal M, Saranadasa M, et al. Curaxins: anticancer compounds that simultaneously suppress NF-kappaB and activate p53 by targeting FACT. *Science translational medicine* 2011;3(95):95ra74.
21. Raouf A, Zhao Y, To K, Stingl J, Delaney A, Barbara M, et al. Transcriptome analysis of the normal human mammary cell commitment and differentiation process. *Cell Stem Cell* 2008;3(1):109–18. [PubMed: 18593563]
22. Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet* 2017;49(12):1779–84. [PubMed: 29083409]
23. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer discovery* 2012;2(5):401–4. [PubMed: 22588877]
24. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* 2017;19(8):649–58. [PubMed: 28732212]
25. Hill SJ, Rolland T, Adelmant G, Xia X, Owen MS, Dricot A, et al. Systematic screening reveals a role for BRCA1 in the response to transcription-associated DNA damage. *Genes Dev* 2014;28(17):1957–75. [PubMed: 25184681]
26. Barrett CW, Reddy VK, Short SP, Motley AK, Lintel MK, Bradley AM, et al. Selenoprotein P influences colitis-induced tumorigenesis by mediating stemness and oxidative damage. *J Clin Invest* 2015;125(7):2646–60. [PubMed: 26053663]

27. Yang H, Das P, Yu Y, Mao W, Wang Y, Baggerly K, et al. NDN is an imprinted tumor suppressor gene that is downregulated in ovarian cancers through genetic and epigenetic mechanisms. *Oncotarget* 2016;7(3):3018–32. [PubMed: 26689988]
28. Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* 2013;45(10):1113–20. [PubMed: 24071849]
29. Tym JE, Mitsopoulos C, Coker EA, Razaz P, Schierz AC, Antolin AA, et al. canSAR: an updated cancer research and drug discovery knowledgebase. *Nucleic Acids Res* 2016;44(D1):D938–43. [PubMed: 26673713]
30. Nakshatri H, Anjanappa M, Bhat-Nakshatri P. Ethnicity-Dependent and -Independent Heterogeneity in Healthy Normal Breast Hierarchy Impacts Tumor Characterization. *Scientific reports* 2015;5:13526. [PubMed: 26311223]
31. Hahn WC, Dessain SK, Brooks MW, King JE, Elenbaas B, Sabatini DM, et al. Enumeration of the simian virus 40 early region elements necessary for human cell transformation. *Mol Cell Biol* 2002;22(7):2111–23. [PubMed: 11884599]
32. Fridriksdottir AJ, Kim J, Villadsen R, Klitgaard MC, Hopkinson BM, Petersen OW, et al. Propagation of oestrogen receptor-positive and oestrogen-responsive normal human breast cells in culture. *Nature communications* 2015;6:8786.
33. Huang R, Zhou PK. DNA damage repair: historical perspectives, mechanistic pathways and clinical translation for targeted cancer therapy. *Signal Transduct Target Ther* 2021;6(1):254. [PubMed: 34238917]
34. Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat Genet* 2006;38(9):1043–8. [PubMed: 16921376]
35. Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. *Clin Cancer Res* 2010;16(21):5222–32. [PubMed: 20837693]
36. Tarsounas M, Sung P. The antitumorigenic roles of BRCA1-BARD1 in DNA repair and replication. *Nat Rev Mol Cell Biol* 2020;21(5):284–99. [PubMed: 32094664]
37. Sheridan C, Kishimoto H, Fuchs RK, Mehrotra S, Bhat-Nakshatri P, Turner CH, et al. CD44+/CD24– breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. *Breast Cancer Res* 2006;8(5):R59. [PubMed: 17062128]
38. Visvader JE, Stingl J. Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev* 2014;28(11):1143–58. [PubMed: 24888586]
39. Burrage LC, Reynolds JJ, Baratang NV, Phillips JB, Wegner J, McFarquhar A, et al. Bi-allelic Variants in TONSL Cause SPONASTRIME Dysplasia and a Spectrum of Skeletal Dysplasia Phenotypes. *Am J Hum Genet* 2019;104(3):422–38. [PubMed: 30773277]
40. Alvino GM, Collingwood D, Murphy JM, Delrow J, Brewer BJ, Raghuraman MK. Replication in hydroxyurea: it's a matter of time. *Mol Cell Biol* 2007;27(18):6396–406. [PubMed: 17636020]
41. Collins AR. The comet assay for DNA damage and repair: principles, applications, and limitations. *Mol Biotechnol* 2004;26(3):249–61. [PubMed: 15004294]
42. Prendergast L, Hong E, Safina A, Poe D, Gurova K. Histone chaperone FACT is essential to overcome replication stress in mammalian cells. *Oncogene* 2020;39(28):5124–37. [PubMed: 32533099]
43. Sarantopoulos J MD, Sharma N, Iyer RV, Ma WW, Ahluwalia MS, Johnson S, Purmal A, Shpigotskaya P, Hards A, Leonov A, Gurova K, Gudkov A, Zakurdaeva K, Miller LL, and Dowlati A. Results of completed phase I trial of CBL0137 administered intrveoulsy to patients with advanced solid tumors. *Journal of Clinical Oncology* 2020;38:3583.
44. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* 2013;41(Database issue):D955–61. [PubMed: 23180760]

45. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest* 2011;121(7):2750–67. [PubMed: 21633166]
46. Straume O, Shimamura T, Lampa MJ, Carretero J, Oyan AM, Jia D, et al. Suppression of heat shock protein 27 induces long-term dormancy in human breast cancer. *Proc Natl Acad Sci U S A* 2012;109(22):8699–704. [PubMed: 22589302]
47. Wagner KU. Know thy cells: commonly used triple-negative human breast cancer cell lines carry mutations in RAS and effectors. *Breast Cancer Res* 2022;24(1):44. [PubMed: 35761335]
48. Sharif S, Moran A, Huson SM, Iddenden R, Shenton A, Howard E, et al. Women with neurofibromatosis 1 are at a moderately increased risk of developing breast cancer and should be considered for early screening. *J Med Genet* 2007;44(8):481–4. [PubMed: 17369502]
49. Xia Y, He X, Renshaw L, Martinez-Perez C, Kay C, Gray M, et al. Integrated DNA and RNA Sequencing Reveals Drivers of Endocrine Resistance in Estrogen Receptor-Positive Breast Cancer. *Clin Cancer Res* 2022;28(16):3618–29. [PubMed: 35653148]
50. Rotondo JC, Mazzoni E, Bononi I, Tognon M, Martini F. Association Between Simian Virus 40 and Human Tumors. *Frontiers in oncology* 2019;9:670. [PubMed: 31403031]
51. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell*;144(1):27–40. [PubMed: 21215367]
52. Barger CJ, Suwala AK, Soczek KM, Wang AS, Kim MY, Hong C, et al. Conserved features of TERT promoter duplications reveal an activation mechanism that mimics hotspot mutations in cancer. *Nature communications* 2022;13(1):5430.
53. Jafri MA, Ansari SA, Alqahtani MH, Shay JW. Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. *Genome Med* 2016;8(1):69. [PubMed: 27323951]
54. Sobinoff AP, Pickett HA. Alternative Lengthening of Telomeres: DNA Repair Pathways Converge. *Trends Genet* 2017;33(12):921–32. [PubMed: 28969871]
55. Martinez AR, Kaul Z, Parvin JD, Groden J. Differential requirements for DNA repair proteins in immortalized cell lines using alternative lengthening of telomere mechanisms. *Genes, chromosomes & cancer* 2017;56(8):617–31. [PubMed: 28398700]
56. Rodgers K, McVey M. Error-Prone Repair of DNA Double-Strand Breaks. *J Cell Physiol* 2016;231(1):15–24. [PubMed: 26033759]
57. Bilal E, Vassallo K, Toppmeyer D, Barnard N, Rye IH, Almendro V, et al. Amplified loci on chromosomes 8 and 17 predict early relapse in ER-positive breast cancers. *PloS one* 2012;7(6):e38575. [PubMed: 22719901]

Statement of Significance

The chr.8q24.3 amplicon resident gene TONSL is upregulated during the initial steps of tumorigenesis to support neoplastic transformation by increasing DNA repair and represents a potential therapeutic target for treating breast cancer.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

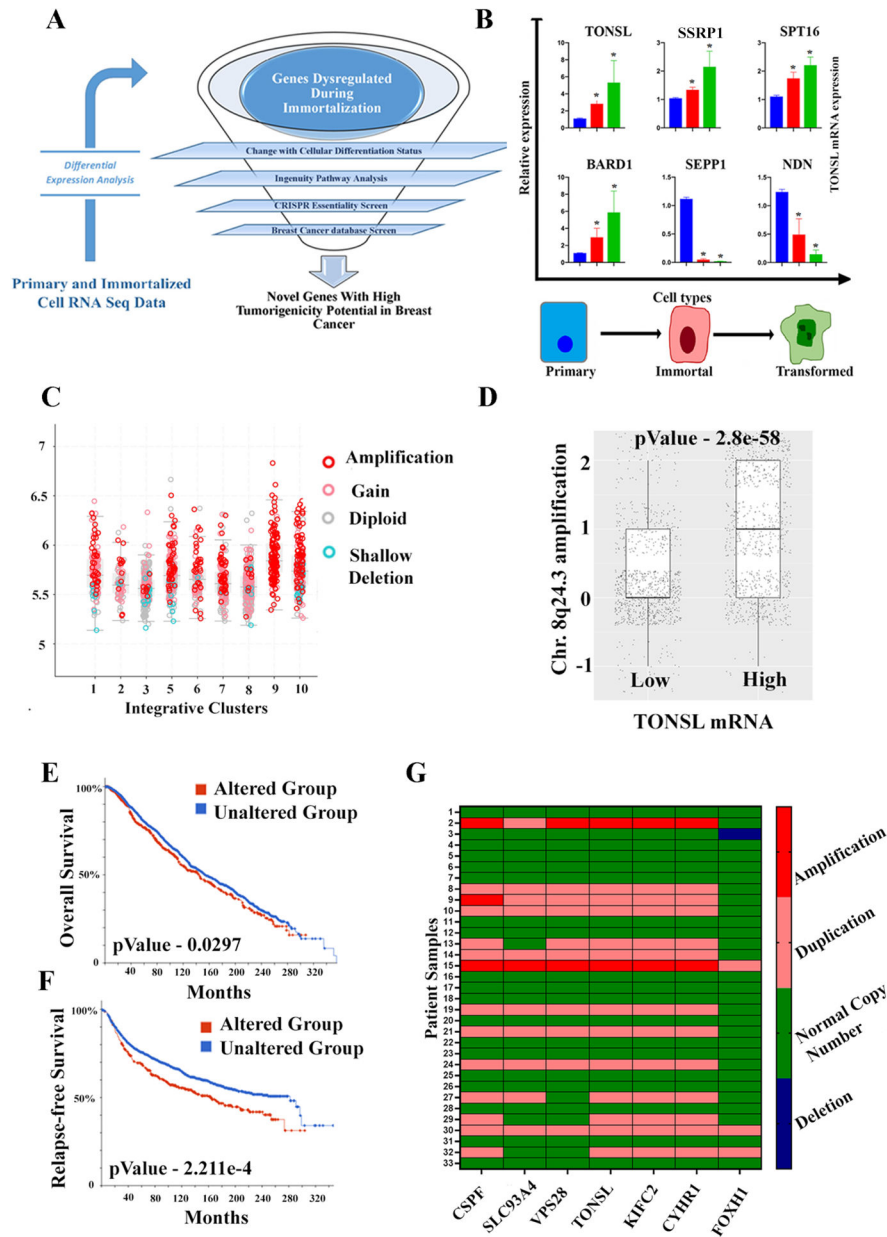


Fig. 1: Deregulated TONSL expression during immortalization and in breast cancer.
 A) Experimental scheme to identify genes aberrantly expressed during immortalization.
 B) Increased expression of TONSL and FACT components SSRP1 and SPT16 in TERT immortalized and transformed cells compared to primary cells. qRT-PCR was performed using indicated isogenic cell lines. * $p=0.05$.
 C) Breast tumors in Integrative cluster 9 demonstrate highest level of *TONSL* amplification. METABRIC datasets were used for this analysis.
 D) *TONSL* amplification correlates with elevated mRNA levels in breast cancers.
 E) *TONSL* amplification is associated with poor overall survival.
 F) *TONSL* amplification is associated with poor recurrence-free survival.
 G) Analysis of *TONSL* and adjacent genes for amplification in primary breast cancers using NanoString nCounter platform.

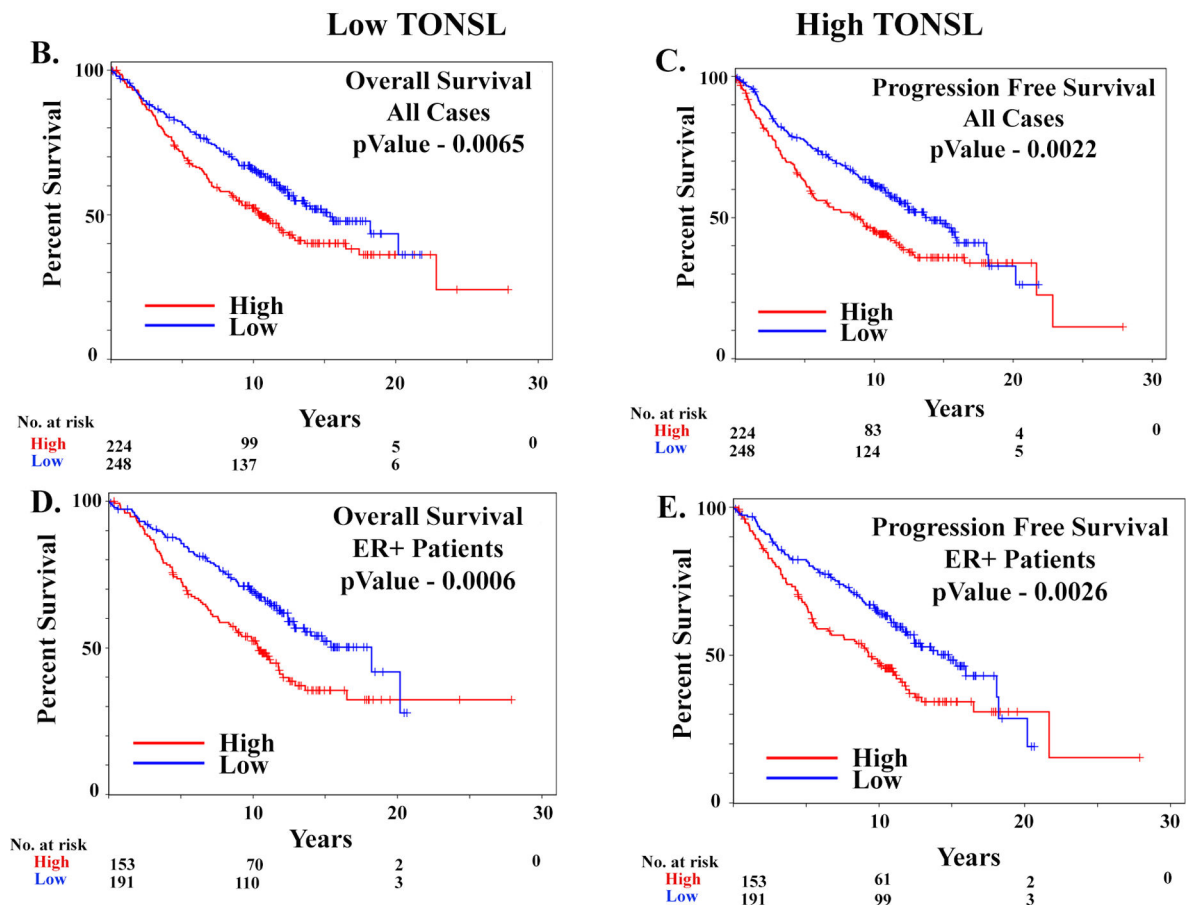
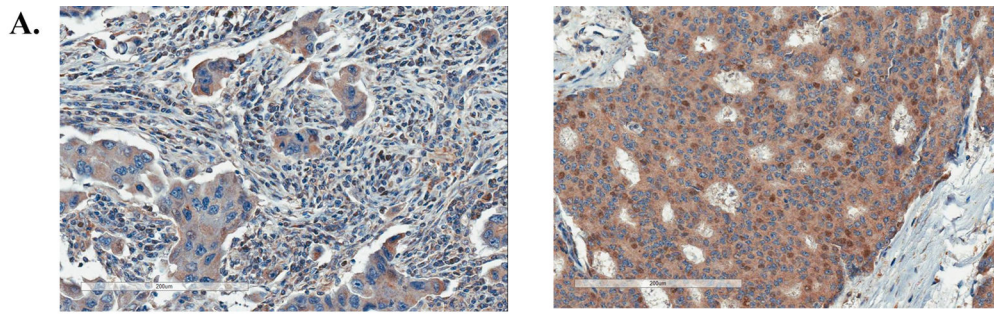


Fig. 2: TONSL overexpression in breast cancer is associated with poor overall and progression-free survival.

A) Representative TONSL staining patterns in breast cancer. B) TONSL overexpression is associated with poor overall survival. C) TONSL overexpression is associated with poor progression-free survival. D) TONSL overexpression is associated with poor overall survival of patients with ER+ breast cancer and treated with endocrine therapy. E) TONSL overexpression is associated with poor progression-free survival of patients with ER+ breast cancer and treated with endocrine therapy.

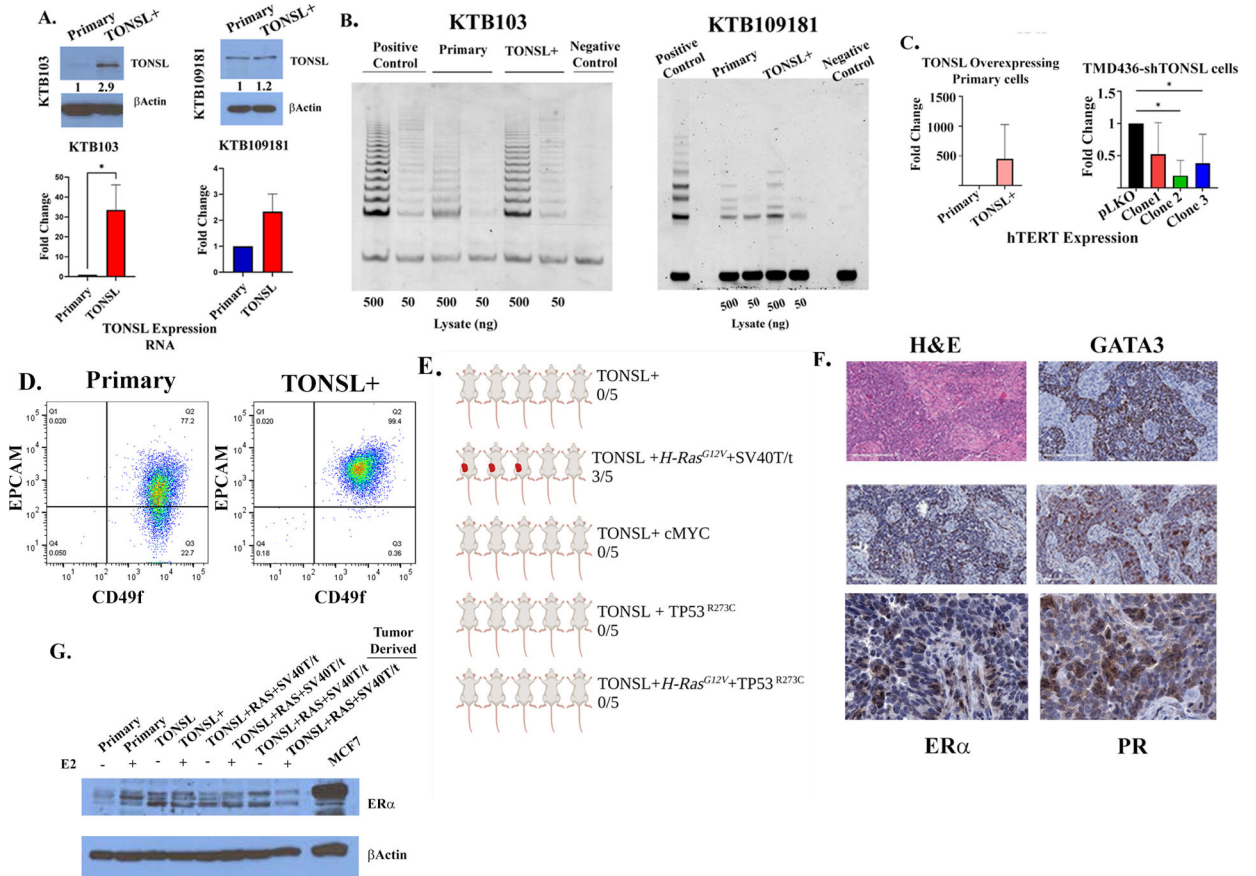


Fig. 3: TONSL is an immortalizing oncogene.

A) TONSL levels in primary cells and those infected with TONSL-overexpressing lentiviruses. Protein (top) and mRNA levels (bottom) were measured. (B) TRAP assay demonstrates elevated telomerase activity in TONSL overexpressing cells compared to parental cells. Results from two clones are shown. (C) TONSL overexpression had modest effects on hTERT mRNA levels in KTB103 cells. TMD436 shTONSL cells showed reduced hTERT levels compared to parental cells. qRT-PCR was used to measure hTERT levels. (D) KTB103 TONSL overexpressing cells are predominantly luminal progenitors based on CD49f and EpCAM staining pattern (CD49f+/EpCAM+). Primary KTB103 cells contained luminal progenitor and basal/stem cell subpopulation. (E) Only H-RAS^{G12V} plus SV40-T/t antigen overexpressing cells generated tumors in NSG mice. Number of mice injected and number of animals that developed tumors are indicated (Created with [BioRender.com](https://www.biorender.com)). (F) TONSL+H-RAS^{G12V}+SV40-T/t antigens expressing cells generate ER+/PR+ tumors in NSG mice. ER and PR staining (middle row) and enlarged view of ER and PR staining (bottom row) is shown. H&E staining pattern of a representative tumor is also shown. (G) Cells lines generated from tumors in F show ER α expression. Cells were treated with charcoal stripped media overnight and treated with vehicle (ethanol) and 10^{-10} M E₂ (Estradiol) for 3hrs.

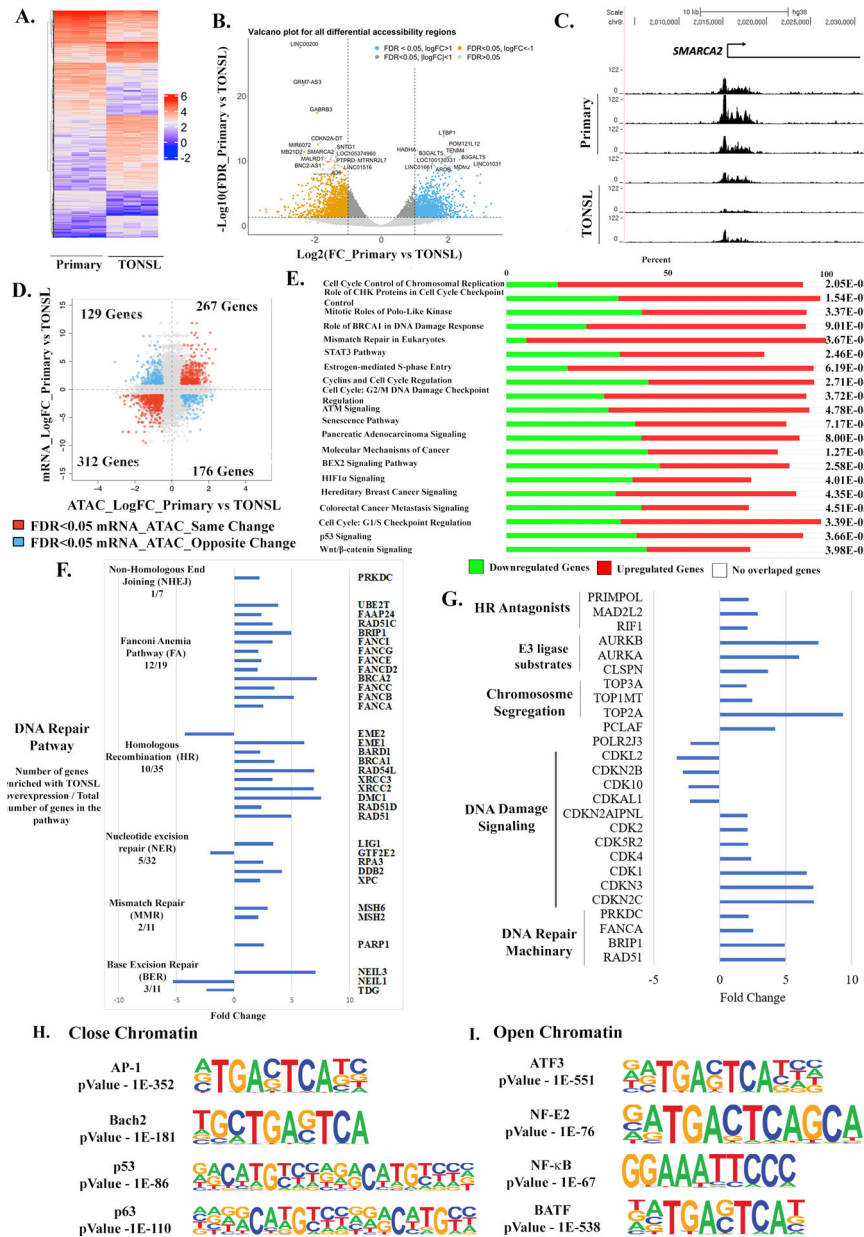


Fig. 4: TONSL induces chromatin reorganization and alters expression of genes associated with DNA repair hubs.

A) Heatmap of TONSL-induced chromatin accessibility changes, measured in triplicate. B) Volcano plot shows TONSL-induced chromatin opening and closing of select genes. C) Chromatin accessibility status of *SMARCA2* gene in primary and TONSL overexpressing cells. D) Integration of ATAC-seq and RNA-seq data shows correlation between chromatin accessibility changes and gene expression. E). Ingenuity pathway analysis reveal effects of TONSL overexpression on specific pathways. F) Genes in different DNA repair hubs affected by TONSL. Pathway genes enriched upon TONSL overexpression with p value < 0.01 and gene expression less than or equal to 2 and/or greater than or equal to 2 were plotted along with pathway name followed by number of genes enriched with TONSL/ Total number of genes involved in the pathway. TONSL increases FA and HR hubs but

reduces BER associated genes. G) TONSL overexpression alters the expression levels of BRCA1-BARD1 interactors and antagonists. Genes were selected as described above. H) Transcription factor binding site enrichment analysis of chromatin regions that became inaccessible upon TONSL overexpression. I) Transcription factor binding site enrichment analysis of chromatin regions that became accessible upon TONSL overexpression.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

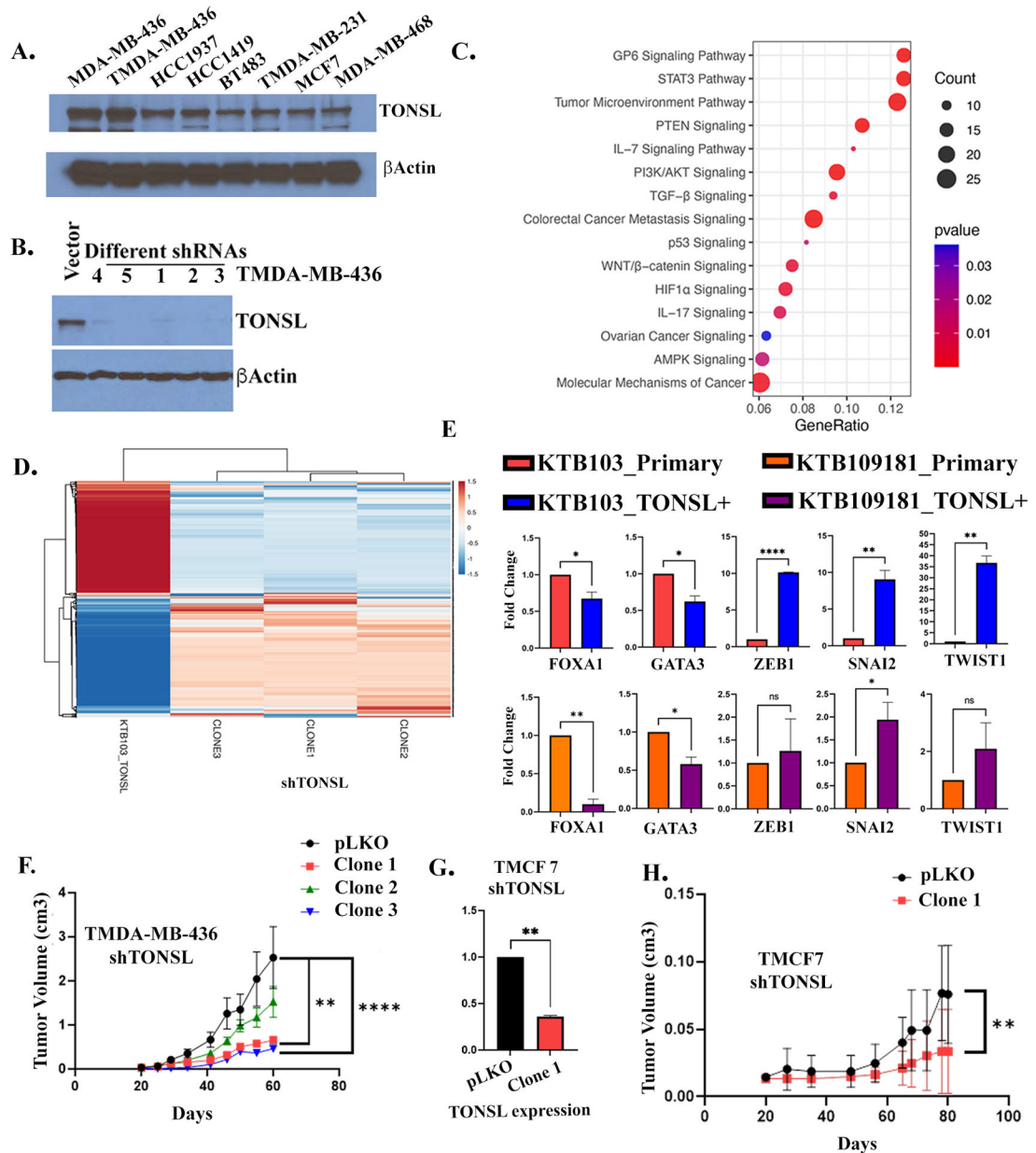


Fig. 5: TONSL is required for *in vivo* growth of TONSL amplified breast cancer cell lines.

A) Expression levels of TONSL in amplified and non-amplified cell lines. B) Knockdown of TONSL in TMD-436 cells through shRNA targeting different regions of *TONSL* gene. C) Ingenuity Pathway Analysis of genes differentially expressed in control shRNA (pLKO) compared to shTONSL clones reveal the role of TONSL in various signaling pathways in cancer cells. D) Overlap analysis of differentially expressed genes in two datasets; one dataset is from primary cells with and without TONSL overexpression and the other set from TMD-436 cells with and without TONSL shRNA. Heatmap shows genes that are induced and repressed by TONSL in both cell system and characterized through two independent means. E) Validation of RNA-seq data through qRT-PCR of select genes with TONSL overexpression. In primary cells, TONSL reduced the expression levels of luminal genes,

while increasing the levels of basal cell-enriched genes. Data from two clones. F) TONSL is essential for optimal growth of TMD-436 cells *in vivo*. Growth patterns of TMD-436 pLKO and shTONSL clones in the mammary fat pad of nude mice. 7–9 animals per group were used. G) Generation of TMCF-7 with TONSL knockdown. H) TONSL is required for the growth of MCF-7 cells *in vivo*. 8–9 animals per group were used. For animals injected with MCF-7, slow-release estrogen pellets were implanted. (ns - $P > 0.05$, * $P = 0.05$, ** $P = 0.01$, *** $P = 0.0001$).

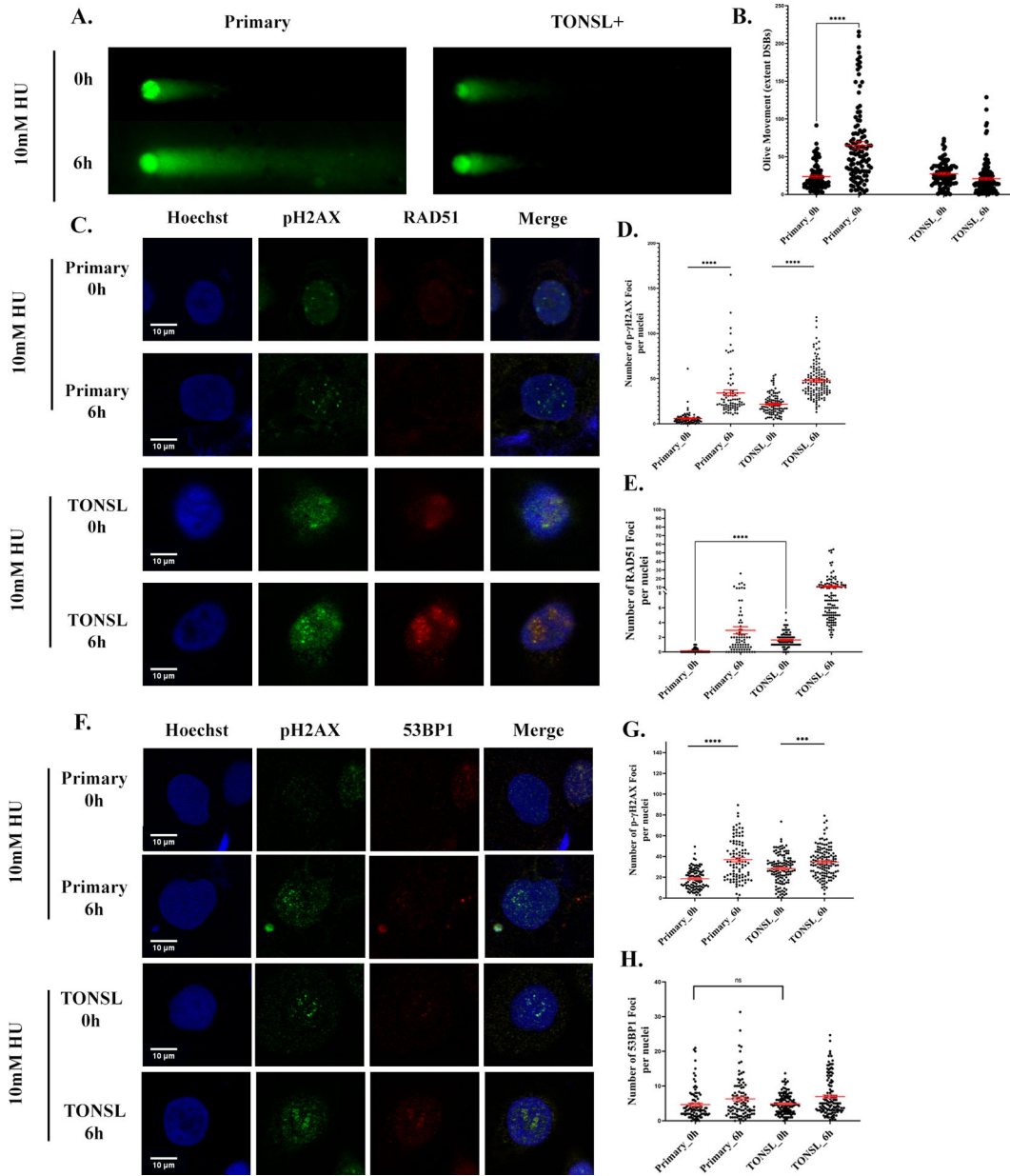


Fig. 6: Cells overexpressing TONSL are resistant to DNA damage.

A) Comet assay showing primary cells are more susceptible to HU treatment compared to TONSL overexpressing cells. B) Quantitative analysis of tail movement/olive movement of primary and TONSL+ cells upon treatment. HU causes significant DNA damage in primary cells after 6h treatment, whereas TONSL overexpressing cells are resistant to HU mediated DNA damage. C) Immunofluorescence images showing γ H2AX, RAD51 and nuclear stain Hoechst. TONSL+ cells display more RAD51 foci with treatment. D, E) Quantitative analysis of γ H2AX and RAD51 foci per nucleus. TONSL+ cells have significantly increased basal levels of RAD51. F) Immunofluorescence images showing γ H2AX, 53BP1 and nuclear stain Hoechst. No significant differences were observed in primary and TONSL+ cells. G, H) Quantitative analysis of γ H2AX and 53BP1 foci per

nucleus. No difference was observed with and without treatment in both cell types. (ns - $P > 0.05$, *** $P = 0.001$, **** $P = 0.0001$)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

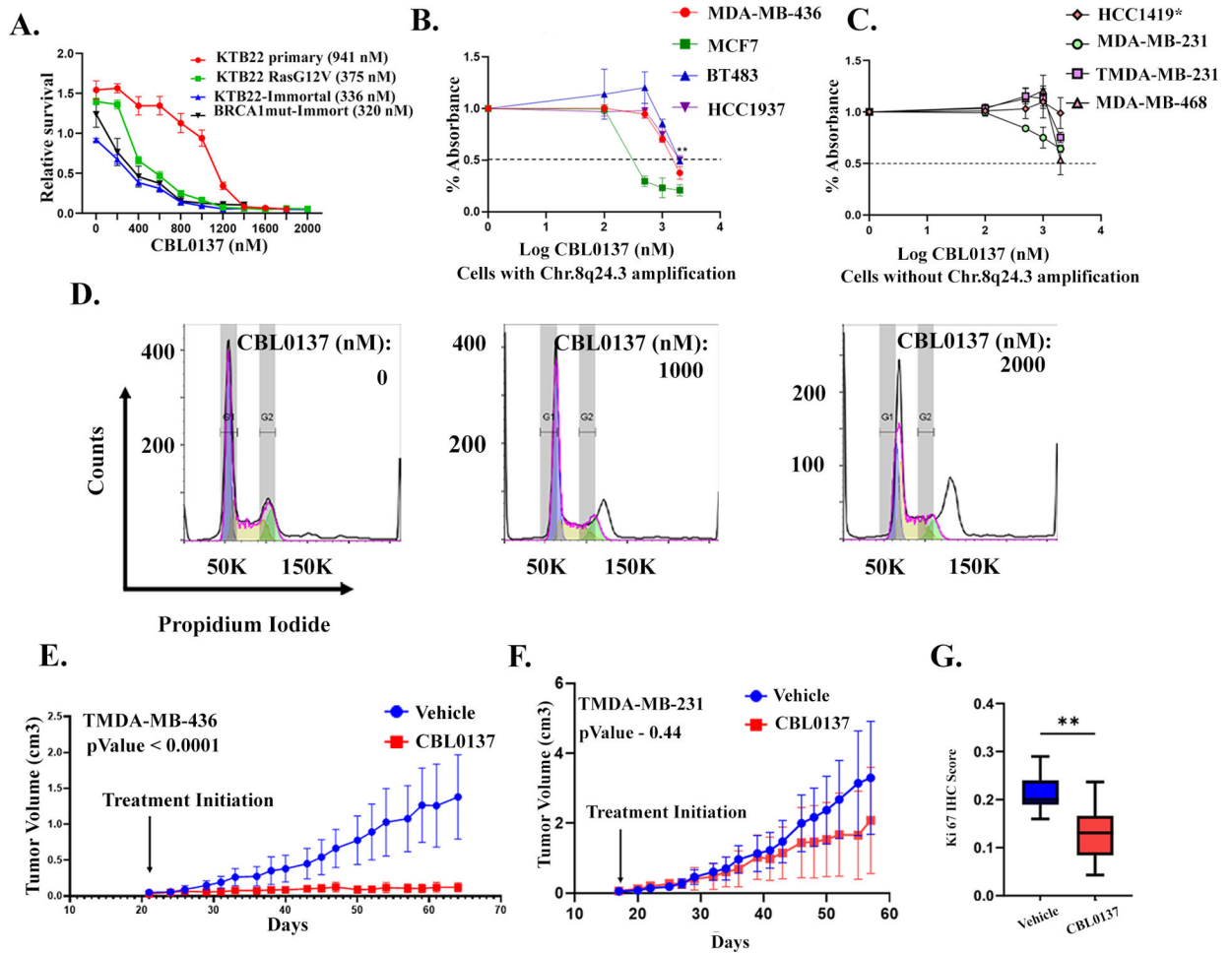


Fig. 7: Cells overexpressing TONSL are sensitive to FACT complex inhibitor CBL0137.
 A) Immortalized and transformed cells are more sensitive to FACT inhibitor CBL0137 compared to isogenic primary cells. BRCA1 mutant cells are also sensitive. Differences in sensitivity between primary cells and other cell lines were statistically significant. IC50 values are indicated. B and C) Breast cancer cell lines with *TONSL* amplification are more sensitive to CBL0137 than cell lines without *TONSL* amplification. (HCC1419 has 8q24.3 amplification but *TONSL* gene is disrupted by translocation). D) CBL0137 induces S and G2/M arrest of MD-436 cells. E and F) Tumors derived from TMD-436 but not TMD-231 cells are growth inhibited by CBL0137. Arrow indicates day of treatment initiation when palpable tumor was apparent. G) TMD-436 derived tumors in mice treated with CBL0137 show lower Ki67 positivity compared to tumors in vehicle treated mice. (**P = 0.01)

Table 1:

Upregulation of TONSL and FACT components in immortalized cells compared to primary cells. Fold changes in immortalized cells compared to primary cells from RNA-seq data ($p < 0.001$ and $FDR < 0.05$) are shown. RNA-seq was done in biologic triplicates and the original data are available publicly (GSE108541).

Gene	Latina ancestry			European ancestry			African Ancestry	Average
	KTB21	KTB22	KTB26	KTB34	KTB36	KTB37		
TONSL	7.87	4.21	6.07	3.74	6.70	15.10	8.18	7.41
BARD1	8.01	2.81	2.42	3.84	4.57	6.31	4.25	4.60
SSRP1	2.64	1.73	2.42	1.78	2.74	2.58	2.25	2.31
SPT16	2.01	1.56	1.70	1.52	1.77	1.40	1.66	1.66
LINC01116	14.49	15.90	19.89	14.65	22.53	44.51	1682.53	259.21
FGFBP1	243.25	17.36	46.31	1.41	63.01	566.37	190.60	161.19
SPRR1A	3.61	51.37	612.89	6.76	48.24	52.11	1.41	110.91
NMU	54.45	11.73	71.87	20.02	289.73	151.60	5.06	86.35
HPGD	17.57	8.72	60.36	101.81	62.36	57.33	-18.52	41.38
SEPP1	-270.01	-6.51	-34.78	-1950.4	-407.9	-183.6	-10323	-1832.3
NDN	-297.3	-881.6	-4.48	-3.84	-10.19	-675.8	-9.98	-269.02
FLJ41200	-532.54	-295.77	-1496.01	-131.72	-1496.43	-491.62	-216.32	-665.77
APBA2	-1698.44	-398.43	-53.94	-54.32	-249.63	-2080.98	-1.42	-648.17
COX7A1	-574.67	-984.58	-146.05	-4.18	-1965.11	-317.86	-20.50	-573.28
GYPC	-766.38	-276.38	-1817.27	-19.20	-233.11	-660.26	-21.24	-541.98