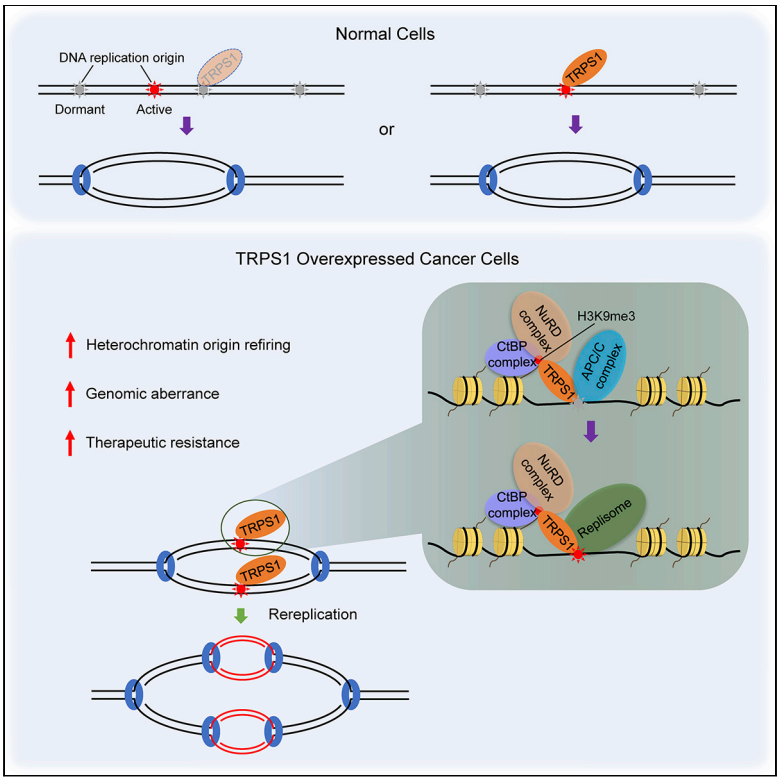
Cell Reports

Article

TRPS1 drives heterochromatic origin refiring and cancer genome evolution

Graphical Abstract



Authors

Jianguo Yang, Xiaoping Liu, Yunchao Huang, ..., Jing Liang, Yu Zhang, Yongfeng Shang

Correspondence

zhang_yu@hsc.pku.edu.cn (Y.Z.), yshang@hsc.pku.edu.cn (Y.S.)

In brief

The tricho-rhino-phalangeal syndrome-associated gene *TRPS1* is also amplified in breast cancer. By leveraging the genetic mutations, Yang et al. identify an unexpected replication boosting role of TRPS1 for H3K9me3-marked heterochromatic origin activation, cancer genome evolution, and therapeutic resistance, thus revealing a critical aspect of TRPS1 in driving breast carcinogenesis through over-replication.

Highlights

- TRPS1 interacts with replication machinery and binds to H3K9me3-marked origins
- TRPS1 encourages chromatin loading of APC/C leading to uncontrolled origin refiring
- TRPS mutations abolish TRPS1-promoted replication boosting and breast carcinogenesis
- TRPS1 overexpression drives cancer genome evolution and therapeutic resistance

Yang et al., 2021, Cell Reports *34*, 108814
March 9, 2021
https://doi.org/10.1016/j.celrep.2021.108814

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

TRPS1 drives heterochromatic origin refiring and cancer genome evolution

Jianguo Yang,[1] Xiaoping Liu,[1] Yunchao Huang,[1] Lin He,[1] Wenting Zhang,[1] Jie Ren,[1] Yue Wang,[2,3] Jiajing Wu,[3] Xiaodi Wu,[3] Lin Shan,[3] Xiaohan Yang,[1] Luyang Sun,[1] Jing Liang,[1] Yu Zhang,[1,*] and Yongfeng Shang[1,2,3,4,*]

[1]Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Peking University Health Science Center, Beijing 100191, China

[2]Department of Biochemistry and Molecular Biology, School of Medicine, Hangzhou Normal University, Hangzhou 311121, China

[3]Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Capital Medical University, Beijing 100069, China

[4]Lead contact

*Correspondence: zhang_yu@hsc.pku.edu.cn (Y.Z.), yshang@hsc.pku.edu.cn (Y.S.)

https://doi.org/10.1016/j.celrep.2021.108814

SUMMARY

Exploitation of naturally occurring genetic mutations could empower the discovery of novel aspects of established cancer genes. We report here that *TRPS1*, a gene linked to the tricho-rhino-phalangeal syndrome (TRPS) and recently identified as a potential breast cancer driver, promotes breast carcinogenesis through regulating replication. Epigenomic decomposition of TRPS1 landscape reveals nearly half of H3K9me3-marked heterochromatic origins are occupied by TRPS1, where it encourages the chromatin loading of APC/C, resulting in uncontrolled origin refiring. TRPS1 binds to the genome through its atypical H3K9me3 reading via GATA and IKAROS domains, while TRPS-related mutations affect its chromatin binding, replication boosting, and tumorigenicity. Concordantly, overexpression of wild-type but not TRPS-associated mutants of TRPS1 is sufficient to drive cancer genome amplifications, which experience an extrachromosomal route and dynamically evolve to confer therapeutic resistance. Together, these results uncover a critical function of TRPS1 in driving heterochromatin origin firing and breast cancer genome evolution.

INTRODUCTION

Despite the enormous amount of data from cancer genomic studies, identification of driver genes remains a major effort in the field of cancer research, and discerning oncogenic drivers from passengers remains a major bottleneck in understanding of the essence of cancer initiation and development. The challenge lies mainly in the genetic composition of distinct tumor subclones that undergo extensive diversifications during tumor development and progression; thus, tremendous genetic heterogeneities within tumors have been documented with the great advances in the depth and resolution of high-throughput sequencing. Ultimately, potential drivers identified by various sources and via different means need mechanistic exploration and experimental validation.

In a recent study that integrated genomic data from primary breast cancers with data from functional RNAi screens to pinpoint potential driver genes within large recurrently amplified regions of DNA, a set of candidate drivers highly enriched with known drivers were identified (Sanchez-Garcia et al., 2014). One of the candidates is *TRPS1*, a gene associated with the tricho-rhino-phalangeal syndrome (TRPS) (Lüdecke et al., 2001; Momeni et al., 2000), an inherited autosomal dominant disorder characterized by a triad of sparse hair, peculiar pear-shaped nose, and brachydactyly, as well as other skeletal abnormalities (Giedion et al., 1973). Indeed, it is reported that *TRPS1* is overexpressed in breast carcinomas (Radvanyi et al., 2005) and osteosarcomas (Li et al., 2015). It is also noted that TRPS is associated with mental retardation and endocrine abnormalities, bilateral renal hypodysplasia, and possibly malformations/malfunctions of other organs (Tasic et al., 2014), suggesting that TRPS1 might be involved in fundamental cellular process(es).

Despite the importance of *TRPS1* in animal pathophysiology, the mechanistic insight into the function of TRPS1 is limited. TRPS1 contains 9 zinc fingers scattering from its N to C terminus (Chang et al., 2000): one resembling the GATA domain and another two analogous to the IKAROS domain (Momeni et al., 2000). Point mutations (Lüdecke et al., 2001) as well as missense mutations (Smaili et al., 2017) in GATA domain are frequently detected in TRPS patients, and the deletion of GATA motif is sufficient to induce TRPS phenotypes in mouse models (Malik et al., 2002), highlighting the importance of GATA motif in the development of TRPS. At the molecular level, TRPS1 has been mainly investigated in transcriptional repression through its association with an array of well-characterized corepressors (Witwicki et al., 2018). However, interestingly, TRPS1 silencing had limited effects on the transcriptome in MCF-7 cells (Wang et al., 2018a), and evidence is emerging to suggest a multifaceted functionality for TRPS1. Indeed, a recent report found a context-dependent regulation of TRPS1 in normal mammary epithelial cell differentiation and breast cancer development (Cornelissen et al., 2020). Clearly, the oncogenic potential of TRPS1 needs experimental validation, and the

molecular mechanism underlying its pathophysiological function needs further investigation.

We investigated in this study the molecular function of TRPS1 in breast carcinogenesis. We identified an unexpected link between TRPS1-evoked heterochromatic origin replication and genomic aberrance attributable to atypical reading of H3K9me3 by TRPS1. We showed that TRPS1 is sufficient to drive cancer genome amplification, generating extrachromosomal circular DNAs (eccDNAs) and intratumoral heterogeneity to confer therapeutic resistance.

RESULTS

TRPS1 is amplified in breast cancer and promotes breast carcinogenesis

We were intrigued by the recent Helios algorithm identification of *TRPS1* as a candidate driver for breast cancer (Sanchez-Garcia et al., 2014). Integrative analysis of *TRPS1* in various cancer types revealed a wide spread of *TRPS1* amplification in breast cancer samples and xenograft models (Figures S1A and S1B), along with significantly overexpressed mRNA levels (Figure S1C). We also performed immunohistochemical staining of TRPS1 using human cancer arrays containing 52 breast carcinoma samples paired with normal mammary tissues and found that TRPS1 protein level significantly elevated in breast carcinomas (Figure 1A), and analysis by western blotting for the expression of TRPS1 showed that the level of TRPS1 is much higher in a panel of breast cancer cell lines than in the human normal mammary epithelial cell line MCF-10A (Figure 1B).

To test the role of TRPS1 in breast carcinogenesis, MCF-10A and MDA-MB-231 cells stably expressing TRPS1 were generated. Measurement by MTS assays showed that TRPS1 overexpression was associated with a significant increase in cell proliferation (Figure 1C). In agreement, stable knockdown of TRPS1 in MCF-7 cells resulted in a significant decrease in cell proliferation (Figure 1D), although TRPS1 depletion had no discernible effect on the senescence and death of cells (Figures S1D and S1E). In addition, mouse xenograft experiments with MDA-MB-231-Luc-D3H2LN-based bioluminescent assays showed that TRPS1 promoted not only the growth of primary tumors (Figure 1E) but also the metastasis of the tumors to lung, liver, and foreleg (Figure 1F). Together, these results indicate that TRPS1 promotes the development and progression of breast cancer.

TRPS-associated mutations abolish the oncogenic potential of TRPS1

To further understand the tumorigenic role of TRPS1, we measured the proliferation of stably transduced MCF-10A cells that were infected with lentivirally delivered TRPS1 truncation mutants (Figures S1F–-S1H). Compared to wild-type (WT) TRPS1, deletion mutants were all associated with a decreased proliferation of MCF-10A cells, albeit to a different extent, with the most significant effect seen in TRPS1ΔGATA (Figure 1G). This scenario agrees with the importance of the GATA domain of TRPS1 in the pathogenesis of TRPS (Malik et al., 2002), suggesting that this module is critical for the pathophysiological function of TRPS1, including its tumorigenic potential.

We thus took advantage of the mutational information from TRPS patients (Table S1) to dissect the molecular function of TRPS1 in breast carcinogenesis. Interestingly, a single amino-acid substitution of cysteine at 896 with tyrosine (C896Y) or arginine at 908 with glutamine (R908Q), two most frequent mutations within the GATA domain seen in TRPS, had greatly compromised effects on the proliferation of MCF-10A and MDA-MB-231 cells (Figure 1C). The importance of these mutations was further verified in mouse xenograft models, in which MDA-MB-231-Luc-D3H2LN tumors overexpressing either TRPS1C896Y or TRPS1R908Q showed compromised effects on promoting the growth of primary tumor (Figure 1E) and distal metastasis (Figure 1F). The above results support the importance of the GATA domain in the pathophysiological function of TRPS1, including its role in breast carcinogenesis.

TRPS1 interacts with the CtBP/NuRD complexes and replication machinery

To understand the molecular basis underlying the tumorigenic potential of TRPS1, we next interrogated the TRPS1 interactome *in vivo* by affinity purification and mass spectrometry. TRPS1 was co-purified with a series of corepressor proteins, including CtBP1/2, components of the CtBP complex, Mi-2β and MTA1/2, subunits of the NuRD complex, as well as HDAC1/2, constituents of both the CtBP and NuRD complex (Figure 2A). Unexpectedly, a large collection of DNA replication factors, including members of the MCM helicase complex MCM3-7, and ORC2, CDC45, RPA1, POLD1/2, PCNA, and RFC2/3/4, as well as factors functioning in origin licensing/firing or elongating replisome, were also detected in TRPS1 interactome (Figure 2A). Remarkably, even some regulatory proteins of DNA replication, such as components of APC/C (Zielke et al., 2008) and FACT, were also identified in TRPS1 interactome (Figure 2A; Table S2). The presence of these proteins in TRPS1-containing complex was verified by western blotting of the affinity elutes with antibodies against the corresponding proteins (Figure 2A, right panel). Notably, the chromatographic profile of TRPS1 in MCF-7 cells was deviated from its monomeric fractions and overlapped with that of the representative subunits of the CtBP/NuRD complexes and replication machinery (Figure S2A). Co-immunoprecipitation experiments detecting endogenous proteins in MCF-7 cells further strengthened this point (Figure 2B). In addition, consistent with the expression pattern of TRPS1 (Figures S2B and S2C), the interaction between TRPS1 and replication factors was also tissue specific (Figure S2D). Moreover, co-immunoprecipitation assays in synchronized cells demonstrated that while TRPS1 was co-purified with the representative APC/C and MCM components, in contrast to its barely detected interaction with RFC and PCNA in the G_1 phase (Figure 2C; Figures S2E and S2F), the association between TRPS1 and replisome components increased in the S phase (Figure 2C). Interestingly, with an antibody against phospho-MCM2 (Ser40) to probe MCM helicase activation (Montagnoli et al., 2006), we found that, in line with MCM inactivation in G_1, only the interaction between TRPS1 and unphosphorylated MCM2 could be detected in this phase (Figure S2G), implying that TRPS1-interacting replication factors in G_1 was not originated from moving replisomes. Collectively, these results indicate that TRPS1 interacts with CtBP/NuRD and the replication machinery *in vivo*.

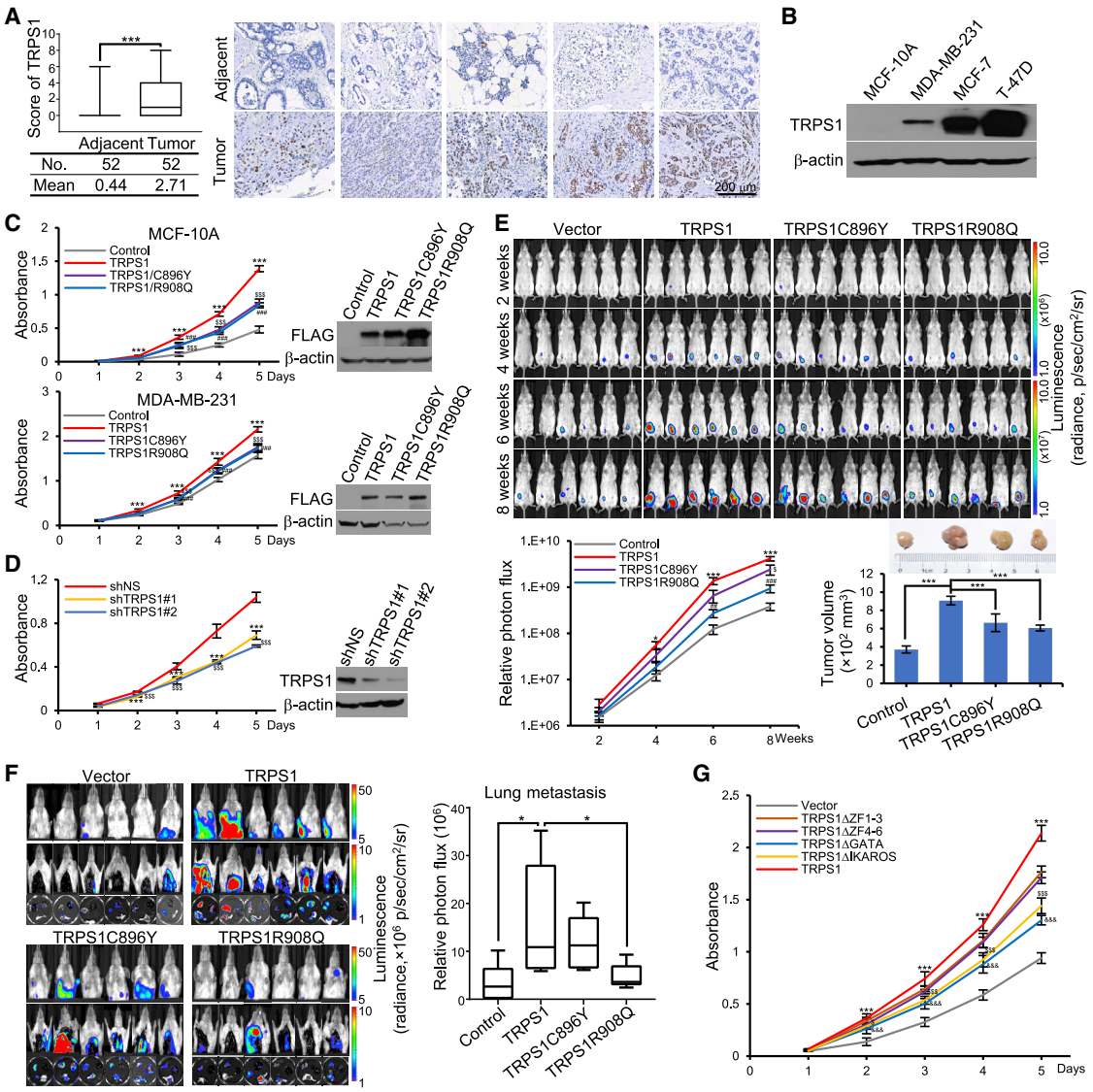


Figure 1. TRPS1 is amplified/overexpressed in breast cancer and promotes breast carcinogenesis

(A) Tissue array analysis for TRPS1 expression (left). ***p < 0.001; two-tailed paired t test. Representative images are shown (right).

(B) Western blotting analysis of TRPS1 expression in the indicated cell lines.

(C) MCF-10A or MDA-MB-231 cells were infected with the indicated lentiviruses for cell-proliferation measurements by MTS assays. Overexpression efficiency was verified by western blotting. TRPS1 versus control, ***p < 0.001; C896Y versus WT, $^{\$\$\$}$p < 0.001; R908Q versus WT, $^{\#\#\#}$p < 0.001; one-way ANOVA.

(D) Proliferation measurements of MCF-7 cells infected with the indicated knockdown lentiviruses by MTS assays. Knockdown efficiency was verified by western blotting. shTRPS1#1 versus shNS, ***p < 0.001; shTRPS1#2 versus shNS, $^{\$\$\$}$p < 0.001; t test.

(E) Xenograft models using MDA-MB-231-Luc-D3H2LN cells infected with the indicated lentiviruses were subjected to bioluminescent imaging analysis. The volume of primary tumors was quantified, and the bioluminescent images are shown. Each bar represents the mean ± SD. left: TRPS1 versus control, *p < 0.05, ***p < 0.001; C896Y versus WT, $^{\$}$p < 0.05; R908Q versus WT, ##p < 0.01, $^{\#\#\#}$p < 0.001. right: ***p < 0.001, n = 6. One-way ANOVA.

(F) Metastases of MDA-MB-231-Luc-D3H2LN tumors to lung, liver, and foreleg. Metastases to lung were quantified and representative bioluminescent images are shown. *p < 0.05, n = 6, one-way ANOVA.

(G) MCF-10A cells were infected with the indicated lentiviruses for cell-proliferation measurements by MTS assays. TRPS1 versus vector, ***p < 0.001; ΔIKAROS versus WT, $^{\$}$p < 0.05, $^{\$\$\$}$p < 0.001; ΔGATA versus WT, $^{\&\&\&}$p < 0.001; one-way ANOVA. Error bars represent mean ± SD for triplicate experiments in (C), (D), and (G).

Importantly, comparative interactomic analysis indicated that, while the association of TRPS1 with CtBP/NuRD complexes was retained for TRPS1C896Y and TRPS1R908Q, these mutants lost their ability to interact with replication factors, especially in S-phase cells (Figure 2D; Figure S2H). Immunoblotting with antibodies against the corresponding proteins confirmed this observation (Figure 2E). These results imply that the decreased potential of TRPS-associated mutants to drive cancer cell growth

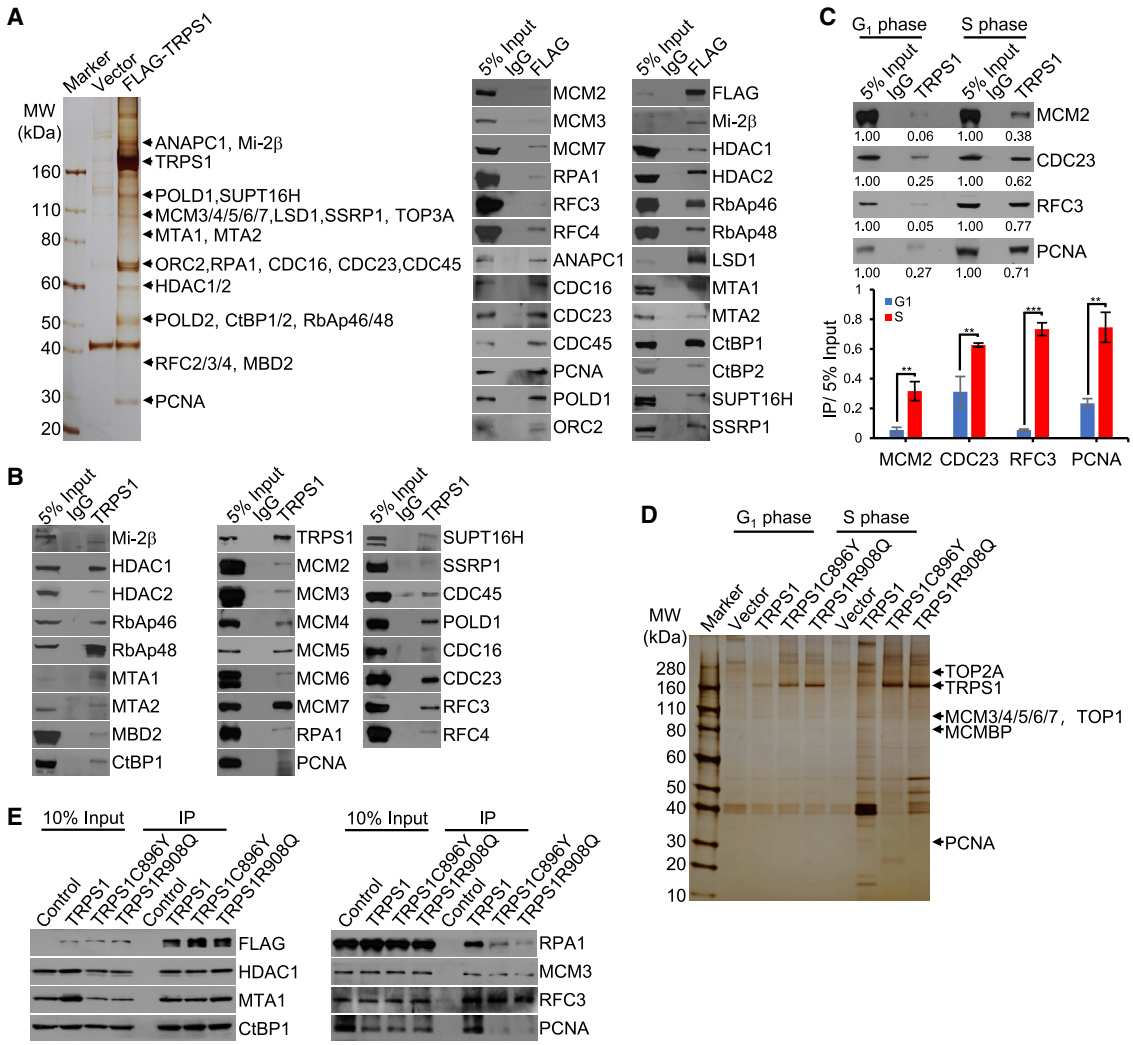


Figure 2. TRPS1 is co-purified with CtBP/NuRD and the replication machinery

(A) TRPS1 interactome was affinity purified from the nuclear extract of HEK293T cells stably expressing FLAG-TRPS1, resolved on SDS-PAGE, and silver stained. The protein bands were retrieved and analyzed by mass spectrometry (left). Co-immunoprecipitation with anti-FLAG followed by immunoblotting with antibodies against the indicated proteins (right).

(B) Co-immunoprecipitation experiments in MCF-7 cells with antibodies against TRPS1 followed by immunoblotting with antibodies against the indicated proteins.

(C) Co-immunoprecipitation experiments in MCF-7 cells synchronized in the G_1 phase by serum starvation or S phase by double thymidine blocking (see also Figures S2E and S2F) (top). Bar plot quantifies triplicate experiments (bottom). Error bars represent mean ± SD. $^{**}p < 0.01$, $^{***}p < 0.001$, t test.

(D) MCF-7 cells stably integrated with the indicated constructs were synchronized in the G_1 phase or S phase for affinity purification and mass spectrometric analysis (see also Figure S2H).

(E) Co-immunoprecipitation with anti-FLAG in S-phase cells described in (C) followed by immunoblotting with antibodies against the indicated proteins.

and metastasis stems from their impaired interaction with the replication machinery, but not with CtBP/NuRD complexes, suggesting a link between TRPS1-promoted breast carcinogenesis and replication abnormalities.

TRPS1 is enriched in H3K9me3-marked replication origins

We next profiled the genomic landscape of TRPS1 in MCF-7 cells using chromatin immunoprecipitation sequencing (ChIP-seq). A total of 10,935 TRPS1 binding peaks were called, with the majority located in distal intergenic regions (62.6%) and gene bodies (37.4%) (Figure 3A). We further categorized the epigenome in MCF-7 cells with ChromHMM (Ernst and Kellis, 2012) into 7 distinct epigenetic domains (E1–E7) (Zhang et al., 2016) and classified TRPS1 binding peaks according to their residence within these domains (Figure 3B). The largest epigenetic compartment enriched in TRPS1 (E2) was characterized by lack of any detectable histone modifications, and the second largest compartment (E3, ~30%) was co-enriched with H3K9me3 (Figure 3B), an epigenetic mark functionally linked to

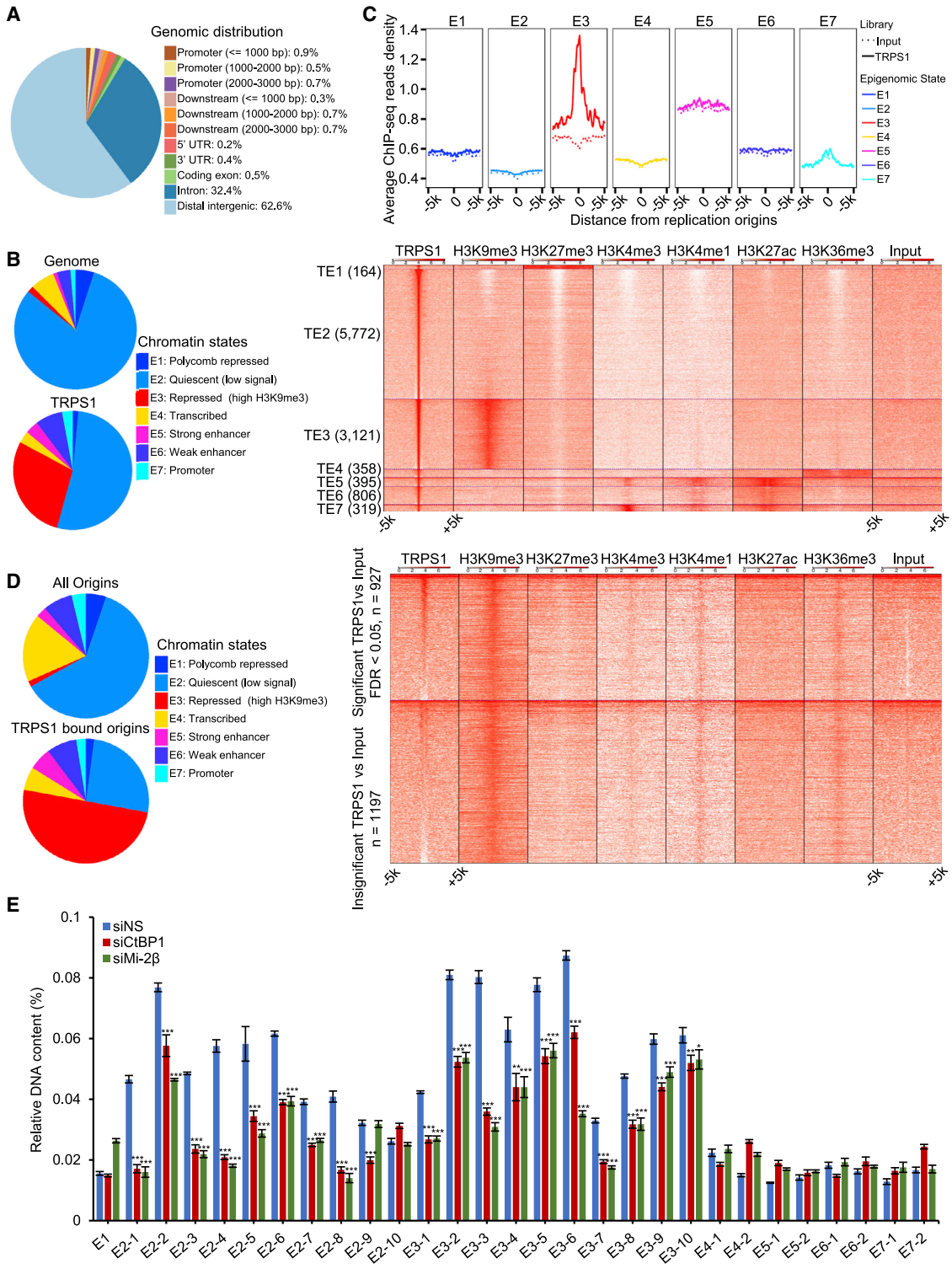


Figure 3. TRPS1 enriches in H3K9me3-marked replication origins

(A) ChIP-seq analysis of the genomic landscape of TRPS1 in MCF-7 cells.

(B) Epigenome analysis of TRPS1 binding in MCF-7 cells. Pie charts represent the fractions of TRPS1 peaks or the whole genome overlapped with the indicated epigenetic domains (left). Heatmaps represent the average ChIP-seq reads density (log2 transformed) of the indicated profiles within ~10-Kb genomic windows centered on TRPS1-binding peaks (right). The color key on top indicates ChIP-seq reads density.

(legend continued on next page)

the CtBP/NuRD complexes, whereas the fraction of TRPS1 peaks in transcriptionally active regions was negligible (Figure 3B). Quantitative ChIP (qChIP) analysis in MCF-7 cells on randomly selected sites of E1–E7 confirmed a strong enrichment of TRPS1 in E2 and E3 (Figure S3A).

To test whether TRPS1 is functionally implicated in DNA replication, we projected the TRPS1 binding signals into the actually fired origins previously identified in MCF-7 cells through high-throughput sequencing of RNA-primed nascent DNAs (Martin et al., 2011; Weddington et al., 2008; Figure S3B). While a clear gross pattern of the co-enrichment of TRPS1 along the whole set of origins could not be drawn, the epigenomic classification strategy exposed a specific enrichment of TRPS1 in H3K9me3-decorated (E3) origins (Figure 3C). Using a false discovery rate (FDR) cutoff of 0.05, we found that nearly half of the fired H3K9me3-marked origins were bound by TRPS1 within the whole set of windows centered at E3-type origin midpoints (Figure 3D; Figures S3C and S3D). Further comparisons of the epigenetic features in origin-containing relative to origin-free E3-type TRPS1 binding sites revealed a significant co-enrichment of H3K4me1 and H3K36me3 at TRPS1-bound E3 origins (Figure S3E), suggesting that TRPS1 binding sites associated with poised weak enhancers or intragenic regions of lowly expressed genes are more likely to be origins within H3K9me3-marked epigenetic domain (Consortium, 2012).

To probe the possible role of the CtBP/NuRD complexes in TRPS1 binding to heterochromatic origins, qChIP showed that depletion of either CtBP1 or Mi-2β resulted in a marked reduction of TRPS1 enrichment on both E2- and E3-type origins (Figure 3E), supporting a function of the CtBP/NuRD complexes in creating/maintaining the heterochromatic state, thereby facilitating TRPS1 loading onto origins residing in these repressive domains.

TRPS1 promotes origin refiring at heterochromatic domains

To determine whether the enrichment of TRPS1 in H3K9me3-marked origins play any active roles in DNA replication, we first compared the amount of nascent DNAs within a 1-Mb window centered at TRPS1 summits stratified by different epigenetic states to the amount of genomic background using the published Repli-seq data in MCF-7 cells that were sorted into 6 fractions spanning the full DNA synthesis phase (G_1b: [G_1/S boundary], S_1, S_2, S_3, S_4, and G_2) (Hansen et al., 2010). We found that nascent DNA productions surrounding TRPS1-bound repressive domains marked by H3K9me3 (E3) or H3K27me3 (E1), or domains devoid of any histone modifications (E2) or strongly activated enhancers (E5) significantly elevated (>3 × SD) at early S phase (Figure 4A, G_1b, S_1, S_2; Figure S4A), suggesting that TRPS1 promotes replication at specific epigenetic domains in early stage of DNA duplication. Consistently, TRPS1 interaction with MCM2 and MCM3 enhanced in early S-phase cells, albeit moderately (Figure S4B).

Considering the large size of replication bubbles versus short stretch of TRPS1 binding sites, increased DNA replication in the 1-Mb window could be a direct, active TRPS1 contribution to origin firing at the central summit. Alternatively, elevated DNA synthesis could be passively contributed by adjacent TRPS1-potentiated replication machinery. To distinguish these two scenarios, we first counted the number of active origins in windows centered at the same TRPS1 peaks but with different size ranges from 0.5 Kb to 1 Mb within each of the 7 epigenomic subgroups. H3K9me3-marked TRPS1 binding sites (E3) had significantly more active origins under a finer scale (Figure 4B), whereas the low signal E2 sites as the largest portion of TRPS1 binding showed an inverse pattern, with more active origins at a coarser scale, implying a passive model for the increased replication within the bulky E2-type replication unit by embedding TRPS1-bound E3-type origins. In agreement, calculation of the observed number of active origins normalized by the expected fractions of epigenomic states revealed that the enhanced DNA replication at any of the other 6 epigenomic subgroups was passively contributed by their inclusion of the H3K9me3-marked and TRPS1-bound origins (Figure 4C). Together, these findings suggest that TRPS1-encouraged DNA synthesis was originated from H3K9me3-decorated replication origins.

We then performed genome-wide analysis of replication by Repli-seq in TRPS1-depleted MCF-7 cells. Comparison of nascent DNAs at early, middle, or late S phase within each of the TRPS1-bound epigenomic subgroups revealed that TRPS1 deficiency resulted in a marked decrease in DNA replication at repressive domains (E2 and E3) at early S phase (Figure 4D), whereas the other epigenomic subgroups showed grossly unchanged DNA synthesis, supporting a specific and active contribution to DNA replication by TRPS1. Consistently, TRPS1 knockdown in MCF-7 cells led to cell-cycle arrests in the G_1 phase, whereas TRPS1 overexpression in MCF-10A cells rendered an increased population of cells in the S phase (Figure 4E). Moreover, TRPS1-boosted replication was compromised in cells overexpressing TRPS1C896Y or TRPS1R908Q (Figure 4E). In concordance, in chondrogenic ATDC5 cells, a cell line used for modeling the molecular contribution of TRPS1 to TRPS-associated defects in cartilage formation (Itoh et al., 2008), deficient of TRPS1 also resulted in reduced DNA replication (Figure S4C). These results point to a possible link of replication dysregulation to the pathogenesis of TRPS and breast cancer.

Further measurement of DNA replication by EdU incorporation showed that TRPS1-deficient cells had much less nuclei with more than 4N than control cells (Figure 4F), supporting a notion

(C) HOMER analysis of average ChIP-seq reads density of TRPS1 based on the stratification by epigenomic states at the actually fired origins that were identified in MCF-7 cells by others.

(D) Fractions of TRPS1-bound origins or the whole set of genomic origins stratified by their epigenomic states (left). Heatmaps represent the average ChIP-seq reads density (log2 transformed) for the indicated profiles within a ~10-Kb genomic windows centered on H3K9me3-marked active replication origins. These origins were classified into TRPS1-bound or unbound based on their enrichment of TRPS1 relative to input with an FDR cutoff of 0.05 (right).

(E) qChIP measurement of the enrichment of the indicated TRPS1-bound origins with antibodies against TRPS1 in MCF-7 cells transfected with control or the indicated small interfering RNAs (siRNAs).

Error bars represent mean ± SD for triplicate experiments. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; t test.

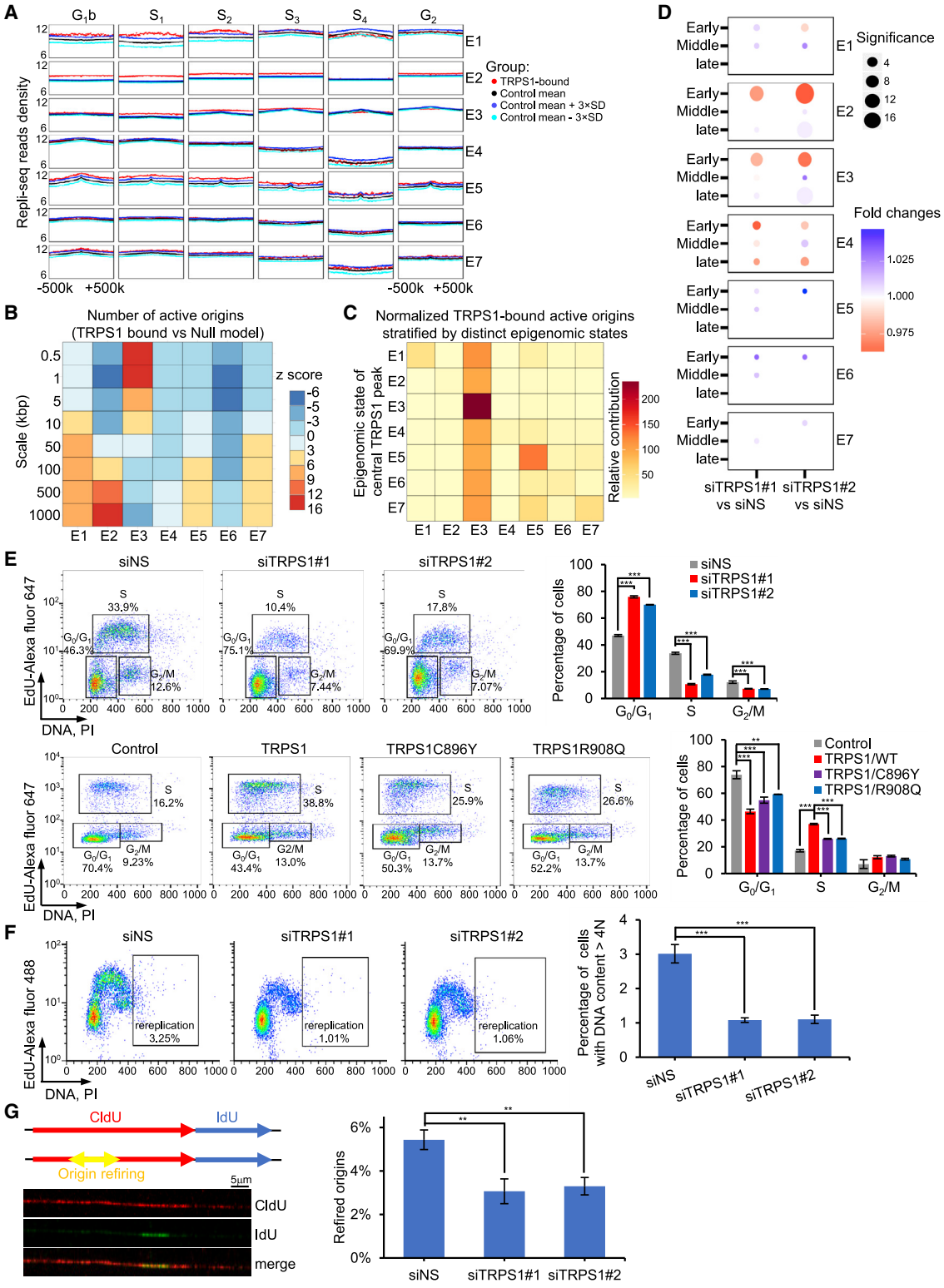


Figure 4. TRPS1 promotes heterochromatic origin refiring

(A) The average Repli-seq read density surrounding TRPS1 peaks stratified by differential epigenetic states. The mean and standard derivations of replication efficiency were calculated based on the background model with 100 groups of randomly shuffled genomic sites with a similar peak length distribution and the same epigenomic state as TRPS1 subgroups.

(legend continued on next page)

that TRPS1 is functionally linked to DNA re-replication. In addition, pulse-labeling MCF-7 cells with 5-chloro-20-deoxyuridine (CldU, red) followed by incorporating 5-Iodo-20-deoxyuridine (IdU, blue), a procedure allowing detection of reactivated origins in replicated DNAs as short yellow signals within long red track by microscopy (Muñoz et al., 2017), revealed that TRPS1 depletion was associated with a significant decrease in origin refiring (Figure 4G). Moreover, CsCl density gradient centrifugation-based partitioning of BrdU-labeled genomic DNAs (Black et al., 2013) revealed that TRPS1-overexpressing cells had a much higher level of re-replicated (heavy/heavy) DNAs than control cells (Figure S4D). Together, these results support the notion that TRPS1 plays an active role in DNA replication by promoting refiring of H3K9me3-marked origins.

TRPS1 incites origin refiring through recruitment of APC/C complex

To investigate the molecular mechanism underlying TRPS1-evoked origin refiring, we first examined the dynamic enrichment of replication factors on chromatin upon TRPS1 removal by chromatin fractionation assays. Representative replication factors detected in TRPS1 interactome exhibited a decreased binding to chromatin in the S phase when TRPS1 was removed (Figure 5A), suggesting that an active role for TRPS1 in replisome assembly. However, TRPS1 depletion resulted in an increased loading of Geminin on S-phase chromatin (Figure 5A). Consistently, the chromatin binding of CDC16 and CDC23, two components of the APC/C complex detected in TRPS1 interactome, diminished in TRPS1-depleted cells (Figure 5A). This TRPS1-supported chromatin loading facilitated Geminin targeting by APC/C, as evidenced by the decreased interaction between FZR1 and Geminin upon CDC23 depletion (Figure S5A). These results point to a theme in which TRPS1 recruits the APC/C complex to target Geminin for proteasomal degradation (Figure S5B), thereby promoting pre-RC assembly and origin refiring. Indeed, glutathione S-transferase (GST) pull-down experiments demonstrated that TRPS1 was able to interact with CDC23 through its C-terminal portion containing the IKAROS domain (Figure 5B; Figure S5C). In this regard, it is interesting to note that a recent genetic study found a heterozygous mutation in the region encoding for IKAROS domain of *TRPS1* in a familial TRPS (Kunotai et al., 2017), implying that this module is also important in the pathogenesis of TRPS. Our observation that TRPS1 interacts with CDC23 through its IKAROS domain supports the functional significance of this domain in replication regulation and breast carcinogenesis. Indeed, the ability to promote the proliferation of MCF-10A cells by TRPS1 was significantly impaired when TRPS1ΔIKAROS was expressed (Figure 1G).

Chromatin fractionation analysis performed in MCF-10A cells showed that, although Geminin had no chromatin binding in control cells, reflecting its dynamic association with chromatin or its functional antagonism with CDT1 in nucleoplasm (Xouri et al., 2007), TRPS1 overexpression resulted in the recruitment of Geminin on chromatin, but only in the presence of TAME (Figure 5C; Figure S5D), a pharmacological inhibitor of the ubiquitin ligase activity of APC/C (Zeng et al., 2010). Consistently, TRPS1 overexpression was associated with an increased chromatin binding of CDC23, favoring the argument that TRPS1 recruits APC/C and enables an enzyme-substrate interface between APC/C and Geminin. The functional connection between TRPS1 and APC/C in DNA replication was further strengthened in cells where TAME nullified, at least partially, TRPS1 overexpression-associated origin refiring (Figure 5D). Together, these results link TRPS1 to Geminin degradation, supporting TRPS1-encouraged origin refiring.

Although we were unable to obtain ChIP-seq data for CDC23, possibly due to the inefficacy of the commercially available antibodies for ChIP-seq experiment, we did successfully generate genome-wide profile of CDC16 binding in the presence or absence of TRPS1. CDC16 enriched in TRPS1 peaks across all epigenetic domains (Figure 5E), whereas TRPS1 depletion led to a diminished binding of CDC16 at TRPS1 peaks (Figure S5E), supporting TRPS1-dependent APC/C recruitment for replication regulation.

TRPS1 recruitment on heterochromatic origins is GATA/IKAROS dependent

The identification of IKAROS not GATA domain for CDC23 docking left the defect of TRPS1C896Y and TRPS1R908Q in replication boosting unexplained. A straight-forward explanation is GATA-dependent DNA binding of TRPS1 on E3-type origins, which is perturbed by C896Y or R908Q mutations. qChIP experiments in MCF-7 cells with ectopically expressed TRPS1, TRPS1C896Y, or TRPS1R908Q indeed failed to detect TRPS1C896Y and TRPS1R908Q at all tested TRPS1-bound origins, especially at heterochromatic regions (Figure S5F). However, this DNA binding-dependent reasoning is complicated by motif analysis showing that GATA motif was populated in E5–E7 but not in E1–E3 of TRPS1 peaks (Figure 5F). In this regard, it is important to note that several readers of histone modifications also contain a GATA-like module (Eustermann et al., 2011; Iwase et al., 2011; Vermeulen et al., 2010). We thus employed peptide pull-down

(B) Number of active origins in the windows centered at the same TRPS1 peaks but with different sizes ranging from 0.5 Kb to 1 Mb.
(C) The relative contribution of TRPS1-associated origins stratified by their epigenetic states defined as the observed number of active origins normalized to expected fractions of epigenomic states within TRPS1 summit-centered windows.
(D) Repli-seq in MCF-7 cells transfected with control or TRPS1 siRNAs. Fold changes of average replication efficiency surrounding the TRPS1 peaks was calculated by use of Repli-seq reads density in TRPS1-depleted versus control cells.
(E) Analysis of EdU incorporation and DNA content in MCF-7 (top) or MCF-10A cells (bottom) with the indicated transient gene manipulations.
(F) MCF-7 cells transfected with control or TRPS1 siRNAs were subjected to cell-cycle analysis after isolating a single cell from aggregated ones using doublet discrimination gates. Re-replicated DNA (>4N) contents are presented as gates, and their percentages are plotted as bars.
(G) MCF-7 cells transfected with control or TRPS1 siRNAs were successively pulsed with CldU (red) and IdU (blue).
Bar plot shows the percentage of refiring events relative to the total number of blue tracks. Error bars represent mean ± SD for triplicate experiments in (E)–(G). $^{**}p < 0.01$, $^{***}p < 0.001$, t test.

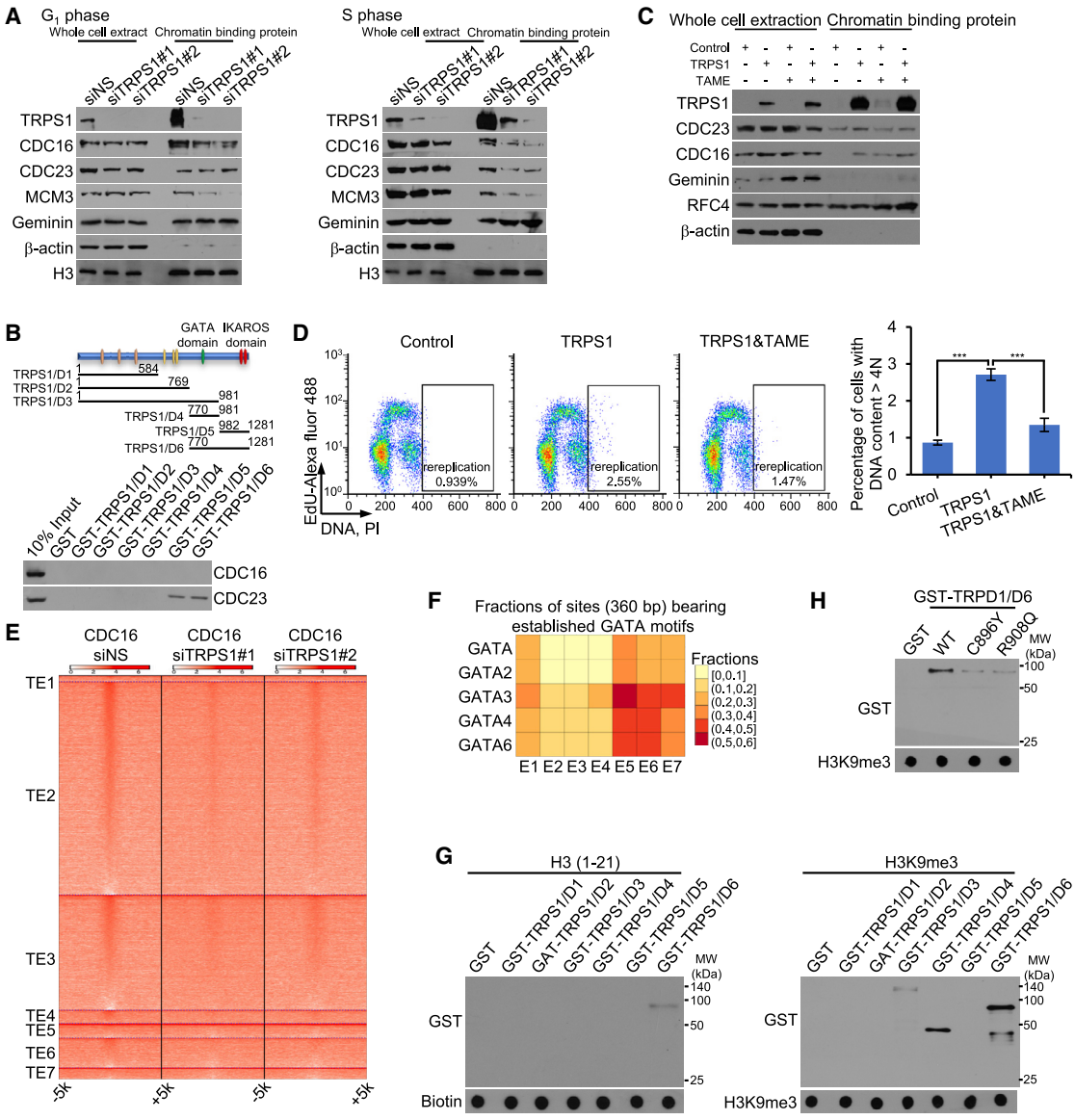


Figure 5. TRPS1 promotes origin refiring and DNA replication through recruitment of the APC/C complex

(A) MCF-7 cells transfected with control or TRPS1 siRNAs were synchronized by a double thymidine block. Chromatin fractionation analysis of TRPS1-associated replication factors was performed in cells collected in G_1 or S phase.

(B) GST pull-down experiments with bacterially expressed GST or the demonstrated GST-fused TRPS1 deletion mutants and *in vitro*-transcribed/translated CDC16 and CDC23 (see also Figure S5C).

(C) Chromatin fractionation analysis of the indicated replication factors in MCF-10A cells stably expressing TRPS1 under the treatment with 100 mM TAME for 12 h to inhibit the enzymatic activity of APC/C.

(D) MCF-10A cells stably expressing FLAG-TRPS1 were untreated or treated with 100 mM TAME for 12 h followed by cell-cycle analysis as in Figure 4F. Re-replicated DNA (>4N) contents were represented as mean ± SD for triplicate experiments. ***$p < 0.001$, one-way ANOVA.

(E) Heatmap analysis of ChIP-seq reads density of CDC16 along TRPS1 peaks stratified by their epigenetic states in MCF-7 cells transfected with two independent siRNAs against TRPS1.

(F) Fraction of TRPS1 peaks stratified by their epigenetic states that possess the indicated GATA motifs.

(G) Peptide pull-down assay was performed using the bacterially expressed GST or GST-fused TRPS1 deletion mutants and biotinylated histone tail peptides with the indicated modifications. Recovered peptides were probed by a dot blotting assay using the indicated antibodies (see also Figure S5C).

(H) H3K9me3 peptide pull-down assay with the indicated GST-fused deletion mutants of TRPS1 (see also Figure S5G).

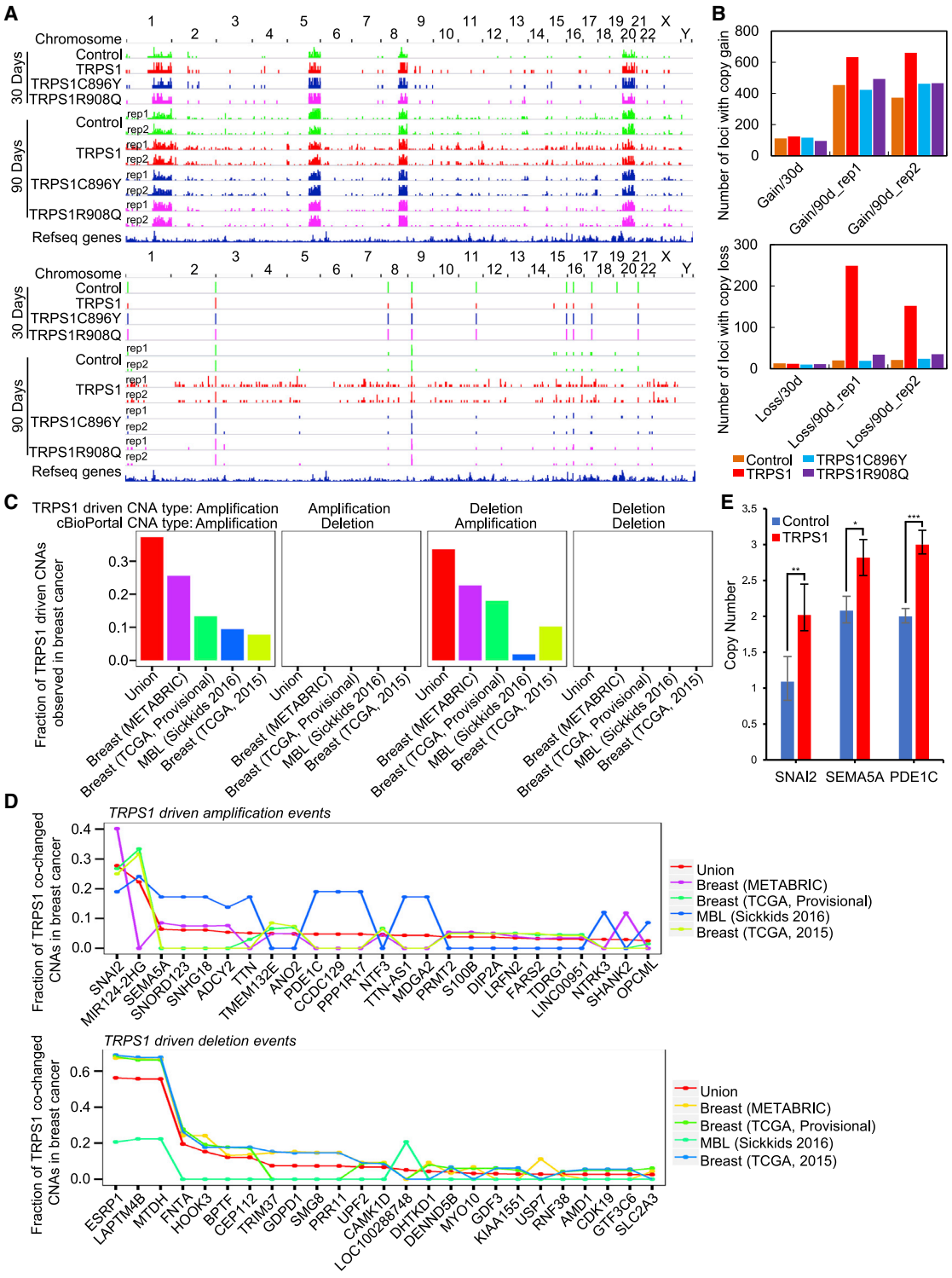


Figure 6. TRPS1 overexpression evokes genome duplications/deletions

(A) TRPS1 overexpression-associated copy-number gain (top) or loss (bottom) analyzed by Affymetrix SNP arrays in MCF-10A cells infected the indicated lentiviruses.

(B) Bar plotting of the number of loci amplified or deleted in MCF-10A cells as in (A).

(C) Fraction of TRPS1-driven CNAs that could be observed in breast cancer samples. Detailed descriptions were provided in the STAR methods. Union represents events that are observed in any of the tested datasets.

(legend continued on next page)

assays with unmodified or K9me3-modified H3 tail to address whether the GATA domain of TRPS1 was similarly involved in reading H3K9me3 instead of recognizing DNA. The GATA domain of TRPS1 exhibited a moderate yet evident affinity toward H3K9me3-modified H3 tail, compared to unmodified H3 peptides (Figure 5G; Figure S5C). Remarkably, further inclusion of the IKAROS domain yielded a strong interaction between the C-terminal portion of TRPS1 and H3K9me3, while the IKAROS domain alone was unable to bind to H3K9me3-modified H3 (Figure 5G). Notably, consistent with the lost binding of TRPS1 mutants to repressive origins, C896Y or R908Q mutations compromised the binding of TRPS1 to H3K9me3-modified histone peptides (Figure 5H; Figure S5G). These findings support a notion that the binding of TRPS1 on heterochromatic origins is fostered by an H3K9me3-marked chromatin environment in a GATA/IKAROS-dependent manner and that mutations in these domains compromise the physical and functional interaction of TRPS1 with the replication machinery.

TRPS1 overexpression provokes genome duplication that mimics genetic aberrances in breast cancer

The intimate link between dysregulated replication and genome aberrances (Kuzminov, 2016) promotes us to investigate whether overexpression of TRPS1, a frequent event occurs in breast cancer, as described earlier, could drive genome amplification. To this end, stable MCF-10A cell lines integrated with TRPS1, TRPS1C896Y, or TRPS1R908Q were established and further cultured for 30 or 90 days before copy-number alteration (CNA) analysis using Affymetrix SNP arrays. After subtracting copy gains in similarly maintained control cells, we found genome amplifications and deletions scattered in different chromosomes in cells overexpressing TRPS1 (Figure 6A). Strikingly, compared to control cells, no dramatic difference of genetic alterations was observed in cells overexpressing TRPS1C896Y or TRPS1R908Q (Figure 6A; Figures S6A and S6B).

We next examined oncogenes or tumor suppressors overlapped with TRPS1-inflicted CNAs that might also be driven by TRPS1-associated DNA re-replication/genomic aberrances in breast cancers. We assembled a set of recurrent CNAs that were associated with TRPS1 overexpression and selected ones that were found in both replicates of TRPS1-overexpressing cells but absent from any CNA lists called from control or TRPS1C896Y- or TRPS1R908Q-overexpressing cells. This yielded a total of 123 amplified and 108 deleted regions (Figure 6B). We retrieved genes co-amplified with TRPS1 within 4 large cohorts of breast cancer samples (Ciriello et al., 2015; Morrissy et al., 2016; Pereira et al., 2016; Cancer Genome Atlas Network, 2012). Mapping significantly co-amplified genes with a co-occurrence FDR less than 0.05 to TPRS1-associated CNA loci showed that a total of ~35% of regions with gained CNAs were also amplified in at least one of the 4 cohorts (Figure 6C; Figure S6C). Intriguingly, however, copy loss associated with TRPS1 overexpression was reflected by a similar proportion of amplifications not deletions in the clinical samples (Figure 6C; Figure S6C). Functional analysis revealed that the top 25 co-amplified genes are eminently implicated in epithelial-mesenchymal transition (EMT), cell migration or angiogenesis (Figure 6D; Figure S6D), and amplification of the selected genes was confirmed by TaqMan copy-number assay (Figure 6E). These observations imply that cancer cells could harness overexpressed TRPS1 to elicit genome duplications and deletions along with TPRS1-evoked origin refiring.

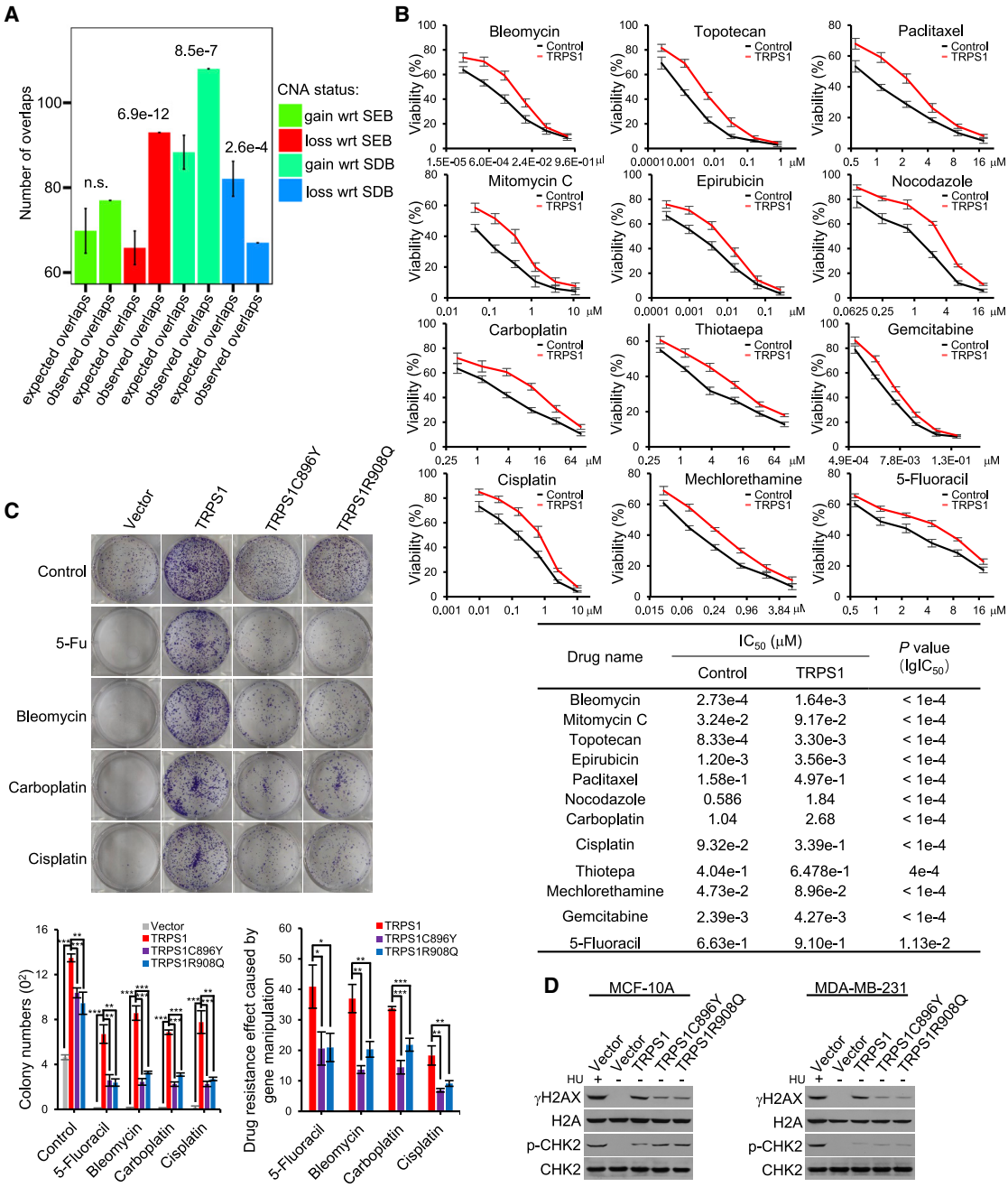
Genome deletions associated with TRPS1 overexpression are found in eccDNAs

The coincidence of copy loss associated with TRPS1 overexpression with copy gain in clinical samples is puzzling. It is possible that the deleted genomic segments under TRPS1 overexpression went to an extrachromosomal route (Sanborn et al., 2013). To test this, we purified non-chromosomal DNA from MCF-10A nuclei using ATP-dependent exonuclease to remove linear genomic DNAs (Shibata et al., 2012). The resultant exonuclease-resistant DNAs were subjected to high-throughput sequencing. Due to compositional heterogeneity of extrachromosomes (Turner et al., 2017), we did not expect to detect full-sized extrachromosomes from the short sequencing reads. However, it is still possible to gauge the statistical significance of the enrichment of relative short extrachromosomal fragments in TRPS1-mediated CNA regions. We thus tiled the genome into 10-Kb windows and called the significantly enriched bins (SEBs) in TRPS1-overexpressing versus control cells as the ones with reads density ratio greater than 2 and p value less than 1e-5 with Poisson test. The enrichment of TRPS1-associated CNAs in extrachromosomes was then evaluated by comparing the number of CNA sites overlapped with at least one SEB to overlaps in null models consisting of 100 groups of randomly shuffled genomic sites with similar length distributions. We found that TRPS1 overexpression-associated genomic deletions were significantly enriched in extrachromosomal DNAs (Figure 7A, $p = 6.9e-12$). A similar test on overlaps with significantly depleted bins (SDBs) further consolidated the depletion of TRPS1 overexpression-associated genomic deletions from genomic DNAs (Figure 7A, $p = 2.6e-4$). Together, these data support an extrachromosomal route for TRPS1-driven genome deletions.

TRPS1-evoked genome aberrance dynamically evolves to confer therapeutic resistance

The occurrence of widespread genome amplification associated with TRPS1 overexpression suggests an evolving selection theme in TRPS1-promoted breast carcinogenesis. Indeed, TRPS1-overexpressing MCF-10A cells exhibited chemoresistance to a broad spectrum of therapeutic compounds (Figure 7B). To validate the link of the chemoresistance to TRPS1-evoked genomic aberrance, we sequenced the whole genome from

(D) Fraction of the alteration in cancer samples of the indicated genes identified from TRPS1 overexpression-associated amplification (top) or deletion (bottom) that also significantly co-varied with TRPS1 CNAs in breast cancer samples.
(E) Copy number of the indicated genes in MCF-10A cells stably expressing empty vector or TRPS1 was measured by TaqMan copy-number assay. Error bars represent mean ± SD for triplicate experiments. *$p < 0.05$; **$p < 0.01$, ***$p < 0.001$, t test.



Drug name	IC_{50} (μM)		*P* value ($lgIC_{50}$)
	Control	TRPS1	
Bleomycin	2.73e-4	1.64e-3	< 1e-4
Mitomycin C	3.24e-2	9.17e-2	< 1e-4
Topotecan	8.33e-4	3.30e-3	< 1e-4
Epirubicin	1.20e-3	3.56e-3	< 1e-4
Paclitaxel	1.58e-1	4.97e-1	< 1e-4
Nocodazole	0.586	1.84	< 1e-4
Carboplatin	1.04	2.68	< 1e-4
Cisplatin	9.32e-2	3.39e-1	< 1e-4
Thiotepa	4.04e-1	6.478e-1	4e-4
Mechlorethamine	4.73e-2	8.96e-2	< 1e-4
Gemcitabine	2.39e-3	4.27e-3	< 1e-4
5-Fluoracil	6.63e-1	9.10e-1	1.13e-2

Figure 7. TRPS1-evoked genome aberrances dynamically evolve to confer therapeutic resistance in breast cancer cells

(A) The enrichment of TRPS1-associated CNAs in extrachromosomes was analyzed by comparing the number of CNA sites overlapped with at least one SEB or SDB to overlap in null models consisting of 100 groups of randomly shuffled genomic sites with a similar length distribution as TRPS1-associated CNAs. n.s., not significant.

(B) TRPS1 stably expressing MCF-10A cells were treated with the indicated antineoplastic compounds for 3 days for cell-proliferation measurement by MTS assays. Error bars represent mean ± SD for triplicate experiments. IC_{50} (half maximal inhibitory concentration) was listed (bottom). p values of $lgIC_{50}$ were calculated based on extra sum-of-squares *F*-test.

(C) MCF-10A cells transduced with the indicated constructs were maintained in medium containing the indicated chemotherapeutic agents for two weeks. Representative images of colony formation were shown (top). The number of colonies was indicated (bottom left). The effect of drug resistance caused by one gene manipulation was calculated as the fraction of colony numbers in drug- versus vehicle-treated cells under this manipulation divided by the fraction under control transduction (bottom right). Error bars represent mean ± SD for triplicate experiments. *p < 0.05, **p < 0.01, ***p < 0.001, t test.

(D) MCF-10A and MDA-MB-231 cells transduced with the indicated constructs were maintained with 0.339 μM (MCF-10A) or 3.515 μM (MDA-MB-231, Figure S7D) cisplatin for 2 weeks. Western blotting was then performed to show the levels of indicated proteins or modifications. Control lentiviruses transduced cells (vector) in the presence Hydroxyurea (HU) were included as a positive control for DNA damage response and checkpoint activation upon genotoxic stress.

TRPS1-overexpressing/chemoresistant MCF-10A cells upon treatment with cisplatin, carboplatin, 5-fluorouracil, gemcitabine, bleomycin, or mitomycin C. We found a large repository of genomic aberrances that were associated with TRPS1 overexpression (Figure S7A). The results agree with the observations in SNP arrays with a significantly improved sensitivity. Compared to untreated cells, treatment of TRPS1-overexpressing with the above 6 compounds yielded common as well as uniquely evolved sets of amplified CNAs (Figure S7B), suggesting an adaptation to different selection pressures. Consistently, functional annotation of CNAs-association genes in cells treated with distinct compounds found, in addition to the commonly identified resistant genes, loci that are known to be involved in resistance to the specific chemicals (Figure S7C). Moreover, tumorigenic analysis indicated that, while WT TRPS1 promoted the colony growth and resistance to genotoxic stress and suppressed DNA damage checkpoint activation in both MCF-10A and MDA-MB-231 cells after culturing and selection for 2 weeks, the two TRPS-related mutants showed compromised effects (Figures 7C and 7D; Figure S7D). Apparently, TRPS1 overexpression-evoked genetic abnormalities enable fitted cancerous cells to clonally outgrow while their genomes are continuing to evolve.

DISCUSSION

Although TRPS is characterized by developmental defects in hair follicle and craniofacial/skeletal structures (Giedion et al., 1973), its pathological influence also involves a broad range of tissues/organs (Fernández et al., 1993; Hansen and Shewmake, 1979; Tasic et al., 2014), suggesting that the molecular activity of TRPS1 might be fundamental. In addition, TRPS1 has been implicated in a number of malignancies including breast cancer. However, possibly due to the possession of a DNA-binding GATA domain, the majority of mechanistic studies regarding TRPS1 have focused on transcription regulation, yielding limited information to neither explain the phenotype of TRPS nor envision TRPS1 as a breast cancer driver. In the current study, we found that TRPS1 is associated with the replication machinery and promotes DNA replication, a fundamental cellular process that is intrinsically linked to genome stability and malignant transformation. We found that TRPS1 overexpression is associated with genome-wide aberrances and CNAs. Since genomic instability/genome aberrance could enable cells to evolve with a possibility of oncogene amplification and malignant transformation (Curtis et al., 2012), TRPS1-encouraged DNA replication and its provoked genetic alterations favor the argument that TRPS1 is a breast cancer driver. We propose that the oncogenic potential of TRPS1 stems from its engagement with H3K9me3-marked heterochromatic replication origins and its enforced replication of repressive chromatin at an early stage of DNA duplication. We showed that TRPS1 does so, through recruitment of the APC/C complex, a molecular machinery functioning to destabilize Geminin, leading to uncontrolled refiring of replication origins. Geminin is the key node in controlling the balance between replication licensing and re-replication (Fragkos et al., 2015). Over-fired origins and stacked re-replicated DNA fragments could cause rear-ending of refired elongating replisomes into previously activated replication fork, leading to replication stalled, chromosome destabilized, and ultimately to DNA double breaks and chromosomal fragmentation (Alexander and Orr-Weaver, 2016).

Several studies implicate TRPS1 in cell proliferation and survival (Gong et al., 2018; Wang et al., 2018a; Witwicki et al., 2018), and we also showed that TRPS1 overexpression promotes cell proliferation in our current study. However, replication origin occupancy and DNA re-replication promotion appear to be decoupled from the reported transcriptional regulation by TRPS1. In addition, drug-associated differential genetic alterations suggest an adaptation to dynamic selection pressures imposed by different drugs that could be better explained by TRPS1-driven genomic instability and cancer genome evolution. Thus, although the possibility of encouraging proliferation and survival could not be excluded, our results point to replication boosting and genome instability as the tumorigenic potential of TRPS1. In support of this, several recent studies indicate that factors well established in control of cell proliferation also play roles in the maintenance of genome integrity. For example, it is reported that overexpression of histone demethylase KDM4A inflicts focalized copy gains dependent of its capacity to recruit replication factors (Black et al., 2013). Moreover, it should be emphasized that it becomes increasingly clear that gene transcription and DNA replication are intimately coupled and that these two processes interplay and are cross-regulated (Fragkos et al., 2015); it is also possible that TRPS1-mediated transcriptional regulation for cell proliferation and survival and TRPS1-promoted replication for genome duplication are co-opted within local chromosomal contexts and eventually synergized to confer functional optimality for cell growth and evolution. It is also worth noting that the vast majority of screenings for replication regulators were based on microorganisms or experimental models such as yeast and fly (Burgers and Kunkel, 2017); these organisms lack a TRPS1 ortholog. Furthermore, unlike ubiquitous expression patterns of the general replication factors such as MCM, TRPS1 exhibits a tissue-specific expression pattern, with high expression in mammary tissues. These might contribute to the missed identification of TRPS1 from genetic screening for replication regulators in yeast and flies, or from large proteomic studies searching for interactomes of known replication factors in cell lines such as HeLa, HEK293T, or U2OS where TRPS1 expression is low (Boos et al., 2013; Dungrawala et al., 2015; Tom et al., 1996). Indeed, co-immunoprecipitation experiments showed that endogenous TRPS1 was readily detected as a MCM2 interacting partner in MCF-7 cells but was barely detected in U2OS, HEK293T, and HeLa cells, although it is quite possible that additional tissue-specific factor(s) exists to bridge the interaction of TRPS1 with the replication regulators in mammary tissues.

While the condensed heterochromatin is generally duplicated in late S phase with postponed origin firing (Rhind and Gilbert, 2013), there are evident examples of early heterochromatin replication across species, including the early replication of HP1-bound pericentromeric regions and silent mating-type locus in *Schizosaccharomyces pombe* (Hayashi et al., 2009), early duplication of centromeres as isolated domains in *Drosophila* cells (Ahmad and Henikoff, 2001), early megareplicons of pericentric

heterochromatin and centromere in mouse (Holló et al., 1996), and arm-specific and subtelomeric structure-influenced variation of replication timing of human telomeres (Arnoult et al., 2010). These data suggest that the association between heterochromatin and late replication is common but not universal, despite the condensed and peripheral nuclear positioning nature of heterochromatin (Kim et al., 2003). In this study, we found that TRPS1-boosted replication of repressive domains takes place at the early stage of the S phase. Although the biological significance of replication timing remains to be understood (Rhind and Gilbert, 2013), we could envision that TRPS1-assisted early replication regions might be assembled into nucleation sites enriched with repressive marks to facilitate epigenomic organization essential for the subsequent DNA duplication and cell mitosis (Aladjem, 2007). Alternatively, a subset of G1 phase silenced genes, especially those that are required for replication itself, needs to be quickly replicated and turned to the active state in early S phase to feed replication thereafter. It is also worth noting that origin-containing relative to origin-free E3-type TRPS1 binding sites are more enriched with H3K4me1 and H3K36me3 marks, suggesting a tendency of TRPS1 binding sites associated with poised weak enhancers or intragenic regions of lowly expressed genes to be origins within H3K9me3 marked domains. These epigenetic features together with the physical interaction between TRPS1 and replication factors support TRPS1's participation in early replication of heterochromatin.

Although GATA domain is characterized as a DNA binding module (Merika and Orkin, 1993), we did not identify any recognizable cognate GATA motif in repressive domains. We did not find any consensus DNA sequence that might account for the TRPS1 binding peaks adjacent to H3K9me3-marked origins either. This issue is also confounded by lack of definable genetic elements for DNA replication origins in mammalian cells (Hyrien, 2015). Inspired by direct reading of H3K9me3 by a GATA-type zinc finger protein/ATP-dependent helicase ATRX (Eustermann et al., 2011; Iwase et al., 2011), we found that the GATA domain of TRPS1 had moderate yet specific affinity to H3K9me3-marked peptides, and, surprisingly, inclusion of the IKAROS domain enhanced the affinity. On the other hand, the modular organization of TRPS1 is amenable to associate with multiple repressive complexes such as NuRD and CtBP, possibly through distinct intra-modular or inter-modular three-dimensional architecture of TRPS1 (Koipally and Georgopoulos, 2002). In effect, NuRD/CtBP might assist TRPS1 targeting at the heterochromatic regions through direct protein-protein interactions and via their chromatin modifying capabilities. Our results support a model in which CtBP/NuRD contribute to the establishment of an H3K9me3-enriched heterochromatic environment for GATA/IKAROS-dependent H3K9me3 reading of heterochromatic origins by TRPS1.

Interestingly, the well-recognized TRPS-associated TRPS1 mutations C896Y or R908Q in the GATA domain reserve the ability to interact with CtBP/NuRD but lost the capacity to promote replication and tumorigenesis. Although the possibility of direct *cis* binding by the GATA domain contributes to TRPS1-promoted carcinogenesis could not be excluded, TRPS1-evoked replication boosting and genome-wide structural alterations disfavor a transcriptional regulatory model in the pathogenesis of TRPS. It is not unlikely that the molecular etiologies of TRPS and breast cancer converge on dysregulated DNA replication (under-replication in TRPS versus over-replication in breast cancer); compromised replication efficiency could entail declined self-renewal or/and premature senescence of embryonic or mesenchymal stem cells, ultimately contributing to developmental defects.

Identification of oncogenic drivers is important to understanding the essence of malignant transformation and to combatting cancer. This effort has been compounded by cancer genome evolution, which gives rise to intra-tumoral heterogeneity and extrachromosomal. Although it is not possible to rebuild the whole recurrent eccDNAs due to sequencing reads length limit, our statistical inference implies significantly higher content of short fragments of eccDNAs in TRPS1-associated deletions. These eccDNAs could be equipped with a chromosomal-independent replication capacity. A direct cancer-causing effect from eccDNAs could be their repeated over-replication and unchecked segregation to confer rapid genetic adaptability for cancer cell survival against metabolic, immunosurveilling, and drug-elicited stresses (Turner et al., 2017). Indeed, we found extraordinary drug resistance in TRPS1-overexpressing cells, supporting the pursuit of TRPS1 as a potential target for breast cancer intervention.

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
  - Mouse models
  - Cell lines
- METHOD DETAILS
  - Cell transfection
  - Cell proliferation assay
  - Immunohistochemistry assay
  - Senescence detection assay
  - Apoptosis detection assay
  - Immunofluorescence assay
  - Cycloheximide chase assay
  - *In vivo* metastasis
  - Immunopurification and mass spectrometry
  - Cell cycle synchronization
  - Fast protein liquid chromatography
  - Co-immunoprecipitation and western blotting
  - EdU incorporation assay and flow cytometry
  - DNA fiber assay
  - Cesium chloride gradient centrifugation
  - Chromatin fractionation
  - GST pull-down assay
  - Peptide pull-down assay
  - Chromatin immunoprecipitation

- ChIP-seq
- Replication timing analysis
- DNA structural variation analysis
- Extrachromosomal DNA analysis
- Drug sensitivity test
- Whole genome sequencing and CNA analysis
- Colony formation assay

- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2021.108814.

ACKNOWLEDGMENTS

This work was supported by grants (31991164, 81730079, and 81530073 to Y.S. and 32070638 to Y.Z.) from the National Natural Science Foundation of China and a grant (2016YFC1302304 to Y.S.) from the Ministry of Science and Technology of the People's Republic of China.

AUTHOR CONTRIBUTIONS

J.Y., Y.Z., and Y.S. conceived the project. Y.Z. and J.Y. designed experiments, and J.Y., X.L., Y.H., L.H., W.Z., J.R., Y.W., J.W., X.W., L.S., X.Y., L.S., and J.L. performed the experiments. J.Y., Y.Z., and Y.S. analyzed the data and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 19, 2020
Revised: December 18, 2020
Accepted: February 10, 2021
Published: March 9, 2021

SUPPORTING CITATIONS

The following references appear in the supplemental information: Arumugam et al. (2009); Chekhun et al. (2005); Dabholkar et al. (1992); Dangi-Garimella et al. (2011); Fischer et al. (2017); Hagmann et al. (2009); Hilton et al. (2002); Huang et al. (2014); Hufeland et al. (2015); Izumi et al. (2010); Kamar et al. (2017); Kantaputra et al. (2008); Kasahara et al. (1991); Kobayashi et al. (2002), (2005); Liang et al. (2008); Oguri et al. (2007); Onozawa et al. (2017); Owatari et al. (2007); Park et al. (2010); Pfleger et al. (2001); Piccione et al. (2009); Rossi et al. (2007); Shibata et al. (2015); Sidler et al. (2012); Singh et al. (2010); Sliutz et al. (1996); Smith (1999); Stewart (2007); Ullah et al. (2018); Van Den Broeck et al. (2012); (Wang et al., 2018b; Wang et al., 2017a, 2017b); Wood et al. (2011); Ye et al. (2017); Zhuo et al. (2008).

REFERENCES

Abyzov, A., Urban, A.E., Snyder, M., and Gerstein, M. (2011). CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. Genome Res. *21*, 974–984.

Ahmad, K., and Henikoff, S. (2001). Modulation of a transcription factor counteracts heterochromatic gene silencing in Drosophila. Cell *104*, 839–847.

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

Alexander, J.L., and Orr-Weaver, T.L. (2016). Replication fork instability and the consequences of fork collisions from rereplication. Genes Dev. *30*, 2241–2252.

Arnoult, N., Schluth-Bolard, C., Letessier, A., Drascovic, I., Bouarich-Bourimi, R., Campisi, J., Kim, S.H., Boussouar, A., Ottaviani, A., Magdinier, F., et al. (2010). Replication timing of human telomeres is chromosome arm-specific, influenced by subtelomeric structures and connected to nuclear localization. PLoS Genet. *6*, e1000920.

Arumugam, T., Ramachandran, V., Fournier, K.F., Wang, H., Marquis, L., Abbruzzese, J.L., Gallick, G.E., Logsdon, C.D., McConkey, D.J., and Choi, W. (2009). Epithelial to mesenchymal transition contributes to drug resistance in pancreatic cancer. Cancer Res. *69*, 5820–5828.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. B *57*, 289–300.

Black, J.C., Manning, A.L., Van Rechem, C., Kim, J., Ladd, B., Cho, J., Pineda, C.M., Murphy, N., Daniels, D.L., Montagna, C., et al. (2013). KDM4A lysine demethylase induces site-specific copy gain and rereplication of regions amplified in tumors. Cell *154*, 541–555.

Boos, D., Yekezare, M., and Diffley, J.F. (2013). Identification of a heteromeric complex that promotes DNA replication origin firing in human cells. Science *340*, 981–984.

Burgers, P.M.J., and Kunkel, T.A. (2017). Eukaryotic DNA Replication Fork. Annu. Rev. Biochem. *86*, 417–438.

Cancer Genome Atlas Network (2012). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61–70.

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.

Chang, G.T., Steenbeek, M., Schippers, E., Blok, L.J., van Weerden, W.M., van Alewijk, D.C., Eussen, B.H., van Steenbrugge, G.J., and Brinkmann, A.O. (2000). Characterization of a zinc-finger protein and its association with apoptosis in prostate cancer cells. J. Natl. Cancer Inst. *92*, 1414–1421.

Chekhun, V.F., Lukyanova, N.Y., Urchenko, O.V., and Kulik, G.I. (2005). The role of expression of the components of proteome in the formation of molecular profile of human ovarian carcinoma A2780 cells sensitive and resistant to cisplatin. Exp. Oncol. *27*, 191–195.

Ciriello, G., Gatza, M.L., Beck, A.H., Wilkerson, M.D., Rhie, S.K., Pastore, A., Zhang, H., McLellan, M., Yau, C., Kandoth, C., et al.; TCGA Research Network (2015). Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. Cell *163*, 506–519.

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Cornelissen, L.M., Drenth, A.P., van der Burg, E., de Bruijn, R., Pritchard, C.E.J., Huijbers, I.J., Zwart, W., and Jonkers, J. (2020). TRPS1 acts as a context-dependent regulator of mammary epithelial cell growth/differentiation and breast cancer development. Genes Dev. *34*, 179–193.

Curtis, C., Shah, S.P., Chin, S.F., Turashvili, G., Rueda, O.M., Dunning, M.J., Speed, D., Lynch, A.G., Samarajiwa, S., Yuan, Y., et al.; METABRIC Group (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature *486*, 346–352.

Dabholkar, M., Bostick-Bruton, F., Weber, C., Bohr, V.A., Egwuagu, C., and Reed, E. (1992). ERCC1 and ERCC2 expression in malignant tissues from ovarian cancer patients. J. Natl. Cancer Inst. *84*, 1512–1517.

Dangi-Garimella, S., Krantz, S.B., Barron, M.R., Shields, M.A., Heiferman, M.J., Grippo, P.J., Bentrem, D.J., and Munshi, H.G. (2011). Three-dimensional collagen I promotes gemcitabine resistance in pancreatic cancer through MT1-MMP-mediated expression of HMGA2. Cancer Res. *71*, 1019–1028.

Dungrawala, H., Rose, K.L., Bhat, K.P., Mohni, K.N., Glick, G.G., Couch, F.B., and Cortez, D. (2015). The Replication Checkpoint Prevents Two Types of Fork Collapse without Regulating Replisome Stability. Mol. Cell *59*, 998–1010.

Elmore, L.W., Rehder, C.W., Di, X., McChesney, P.A., Jackson-Cook, C.K., Gewirtz, D.A., and Holt, S.E. (2002). Adriamycin-induced senescence in breast tumor cells involves functional p53 and telomere dysfunction. J. Biol. Chem. *277*, 35509–35515.

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. Nat. Methods *9*, 215–216.

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

Fernández, J.M., Santolaya, J.M., Sádaba, F., Areitio, E., Alvaro, L.C., and Cortina, C. (1993). Basilar impression and syringomyelia in a patient with tricho-rhino-phalangeal syndrome type I. Clin. Genet. *43*, 324–325.

Fischer, S.B., Attenhofer, M., Gultekin, S.H., Ross, D.A., and Heinimann, K. (2017). TRPS1 gene alterations in human subependymoma. J. Neurooncol. *134*, 133–138.

Fragkos, M., Ganier, O., Coulombe, P., and Méchali, M. (2015). DNA replication origin activation in space and time. Nat. Rev. Mol. Cell Biol. *16*, 360–374.

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.

Giedion, A., Burdea, M., Fruchter, Z., Meloni, T., and Trosc, V. (1973). Autosomal-dominant transmission of the tricho-rhino-phalangeal syndrome. Report of 4 unrelated families, review of 60 cases. Helv. Paediatr. Acta *28*, 249–259.

Gong, X., Liu, W., Wu, L., Ma, Z., Wang, Y., Yu, S., Zhang, J., Xie, H., Wei, G., Ma, F., et al. (2018). Transcriptional repressor GATA binding 1-mediated repression of SRY-box 2 expression suppresses cancer stem cell functions and tumor initiation. J. Biol. Chem. *293*, 18646–18654.

Hagmann, W., Jesnowski, R., Faissner, R., Guo, C., and Löhr, J.M. (2009). ATP-binding cassette C transporters in human pancreatic carcinoma cell lines. Upregulation in 5-fluorouracil-resistant cells. Pancreatology *9*, 136–144.

Hansen, D.D., and Shewmake, S.W. (1979). Tricho-rhino-phalangeal syndrome. Int. J. Dermatol. *18*, 561–564.

Hansen, R.S., Thomas, S., Sandstrom, R., Canfield, T.K., Thurman, R.E., Weaver, M., Dorschner, M.O., Gartler, S.M., and Stamatoyannopoulos, J.A. (2010). Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. Proc. Natl. Acad. Sci. USA *107*, 139–144.

Hayashi, M.T., Takahashi, T.S., Nakagawa, T., Nakayama, J., and Masukata, H. (2009). The heterochromatin protein Swi6/HP1 activates replication origins at the pericentromeric region and silent mating-type locus. Nat. Cell Biol. *11*, 357–362.

He, L., Liu, X., Yang, J., Li, W., Liu, S., Liu, X., Yang, Z., Ren, J., Wang, Y., Shan, L., et al. (2018). Imbalance of the reciprocally inhibitory loop between the ubiquitin-specific protease USP43 and EGFR/PI3K/AKT drives breast carcinogenesis. Cell Res. *28*, 934–951.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Hilton, M.J., Sawyer, J.M., Gutiérrez, L., Hogart, A., Kung, T.C., and Wells, D.E. (2002). Analysis of novel and recurrent mutations responsible for the tricho-rhino-phalangeal syndromes. J. Hum. Genet. *47*, 103–106.

Holló, G., Keresö, J., Praznovszky, T., Cserpán, I., Fodor, K., Katona, R., Csonka, E., Fátyol, K., Szeles, A., Szalay, A.A., and Hadlaczky, G. (1996). Evidence for a megareplicon covering megabases of centromeric chromosome segments. Chromosome Res. *4*, 240–247.

Huang, Y., Yu, H., Lei, H., Xie, C., and Zhong, Y. (2014). Matrix metalloproteinase 7 is a useful marker for 5-fluorouracil-based adjuvant chemotherapy in stage II and stage III colorectal cancer patients. Med. Oncol. *31*, 824.

Hufeland, M., Rahner, N., and Krauspe, R. (2015). Trichorhinophalangeal syndrome type I: a novel mutation and Perthes-like changes of the hip in a family with 4 cases over 3 generations. J. Pediatr. Orthop. *35*, e1–e5.

Hyrien, O. (2015). Peaks cloaked in the mist: the landscape of mammalian replication origins. J. Cell Biol. *208*, 147–160.

Itoh, S., Kanno, S., Gai, Z., Suemoto, H., Kawakatsu, M., Tanishima, H., Morimoto, Y., Nishioka, K., Hatamura, I., Yoshida, M., and Muragaki, Y. (2008). Trps1 plays a pivotal role downstream of Gdf5 signaling in promoting chondrogenesis and apoptosis of ATDC5 cells. Genes Cells *13*, 355–363.

Iwase, S., Xiang, B., Ghosh, S., Ren, T., Lewis, P.W., Cochrane, J.C., Allis, C.D., Picketts, D.J., Patel, D.J., Li, H., and Shi, Y. (2011). ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. Nat. Struct. Mol. Biol. *18*, 769–776.

Izumi, K., Takagi, M., Parikh, A.S., Hahn, A., Miskovsky, S.N., Nishimura, G., Torii, C., Kosaki, K., Hasegawa, T., and Neilson, D.E. (2010). Late manifestations of tricho-rhino-pharangeal syndrome in a patient: Expanded skeletal phenotype in adulthood. Am. J. Med. Genet. A. *152A*, 2115–2119.

Jenkins, D.E., Hornig, Y.S., Oei, Y., Dusich, J., and Purchio, T. (2005). Bioluminescent human breast cancer cell lines that permit rapid and sensitive in vivo detection of mammary tumors and multiple metastases in immune deficient mice. Breast Cancer Res. *7*, R444–R454.

Kamar, A., Fahed, A.C., Shibbani, K., El-Hachem, N., Bou-Slaiman, S., Arabi, M., Kurban, M., Seidman, J.G., Seidman, C.E., Haidar, R., et al. (2017). A Novel Role for *CSRP1* in a Lebanese Family with Congenital Cardiac Defects. Front. Genet. *8*, 217.

Kantaputra, P., Miletich, I., Lüdecke, H.J., Suzuki, E.Y., Praphanphoj, V., Shivdasani, R., Wuelling, M., Vortkamp, A., Napierala, D., and Sharpe, P.T. (2008). Tricho-rhino-phalangeal syndrome with supernumerary teeth. J. Dent. Res. *87*, 1027–1031.

Kasahara, K., Fujiwara, Y., Nishio, K., Ohmori, T., Sugimoto, Y., Komiya, K., Matsuda, T., and Saijo, N. (1991). Metallothionein content correlates with the sensitivity of human small cell lung cancer cell lines to cisplatin. Cancer Res. *51*, 3237–3242.

Kim, S.M., Dubey, D.D., and Huberman, J.A. (2003). Early-replicating heterochromatin. Genes Dev. *17*, 330–335.

Kobayashi, H., Hino, M., Shimodahira, M., Iwakura, T., Ishihara, T., Ikekubo, K., Ogawa, Y., Nakao, K., and Kurahachi, H. (2002). Missense mutation of TRPS1 in a family of tricho-rhino-phalangeal syndrome type III. Am. J. Med. Genet. *107*, 26–29.

Kobayashi, H., Hino, M., Inoue, T., Nii, E., Ikeda, K., Son, C., Iwakura, T., Ishihara, T., and Ogawa, Y. (2005). GC79/TRPS1 and tumorigenesis in humans. Am. J. Med. Genet. A. *134*, 341–343.

Koipally, J., and Georgopoulos, K. (2002). Ikaros-CtIP interactions do not require C-terminal binding protein and participate in a deacetylase-independent mode of repression. J. Biol. Chem. *277*, 23143–23149.

Kunotai, W., Ananpornruedee, P., Lubinsky, M., Pruksametanan, A., and Kantaputra, P.N. (2017). Making extra teeth: Lessons from a TRPS1 mutation. Am. J. Med. Genet. A. *173*, 99–107.

Kuzminov, A. (2016). Chromosomal Replication Complexity: A Novel DNA Metrics and Genome Instability Factor. PLoS Genet. *12*, e1006229.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput. Biol. *9*, e1003118.

Lefebvre, C., Bachelot, T., Filleron, T., Pedrero, M., Campone, M., Soria, J.C., Massard, C., Lévy, C., Arnedos, M., Lacroix-Triki, M., et al. (2016). Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. PLoS Med. *13*, e1002201.

Li, Z., Jia, M., Wu, X., Cui, J., Pan, A., and Li, L. (2015). Overexpression of Trps1 contributes to tumor angiogenesis and poor prognosis of human osteosarcoma. Diagn. Pathol. *10*, 167.

Liang, X.J., Finkel, T., Shen, D.W., Yin, J.J., Aszalos, A., and Gottesman, M.M. (2008). SIRT1 contributes in part to cisplatin resistance in cancer cells by altering mitochondrial metabolism. Mol. Cancer Res. *6*, 1499–1506.

Lüdecke, H.J., Schaper, J., Meinecke, P., Momeni, P., Gross, S., Hirche, H., Abramowicz, M.J., Albrecht, B., Apacik, C., Christen, H.J., et al. (2001). Genotypic and phenotypic spectrum in tricho-rhino-phalangeal syndrome types I and III. Am. J. Hum. Genet. *68*, 81–91.

Malik, T.H., Von Stechow, D., Bronson, R.T., and Shivdasani, R.A. (2002). Deletion of the GATA domain of TRPS1 causes an absence of facial hair and provides new insights into the bone disorder in inherited tricho-rhino-phalangeal syndromes. Mol. Cell. Biol. *22*, 8592–8600.

Martin, M.M., Ryan, M., Kim, R., Zakas, A.L., Fu, H., Lin, C.M., Reinhold, W.C., Davis, S.R., Bilke, S., Liu, H., et al. (2011). Genome-wide depletion of replication initiation events in highly transcribed regions. Genome Res. *21*, 1822–1832.

Merika, M., and Orkin, S.H. (1993). DNA-binding specificity of GATA family transcription factors. Mol. Cell. Biol. *13*, 3999–4010.

Momeni, P., Glöckner, G., Schmidt, O., von Holtum, D., Albrecht, B., Gillessen-Kaesbach, G., Hennekam, R., Meinecke, P., Zabel, B., Rosenthal, A., et al. (2000). Mutations in a new gene, encoding a zinc-finger protein, cause tricho-rhino-phalangeal syndrome type I. Nat. Genet. *24*, 71–74.

Montagnoli, A., Valsasina, B., Brotherton, D., Troiani, S., Rainoldi, S., Tenca, P., Molinari, A., and Santocanale, C. (2006). Identification of Mcm2 phosphorylation sites by S-phase-regulating kinases. J. Biol. Chem. *281*, 10281–10290.

Morrissy, A.S., Garzia, L., Shih, D.J., Zuyderduyn, S., Huang, X., Skowron, P., Remke, M., Cavalli, F.M., Ramaswamy, V., Lindsay, P.E., et al. (2016). Divergent clonal selection dominates medulloblastoma at recurrence. Nature *529*, 351–357.

Muñoz, S., Búa, S., Rodríguez-Acebes, S., Megías, D., Ortega, S., de Martino, A., and Méndez, J. (2017). In Vivo DNA Re-replication Elicits Lethal Tissue Dysplasias. Cell Rep. *19*, 928–938.

Neph, S., Kuehn, M.S., Reynolds, A.P., Haugen, E., Thurman, R.E., Johnson, A.K., Rynes, E., Maurano, M.T., Vierstra, J., Thomas, S., et al. (2012). BEDOPS: high-performance genomic feature operations. Bioinformatics *28*, 1919–1920.

Nishitani, H., Morino, M., Murakami, Y., Maeda, T., and Shiomi, Y. (2014). Chromatin fractionation analysis of licensing factors in mammalian cells. Methods Mol. Biol. *1170*, 517–527.

Oguri, T., Bessho, Y., Achiwa, H., Ozasa, H., Maeno, K., Maeda, H., Sato, S., and Ueda, R. (2007). MRP8/ABCC11 directly confers resistance to 5-fluorouracil. Mol. Cancer Ther. *6*, 122–127.

Onozawa, H., Saito, M., Saito, K., Kanke, Y., Watanabe, Y., Hayase, S., Sakamoto, W., Ishigame, T., Momma, T., Ohki, S., and Takenoshita, S. (2017). Annexin A1 is involved in resistance to 5-FU in colon cancer cells. Oncol. Rep. *37*, 235–240.

Owatari, S., Akune, S., Komatsu, M., Ikeda, R., Firth, S.D., Che, X.F., Yamamoto, M., Tsujikawa, K., Kitazono, M., Ishizawa, T., et al. (2007). Copper-transporting P-type ATPase, ATP7A, confers multidrug resistance and its expression is related to resistance to SN-38 in clinical colon cancer. Cancer Res. *67*, 4860–4868.

Park, J.T., Chen, X., Tropè, C.G., Davidson, B., Shih, IeM., and Wang, T.L. (2010). Notch3 overexpression is related to the recurrence of ovarian cancer and confers resistance to carboplatin. Am. J. Pathol. *177*, 1087–1094.

Pereira, B., Chin, S.F., Rueda, O.M., Vollan, H.K., Provenzano, E., Bardwell, H.A., Pugh, M., Jones, L., Russell, R., Sammut, S.J., et al. (2016). The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat. Commun. *7*, 11479.

Pfleger, C.M., Lee, E., and Kirschner, M.W. (2001). Substrate recognition by the Cdc20 and Cdh1 components of the anaphase-promoting complex. Genes Dev. *15*, 2396–2407.

Piccione, M., Niceta, M., Antona, V., Di Fiore, A., Cariola, F., Gentile, M., and Corsello, G. (2009). Identification of two new mutations in TRPS 1 gene leading to the tricho-rhino-phalangeal syndrome type I and III. Am. J. Med. Genet. A. *149A*, 1837–1841.

Radvanyi, L., Singh-Sandhu, D., Gallichan, S., Lovitt, C., Pedyczak, A., Mallo, G., Gish, K., Kwok, K., Hanna, W., Zubovits, J., et al. (2005). The gene associated with trichorhinophalangeal syndrome in humans is overexpressed in breast cancer. Proc. Natl. Acad. Sci. USA *102*, 11005–11010.

Rhind, N., and Gilbert, D.M. (2013). DNA replication timing. Cold Spring Harb. Perspect. Biol. *5*, a010132.

Rossi, A., Devirgiliis, V., Panasiti, V., Borroni, R.G., Carlesimo, M., Gentile, M., Cariola, F., and Calvieri, S. (2007). Missense mutation in exon 7 of TRPS1 gene in an Italian family with a mild form of trichorhinophalangeal syndrome type I. Br. J. Dermatol. *157*, 1021–1024.

Sanborn, J.Z., Salama, S.R., Grifford, M., Brennan, C.W., Mikkelsen, T., Jhanwar, S., Katzman, S., Chin, L., and Haussler, D. (2013). Double minute chromosomes in glioblastoma multiforme are revealed by precise reconstruction of oncogenic amplicons. Cancer Res. *73*, 6036–6045.

Sanchez-Garcia, F., Villagrasa, P., Matsui, J., Kotliar, D., Castro, V., Akavia, U.D., Chen, B.J., Saucedo-Cuevas, L., Rodriguez Barrueco, R., Llobet-Navas, D., et al. (2014). Integration of genomic data enables selective discovery of breast cancer drivers. Cell *159*, 1461–1475.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675.

Shan, L., Zhou, X., Liu, X., Wang, Y., Su, D., Hou, Y., Yu, N., Yang, C., Liu, B., Gao, J., et al. (2016). FOXK2 Elicits Massive Transcription Repression and Suppresses the Hypoxic Response and Breast Cancer Carcinogenesis. Cancer Cell *30*, 708–722.

Shibata, Y., Kumar, P., Layer, R., Willcox, S., Gagan, J.R., Griffith, J.D., and Dutta, A. (2012). Extrachromosomal microDNAs and chromosomal microdeletions in normal tissues. Science *336*, 82–86.

Shibata, A., Tanahashi, K., Sugiura, K., and Akiyama, M. (2015). TRPS1 Haploinsufficiency Results in Increased STAT3 and SOX9 mRNA Expression in Hair Follicles in Trichorhinophalangeal Syndrome. Acta Derm. Venereol. *95*, 620–621.

Sidler, J.A., Filges, I., Boesch, N., Ramelli, G.P., Röthlisberger, B., Huber, A.R., Tercanli, S., Bronz, L., Miny, P., and Heinimann, K. (2012). TRPS1 codon 952 constitutes a mutational hot spot in trichorhinophalangeal syndrome type I and could be associated with intellectual disability. Clin. Dysmorphol. *21*, 87–90.

Singh, S., Srivastava, S.K., Bhardwaj, A., Owen, L.B., and Singh, A.P. (2010). CXCL12-CXCR4 signalling axis confers gemcitabine resistance to pancreatic cancer cells: a novel target for therapy. Br. J. Cancer *103*, 1671–1679.

Sliutz, G., Karlseder, J., Tempfer, C., Orel, L., Holzer, G., and Simon, M.M. (1996). Drug resistance against gemcitabine and topotecan mediated by constitutive hsp70 overexpression in vitro: implication of quercetin as sensitiser in chemotherapy. Br. J. Cancer *74*, 172–177.

Smaili, W., Elalaoui, S.C., Meier, S., Zerkaoui, M., Sefiani, A., and Heinimann, K. (2017). A novel TRPS1 mutation in a Moroccan family with Tricho-rhino-phalangeal syndrome type III: case report. BMC Med. Genet. *18*, 50.

Smith, M.L. (1999). Mdm2 sensitizes MCF7 breast cancer cells to cisplatin or carboplatin. Breast Cancer Res. Treat. *58*, 99–105.

Stephens, P.J., Tarpey, P.S., Davies, H., Van Loo, P., Greenman, C., Wedge, D.C., Nik-Zainal, S., Martin, S., Varela, I., Bignell, G.R., et al.; Oslo Breast Cancer Consortium (OSBREAC) (2012). The landscape of cancer genes and mutational processes in breast cancer. Nature *486*, 400–404.

Stewart, D.J. (2007). Mechanisms of resistance to cisplatin and carboplatin. Crit. Rev. Oncol. Hematol. *63*, 12–31.

Tasic, V., Gucev, Z., Ristoska-Bojkovska, N., Janchevska, A., and Lüdecke, H.J. (2014). Tricho-rhino-phalangeal syndrome in a 13-year-old girl with chronic renal failure and severe growth retardation. Ren. Fail. *36*, 619–622.

Tom, T.D., Malkas, L.H., and Hickey, R.J. (1996). Identification of multiprotein complexes containing DNA replication factors by native immunoblotting of HeLa cell protein preparations with T-antigen-dependent SV40 DNA replication activity. J. Cell. Biochem. *63*, 259–267.

Turner, K.M., Deshpande, V., Beyter, D., Koga, T., Rusert, J., Lee, C., Li, B., Arden, K., Ren, B., Nathanson, D.A., et al. (2017). Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. Nature *543*, 122–125.

Ullah, A., Umair, M., Hussain, S., Jan, A., and Ahmad, W. (2018). Sequence variants in GDF5 and TRPS1 underlie brachydactyly and tricho-rhino-phalangeal syndrome type III. Pediatr. Int. (Roma) *60*, 304–306.

Van Den Broeck, A., Nissou, D., Brambilla, E., Eymin, B., and Gazzeri, S. (2012). Activation of a Tip60/E2F1/ERCC1 network in human lung adenocarcinoma cells exposed to cisplatin. Carcinogenesis *33*, 320–325.

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., and Mann, M. (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. Cell *142*, 967–980.

Wang, X., Wang, Y., Gu, J., Zhou, D., He, Z., Wang, X., and Ferrone, S. (2017a). ADAM12-L confers acquired 5-fluorouracil resistance in breast cancer cells. Sci. Rep. *7*, 9687.

Wang, Z., Chen, Y., Lin, Y., Wang, X., Cui, X., Zhang, Z., Xian, G., and Qin, C. (2017b). Novel crosstalk between KLF4 and ZEB1 regulates gemcitabine resistance in pancreatic ductal adenocarcinoma. Int. J. Oncol. *51*, 1239–1248.

Wang, Y., Zhang, J., Wu, L., Liu, W., Wei, G., Gong, X., Liu, Y., Ma, Z., Ma, F., Thiery, J.P., and Chen, L. (2018a). Tricho-rhino-phalangeal syndrome 1 protein functions as a scaffold required for ubiquitin-specific protease 4-directed histone deacetylase 2 de-ubiquitination and tumor growth. Breast Cancer Res. *20*, 83.

Wang, Z., Zhao, X., Wang, W., Liu, Y., Li, Y., Gao, J., Wang, C., Zhou, M., Liu, R., Xu, G., and Zhou, Q. (2018b). ZBTB7 evokes 5-fluorouracil resistance in colorectal cancer through the NF-κB signaling pathway. Int. J. Oncol. *53*, 2102–2110.

Weddington, N., Stuy, A., Hiratani, I., Ryba, T., Yokochi, T., and Gilbert, D.M. (2008). ReplicationDomain: a visualization tool and comparative database for genome-wide replication timing data. BMC Bioinformatics *9*, 530.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer International Publishing).

Witwicki, R.M., Ekram, M.B., Qiu, X., Janiszewska, M., Shu, S., Kwon, M., Trinh, A., Frias, E., Ramadan, N., Hoffman, G., et al. (2018). TRPS1 Is a Lineage-Specific Transcriptional Dependency in Breast Cancer. Cell Rep. *25*, 1255–1267.e5.

Wood, J., Pring, M., Eveson, J.W., Price, N., Proby, C.M., and Hague, A. (2011). Co-overexpression of Bag-1 and heat shock protein 70 in human epidermal squamous cell carcinoma: Bag-1-mediated resistance to 5-fluorouracil-induced apoptosis. Br. J. Cancer *104*, 1459–1471.

Xouri, G., Squire, A., Dimaki, M., Geverts, B., Verveer, P.J., Taraviras, S., Nishitani, H., Houtsmuller, A.B., Bastiaens, P.I., and Lygerou, Z. (2007). Cdt1 associates dynamically with chromatin throughout G1 and recruits Geminin onto chromatin. EMBO J. *26*, 1303–1314.

Ye, D., Fei, Y., Sheng, Y.E., Qiao, J.J., and Dong, F.Q. (2017). Analysis of a Chinese pedigree with trichorhinophalangeal syndrome derived from a missense mutation in the TRPS1 gene. Clin. Exp. Dermatol. *42*, 432–434.

Zeng, X., Sigoillot, F., Gaur, S., Choi, S., Pfaff, K.L., Oh, D.C., Hathaway, N., Dimova, N., Cuny, G.D., and King, R.W. (2010). Pharmacologic inhibition of the anaphase-promoting complex induces a spindle checkpoint-dependent mitotic arrest in the absence of spindle damage. Cancer Cell *18*, 382–395.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhang, Y., Yang, X., Gui, B., Xie, G., Zhang, D., Shang, Y., and Liang, J. (2011). Corepressor protein CDYL functions as a molecular bridge between polycomb repressor complex 2 and repressive chromatin mark trimethylated histone lysine 27. J. Biol. Chem. *286*, 42414–42425.

Zhang, Y., Zhang, D., Li, Q., Liang, J., Sun, L., Yi, X., Chen, Z., Yan, R., Xie, G., Li, W., et al. (2016). Nucleation of DNA repair factors by FOXA1 links DNA demethylation to transcriptional pioneering. Nat. Genet. *48*, 1003–1013.

Zhuo, W.L., Wang, Y., Zhuo, X.L., Zhang, Y.S., and Chen, Z.T. (2008). Short interfering RNA directed against TWIST, a novel zinc finger transcription factor, increases A549 cell sensitivity to cisplatin via MAPK/mitochondrial pathway. Biochem. Biophys. Res. Commun. *369*, 1098–1102.

Zielke, N., Querings, S., Rottig, C., Lehner, C., and Sprenger, F. (2008). The anaphase-promoting complex/cyclosome (APC/C) is required for rereplication control in endoreplication cycles. Genes Dev. *22*, 1690–1703.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit monoclonal anti-TRPS1 (clone EPR16171)	Abcam	Cat#ab209664
Rabbit monoclonal anti-MCMC3 (clone EPR7081)	Abcam	Cat#ab126723; RRID: AB_11127222
Rabbit monoclonal anti-MCMC5 (clone EP2682Y)	Abcam	Cat#ab76023; RRID: AB_1310438
Rabbit monoclonal anti-MCMC6 (clone EPR17686)	Abcam	Cat#ab201683
Rabbit monoclonal anti-MCMC7 (clone EP1974Y)	Abcam	Cat#ab52489; RRID: AB_881187
Rabbit monoclonal anti-CDC16 (clone EPR11168)	Abcam	Cat#ab169536
Rabbit monoclonal anti-CDC23 (clone EPR11827)	Abcam	Cat#ab177148
Rabbit monoclonal anti-CDC45 (clone EPR5759)	Abcam	Cat#ab126762; RRID: AB_11140216
Rabbit monoclonal anti-RPA1 (clone EPR3472)	Abcam	Cat#ab79398; RRID: AB_1603759
Mouse monoclonal anti-PCNA (clone PC10)	Abcam	Cat#ab29; RRID: AB_303394
Rabbit monoclonal anti-Mi-2β (clone EPR12229)	Abcam	Cat#ab181370
Rabbit polyclonal anti-HDAC1	Abcam	Cat#ab7028; RRID: AB_305705
Rabbit polyclonal anti-HDAC2	Abcam	Cat#ab7029; RRID: AB_305706
Rabbit polyclonal anti-LSD1	Abcam	Cat#ab17721; RRID: AB_443964
Rabbit monoclonal anti-CtBP1 (clone EPR6800)	Abcam	Cat#ab129181; RRID: AB_11156048
Rabbit monoclonal anti-CtBP2 (clone EPR7611(B))	Abcam	Cat#ab128871; RRID: AB_11143956
Rabbit polyclonal anti-H3	Abcam	Cat#ab1791; RRID: AB_302613
Rabbit monoclonal anti-H3K9me3 (clone EPR16601)	Abcam	Cat#ab176916; RRID: AB_2797591
Rabbit polyclonal anti-H2A	Abcam	Cat#ab88770; RRID: AB_10672053
Rabbit monoclonal anti- Phospho-MCM2 (S40) (clone EPR4170(2))	Abcam	Cat#ab133243; RRID: AB_11154969
Rabbit polyclonal anti-ORC1	Abcam	Cat#ab85830; RRID: AB_2157593
Rabbit monoclonal anti-BrdU (clone BU1/75 (ICR1))	Abcam	Cat#ab6326; RRID: AB_305426
Rabbit polyclonal anti-SUPT16H	Abcam	Cat#ab117081; RRID: AB_10899485
Rabbit polyclonal anti-MCM4	Abcam	Cat#ab154315
Rabbit monoclonal anti-RbAp48 (clone EPR3412)	Abcam	Cat#ab92344; RRID: AB_2177634
Mouse monoclonal anti-β-Actin (clone C-4)	Santa Cruz Biotechnology	Cat#sc-47778; RRID: AB_626632
Mouse monoclonal anti-POLD1 (clone A-9)	Santa Cruz Biotechnology	Cat#sc-17776; RRID: AB_675487
Mouse monoclonal anti-ORC2 (clone H-300)	Santa Cruz Biotechnology	Cat#sc-28742; RRID: AB_2157712

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse monoclonal anti-MTA1 (clone E-12)	Santa Cruz Biotechnology	Cat#sc-373765, RRID: AB_10917039
Mouse monoclonal anti-MTA2 (clone F-9)	Santa Cruz Biotechnology	Cat#sc-55566; RRID: AB_831564
Mouse monoclonal anti-RbAp46 (clone E-9)	Santa Cruz Biotechnology	Cat#sc-377197
Mouse monoclonal anti-Geminin (clone F-7)	Santa Cruz Biotechnology	Cat#sc-74456; RRID: AB_1124963
Mouse monoclonal anti-FLAG (clone M2)	Sigma-Aldrich	Cat# F3165; RRID: AB_259529
Rabbit polyclonal anti-RFC3	Proteintech	Cat#11814-1-AP; RRID: AB_2178470
Rabbit polyclonal anti-RFC4	Proteintech	Cat#10806-1-AP; RRID: AB_2178477
Rabbit polyclonal anti-ANAPC1	Proteintech	Cat#21748-1-AP; RRID: AB_10733241
Rabbit polyclonal anti-MBD2	Proteintech	Cat#55200-1-AP; RRID: AB_10852963
Rabbit polyclonal anti-SSRP1	Proteintech	Cat# 15696-1-AP; RRID: AB_2195756
Rabbit monoclonal anti-MCM2 (clone D7G11)	Cell Signaling Technology	Cat#3619; RRID: AB_2142137
Rabbit polyclonal anti-Phospho-CHK2 (Thr68)	Cell Signaling Technology	Cat#2661; RRID: AB_331479
Rabbit polyclonal anti-CHK2	Cell Signaling Technology	Cat#2662; RRID: AB_2080793
Mouse monoclonal anti-GST (clone 1B10)	Biodragon	Cat#B1007
Mouse monoclonal anti-BrdU (clone B44)	BD Biosciences	Cat#347580; RRID: AB_10015219
Mouse monoclonal anti- Phospho-H2A.X (Ser139) (clone JBW301)	Millipore	Cat#05-636; RRID: AB_309864
Mouse monoclonal anti- Biotin	Jackson ImmunoResearch Labs	Cat#200-002-211; RRID: AB_2339006
Chemicals, peptides, and recombinant proteins		
XenoLight D-Luciferin Potassium Salt	Perkin Elmer	Cat#122799
Thymidine	Sigma Aldrich	Cat#T9250; CAS: 50-89-5
EdU (5-ethynyl-2′-deoxyuridine)	Sigma Aldrich	Cat#900584; CAS: 61135-33-9
BrdU (5-Bromo-2′-deoxyuridine)	Sigma Aldrich	Cat#B9285; CAS: 59-14-3
IdU (5-Iodo-2′-deoxyuridine)	Sigma Aldrich	Ca#I7125; CAS: 54-42-2
FLAG® Peptide	Sigma Aldrich	Ca#F3290
TAME (tosyl arginine methyl ester)	Selleck Chemicals	Ca#S2225; CAS: 901-47-3
Bleomycin Sulfate	Selleck Chemicals	Ca#S1214; CAS: 9041-93-4
Mitomycin C	Selleck Chemicals	Ca#S8146; CAS: 50-07-7
Topotecan HCl	Selleck Chemicals	Ca#S1231; CAS: 119413-54-6
Epirubicin HCl	Selleck Chemicals	Ca#S1223; CAS: 56390-09-1
Paclitaxel	Selleck Chemicals	Ca#S1150; CAS: 33069-62-4
Nocodazole	Selleck Chemicals	Ca#S2775; CAS: 31430-18-9
Carboplatin	Selleck Chemicals	Ca#S1215; CAS: 41575-94-4
Cisplatin	Selleck Chemicals	Ca#S1166; CAS: 15663-27-1
Thiotepa	Selleck Chemicals	Ca#S1775; CAS: 52-24-4
Mechlorethamine HCl	Selleck Chemicals	Ca#S4252; CAS: 55-86-7
Gemcitabine	Selleck Chemicals	Ca#S1714; CAS: 95058-81-4
5-FU (5-Fluorouracil)	Selleck Chemicals	Ca#S1209; CAS: 51-21-8
CHX (Cycloheximide)	Selleck Chemicals	Ca#S7418; CAS: 66-81-9
Hydrocortisone	Selleck Chemicals	Ca#S1696; CAS: 50-23-7
Human EGF Recombinant Protein	GIBCO	Ca#PHG0311
Insulin	M&C GENE TECHNOLOGY	Ca#CC101
Cholera Toxin	M&C GENE TECHNOLOGY	Ca#CC104
Critical commercial assays		
CellTiter 96® AQ_{ueous} One Solution Cell Proliferation Assay (MTS)	Promega	Ca#G3582

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
TnT® Quick Coupled Transcription/ Translation System	Promega	Ca#L1170
Senescence Detection Kit	Abcam	Ca#ab65351
Pierce Silver Stain Kit	Thermo Scientific	Ca#24612
Pierce BCA Protein Assay Kit	Thermo Scientific	Ca#23225
Click-iT EdU Alexa Fluor 488 Flow Cytometry Assay Kit	Invitrogen	Ca#C10420
Click-iT EdU Alexa Fluor 647 Flow Cytometry Assay Kit	Invitrogen	Ca#C10419
Lipofectamine 3000 Transfection Reagent	Invitrogen	Ca#L3000015
Lipofectamine RNAiMAX Transfection Reagent	Invitrogen	Ca#13778030
TaqMan Gene Expression Master Mix	Applied Biosystems	Ca#4369016
HiSpeed Plasmid MidiKit	QIAGEN	Ca#12643
REPLI-g Mini Kit	QIAGEN	Ca#150023
QIAquick PCR Purification Kit	QIAGEN	Ca#28106
FITC Annexin V Apoptosis Detection Kit I	BD Biosciences	Ca#556547
Apoptosis Inducer Kit (TS, TNF-α+SM-164)	Beyotime Biotechnology	Ca#C0006S
Deposited data		
ChIP-seq, Repli-seq, eccDNA sequencing and CNA microarray data	This paper	GEO: GSE118643, GSE135598
Whole genome sequencing data	This paper	SRA: SRP217294
Experimental models: cell lines		
Human: MCF-10A	ATCC	ATCC CRL-10317
Human: MCF-7	ATCC	ATCC HTB-22
Human: MDA-MB-231	ATCC	ATCC HTB-26
Human: T-47D	ATCC	ATCC HTB-133
Human: U2OS	ATCC	ATCC HTB-96
Human: HEK293T	ATCC	ATCC CRL-3216
Human: HeLa	ATCC	ATCC CCL-2
MDA-MB-231-Luc-D3H2LN cell	(Jenkins et al., 2005)	N/A
Mouse: ATDC5	KeyGEN BioTECH	Ca#KG445
Experimental models: organisms/strains		
Mouse: SCID beige	Vital River Laboratories	N/A
Oligonucleotides		
siRNA targeting sequence: siControl: UUCUCCGAACGUGUCACGU	This paper	N/A
siRNA targeting sequence: siTRPS1#1: GCAAAUCACCUGCAAGGAU	This paper	N/A
siRNA targeting sequence: siTRPS1#2: GUACAUGAGACCUGCGAAA	This paper	N/A
siRNA targeting sequence: siCtBP1: ACGACUUCACCGUCAAGCA	This paper	N/A
siRNA targeting sequence: siMi-2β: CCGCAAGAAACUCCGAACCACUAAA	This paper	N/A
siRNA targeting sequence: siCDC23#1: GCUGCCCAGUGUUACAUCAAAUAUA	This paper	N/A
siRNA targeting sequence: siCDC23#2: CCAAGCUCGAGAACUUGAUGGAUUU	This paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
shRNA targeting sequence: siTRPS1#1: ATATGGTAACGAGCTATAATT	This paper	N/A
shRNA targeting sequence: siTRPS1#2: GGCAGGACAAGATAACAGTCA	This paper	N/A
Primers for qPCR, See Table S3	This paper	N/A
Recombinant DNA		
pLVX-PGK-Puro-FLAG-Vector	This paper	N/A
pLVX-PGK-Puro-FLAG-TRPS1	This paper	N/A
pLVX-PGK-Puro-FLAG-TRPS1C896Y	This paper	N/A
pLVX-PGK-Puro-FLAG-TRPS1R908Q	This paper	N/A
pLVX-PGK-Puro-FLAG-TRPS1ΔZF1-3	This paper	N/A
pLVX-PGK-Puro-FLAG-TRPS1ΔZF4-6	This paper	N/A
pLVX-PGK-Puro-FLAG-TRPS1ΔGATA	This paper	N/A
pLVX-PGK-Puro-FLAG-TRPS1ΔIKAROS	This paper	N/A
pLVX-shRNA2-Puro-shControl	This paper	N/A
pLVX-shRNA2-Puro-TRPS1#1	This paper	N/A
pLVX-shRNA2-Puro-TRPS1#2	This paper	N/A
pGEX-6P1-TRPS1/D1	This paper	N/A
pGEX-6P1-TRPS1/D2	This paper	N/A
pGEX-6P1-TRPS1/D3	This paper	N/A
pGEX-6P1-TRPS1/D4	This paper	N/A
pGEX-6P1-TRPS1/D5	This paper	N/A
pGEX-6P1-TRPS1/D6	This paper	N/A
pGEX-6P1-TRPS1/D6C896Y	This paper	N/A
pGEX-6P1-TRPS1/D6R908Q	This paper	N/A
pCDNA3.1(+)-TRPS1	This paper	N/A
pCDNA3.1(+)-TRPS1/C896Y	This paper	N/A
pCDNA3.1(+)-TRPS1/R908Q	This paper	N/A
Software and algorithms		
Image-Pro Plus 7	Media Cybernetics	N/A
FlowJo v10.0.7	Tree Star	N/A
Living Image v4.3.1	Perkin Elmer	N/A
ZEN v2.3 (blue edition)	Carl Zeiss	N/A
GraphPad v7.0	Prism	N/A
ImageJ v1.50i	Schneider et al., 2012	https://imagej.nih.gov/ij/
MACS v2	Zhang et al., 2008	https://github.com/macs3-project/MACS
ChromHMM v1.22	Ernst and Kellis, 2012	http://compbio.mit.edu/ChromHMM/
BEDOPS v2.4.39	Neph et al., 2012	https://bedops.readthedocs.io/en/latest/
GenomicRanges v1.42.0	Lawrence et al., 2013	https://bioconductor.org/packages/release/bioc/html/GenomicRanges.html
R v3.6.3	R Core Team	http://www.R-project.org
ggplot2 v3.3.2	Wickham, 2016	https://ggplot2.tidyverse.org/
HOMER v4.4	Heinz et al., 2010	http://homer.ucsd.edu/homer/
CNVnator v0.2.7	Abyzov et al., 2011	http://sv.gersteinlab.org/cnvnator/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Yongfeng Shang (yshang@hsc.pku.edu.cn).

Materials availability

All reagents generated in this study are available upon request to the Lead Contact.

Data and code availability

ChIP-seq, Repli-seq, eccDNA sequencing and CNA microarray data are deposited at the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) with accession number GSE118643 and GSE135598. Whole genome sequencing data was deposited at the Sequence Read Archive (SRA) database (https://www.ncbi.nlm.nih.gov/sra/) with accession number SRP217294.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mouse models

Six-week-old immunocompromised female SCID beige mice were purchased from the Charles River. Animal handling and procedures were approved by the Peking University Health Science Center Institutional Animal Care and Use Committee (LA2015019).

Cell lines

HEK293T, MCF-7, MDA-MB-231, MDA-MB-231-Luc-D3H2LN, U2OS and HeLa cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS). ATDC5 cells were cultured in DMEM/F12 medium supplemented with 10% FBS. T-47D cells were cultured in RPMI-1640 medium supplemented with 10% FBS. MCF-10A were cultured in DMEM/F12 medium supplemented with 5% horse serum, 20 ng/mL EGF, 500 ng/mL hydrocortisone, 100 ng/mL cholera toxin, and 10 μg/mL insulin. Cells were maintained in a humidified incubator equilibrated with 5% CO_2 at 37°C.

METHOD DETAILS

Cell transfection

Transfections were conducted with Lipofectamine 3000 Transfection Reagent (Invitrogen) or Lipofectamine RNAiMAX Transfection Reagent (Invitrogen) according to the manufacturer's instruction. The siRNAs were synthesized by Shanghai GenePharma.

Cell proliferation assay

Cells were seeded into 96-well plates. The CellTiter 96® AQueous ONE Solution Reagent (MTS, Promega) was added according to the manufacturer's instruction before cell harvesting. Plates were incubated at 37°C for 3 h and cell viability was determined by measuring the absorbance of converted dye at a wavelength of 490 nm. Each experiment was performed in at least triplicates.

Immunohistochemistry assay

Human breast tumor tissue array was purchased from Shanghai Outdo Biotech. The immunohistochemical analysis was performed as previously described with minor modifications (He et al., 2018). Briefly, after antigen retrieval by use of sodium citrate, the samples were blocked in 10% normal goat serum for 10 min at room temperature (RT) followed by an overnight incubation with anti-TRPS1 antibody (1:500) at 4°C. After washing with PBS, the samples were incubated with goat anti-rabbit HRP (DAKO, Agilent) at RT for 30 min. The samples were then developed with DAB (3,3′-diaminobenzide tetrahydrochloride) buffer before counterstaining with hematoxylin. The staining intensity score was defined as follows: 0, negative; 1, weak; 2, moderate; 3, strong. The positive staining ratio score was defined according to the percentage of positive tumor cells as follows: 0, < 5%; 1, 5%–25%; 2, 25%–50%; 3, 50%–75%; 4, >75%. The final histopathologic score was calculated as product of these two values. The statistical analysis was performed by comparing the histopathologic score from tumor samples to adjacent normal tissues with paired t test (two-tailed).

Senescence detection assay

Cellular senescence was examined by use of Senescence Detection Kit (Abcam) following the manufacturer's instruction. Briefly, 48 h after transfection, the cells were washed once with PBS, fixed with fixative solution for 10 min at RT, and then incubated with staining solution mix at 37°C for 2 h (without CO_2). Senescent cells showing blue color were counted under a microscope. Adriamycin was treated to induce cellular senescence as a positive control (Elmore et al., 2002). The experiment was performed in triplicates.

Apoptosis detection assay

Cell apoptosis was analyzed with the FITC Annexin V Apoptosis Detection Kit I (BD Biosciences) following the manufacturer's instruction. Briefly, 48 h after transfection, the cells were resuspended in binding buffer at a concentration of 1×10^6 cells/mL. Then ~1×10^5 cells were transferred to a 5 mL culture tube and labeled with FITC Annexin V and propidium iodide for 15 min at RT in the dark. A FACS Canto II cytometer (BD) was then used for flow cytometry analysis and the data was processed by use of FlowJo 10.0.7 (Tree Star). TNF-α/SM-164 was treated to induce apoptosis as a positive control. The experiment was performed in triplicates.

Immunofluorescence assay

MCF-10A cells stably expressing control, wild-type TRPS1, or TRPS1 mutants were fixed in 4% paraformaldehyde for 10 min. After being permeabilized with 0.2% Triton X-100, the cells were successively incubated with primary antibodies, fluorescent-dye conjugated secondary antibodies, and DAPI (4',6-diamidino-2-phenylindole, Dihydrochloride, Sigma). Images were acquired with a Zeiss LSM880 fluorescence microscope (ZEN software, Carl Zeiss Inc).

Cycloheximide chase assay

MCF-10A cells stably expressing control, wild-type TRPS1, or TRPS1 mutants were treated with 50 μg/mL cycloheximide (CHX) for a variety of time intervals before collecting cell extracts. Then protein lysate was resolved on SDS-PAGE and analyzed by western blotting.

In vivo metastasis

MDA-MB-231-Luc-D3H2LN cells (Xenogen Corporation) were infected with lentiviruses carrying empty vector or FLAG-tagged TRPS1/WT-, TRPS1/C896Y-, or TRPS1/R908Q-expressing plasmid. Six-week-old immunocompromised female SCID beige mice (Charles River, Beijing, China) were divided into 4 groups and inoculated the above cells (5×10^6 cells per mouse) at the fourth right abdominal mammary fat pad. For bioluminescence imaging, mice were anesthetized and injected with D-luciferin intraperitoneally, and bioluminescence images were obtained with IVIS *in vivo* imaging system (Living Image Software, Perkin Elmer). Bioluminescence from relative optical intensity was defined manually, and data was expressed as photon flux (photons/s/cm2/Steradian). All the 6 mice used for xenograft tumor growth and metastasis analysis in each group of gene manipulations were shown. They were not selected based on whether they had metastases or not. Except for the inoculated cancer cells with different infections, the treatment, including operation and measurement on each group of mice was same. However, to facilitate the detection of metastatic events that were relatively weak compared to the primary tumors, a longer exposure was used in Figure 1F than in Figure 1E.

Immunopurification and mass spectrometry

HEK293T cells stably expressing FLAG-TRPS1 were lysed in lysis buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.3% Nonidet P-40, 1 mM DTT, and 5 mM EDTA) containing protease inhibitor cocktail (Roche) at 4°C for 40 min. The lysates were centrifuged at 14,000 g at 4°C for 15 min and protein supernatant was incubated with anti-FLAG M2 gel (Sigma) at 4°C for 4 h. After washing 5 times with lysis buffer, protein complexes were eluted from the beads with 0.1 mg/mL FLAG peptides. The eluted protein complexes were resolved on SDS-PAGE and silver stained with the Pierce Silver Stain Kit (Thermo Scientific) and subjected to LC-MS/MS for sequencing and data analysis. In Figure 2D, MCF-7 cells stably expressing empty vector or FLAG-TRPS1/WT, FLAG-TRPS1/C896Y, or FLAG-TRPS1/R908Q were synchronized in G_1 phase by serum starvation or S phase by double thymidine blocking before immunopurification and mass spectrometry analysis.

Cell cycle synchronization

Serum starvation was used for cell cycle synchronization in G_0/G_1. Briefly, cells were seeded and maintained in complete culture medium (DMEM+10% FBS) for 6-8 h before culturing with low serum culture medium (DMEM+0.1% FBS) for 36 h. The cells were then released by replacing back to complete culture medium for a variety of time intervals for cell cycle analysis. Cells in G_1 phase validated by flow cytometry were then collected for following experiments. The double thymidine block method was used for synchronization in S. Briefly, cells were seeded and cultured for 6-8 h before first round of treatment with 2 mM thymidine for 16 h. Then cells were released by removing thymidine and culturing in complete medium for 12 h. After second round of thymidine treatment for another 16 h, cells were released and collected for cell cycle analysis. Cells in early/middle/late S phase validated by flow cytometry were collected after release for ~4/6/8 h.

Fast protein liquid chromatography

FPLC was performed according to the procedures described previously with minor modifications (Shan et al., 2016). Briefly, nuclear extracts of MCF-7 cells were prepared and applied to Superose 6 size exclusion column (Amersham Biosciences). The column was eluted at a flow rate of 0.5 ml/min and fractions were collected separately. The eluted protein complexes were resolved on SDS-PAGE and analyzed by western blotting.

Co-immunoprecipitation and western blotting

Cellular lysates were prepared by incubating the cells in lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.3% NP-40, and 2 mM EDTA) containing protease inhibitor cocktail (Roche), followed by centrifugation at 13,400 rpm at 4°C for 15 min. The protein concentration of the lysates was determined using the Pierce BCA Protein Assay Kit (Thermo Scientific). For immunoprecipitation, 500 μg of protein was incubated with 2 μg specific antibodies or normal IgG at 4°C for 12 h with rotation. Protein A or G beads (Thermo Scientific) were added, and the incubation was continued for an additional 3 h. Beads were washed 5 times using the lysis buffer. The precipitated proteins were eluted from the beads by resuspending the beads in SDS sample buffer and boiling for 10 min. The resultant materials from immunoprecipitation or cell lysates were resolved on SDS-PAGE and analyzed by western blotting.

EdU incorporation assay and flow cytometry

EdU incorporation assay was performed according to the Click-iT® EdU Flow Cytometry Assay Kits (Invitrogen). Briefly, cells were pulse-labeled with 10 μM EdU for 2 h. After harvesting and washing with 1% BSA in PBS, cells were fixed in Click-iT® fixative for 15 min at RT. Cells were dislodged in 1 × Click-iT® saponin-based permeabilization and wash reagent at RT for 15 min. Click-iT® reactions were performed with Alexa Fluor® 488/647 azide at RT for 30 min. After digesting RNAs with RNase A and staining cells with propidium iodide, flow cytometry was performed in a FACS Canto II cytometer (BD) and data were analyzed with FlowJo (Tree Star).

DNA fiber assay

Single molecule analysis of DNA replication was performed by DNA fiber assay with minor modifications (Muñoz et al., 2017). Briefly, MCF-7 cells were sequentially pulse-labeled with 50 μM CldU for 2 h and 250 μM IdU for 30 min. Cells were then trypsinized and resuspended in PBS at a concentration of 500 cells/μL. ~1000 cells were lysed in lysis buffer (0.2 M Tris-HCl, pH 7.4, 50 mM EDTA, and 0.5% SDS) on each microscope slide at RT for 5 min, and the slides were subsequently tilted at 15 degrees to spread DNA fibers. Slides were air-dried and fixed in methanol/acetic acid (3:1) at RT for 10 min. After denatured with 2.5 M HCl at RT for 80 min, DNA was incubated in blocking buffer (0.1% Triton X-100, 1% BSA in PBS) at RT for 1 h. The slides were incubated with primary antibodies (anti-CldU and anti-IdU) at 4°C overnight and then secondary antibodies (Alexa Fluors 488 and Alexa Fluors 555 from Invitrogen) at RT for 1 h. Slides were air-dried and mounted with Prolong (Invitrogen). Images were obtained with Leica TCS-SP8 STED 3X microscope and analyzed with ImageJ software.

Cesium chloride gradient centrifugation

Rereplication analysis of DNA was performed by cesium chloride gradient centrifugation assay with minor modifications (Black et al., 2013). Briefly, MCF-7 cells were treated with 100 mM BrdU for 14 h and lysed in lysis buffer (20 mM Tris-HCl, pH 8.0, 4 mM EDTA, 20 mM NaCl, and 2% SDS) supplemented with RNase A at 37°C for 2 h followed by incubatation with proteinase K at 55°C for 3 h. DNA was extracted and then digested with *Eco*R I and *Bam*H I (NEB) at 37°C overnight. ~100 μg DNA was mixed with CsCl (1 g/mL) in TE buffer (refractive index of 1.4015-1.4031). The CsCl gradient was centrifuged at 50,000 rpm in an MLN-80 rotor at 25°C for 20 h. Fractions were collected from the bottom of the gradient in 200 μL aliquots and dialyzed against 0.1 × TE buffer. DNA concentration was measured by Qubit 3.0.

Chromatin fractionation

Chromatin fractionation was performed as described elsewhere with minor modifications (Nishitani et al., 2014). Briefly, cells were synchronized in a specific phase of cell cycle before collection. ~3 × 10^6 cells were lysed with 150 μL CSK buffer (0.5% Triton X-100, 0.1 M NaCl, 10 mM PIPES pH 7.0, 300 mM sucrose, 1 mM $MgCl_2$, 1 mM EDTA, 2 mM phenylmethylsulfonyl fluoride, 10 mM NaF, 20 mM β-glycerophosphate, and 100 μM Na_3VO_4) containing protease inhibitor cocktail (Roche). The lysates were kept on ice for 20 min while vortexing about 10 s every 5 min. After pipetting the lysates 5 more times, 50 μL of lysates were saved in a new tube as a whole cell extract (WCE) sample, and the rest of lysates were centrifuged at 5,000 rpm for 5 min. The pellets were washed with 200 μL CSK buffer and resuspended in 100 μL CSK buffer as a chromatin binding sample. Finally, WCE and chromatin binding samples were boiled in SDS sample buffer, resolved on SDS-PAGE, and analyzed by western blotting.

GST pull-down assay

GST or GST fusion proteins were expressed in BL21 *E. coli*. HA-tagged CDC16 and CDC23 proteins were obtained using TnT® Quick Coupled Transcription/Translation System (Promega). Equal amounts of GST alone or GST fusion proteins were immobilized on glutathione-Sepharose 4B beads (Amersham Biosciences) in binding buffer (10 mM HEPES, pH 7.6, 3 mM MgCl2, 100 mM KCl, 5 mM EDTA, 5% glycerol, and 0.5% Nonidet P-40) containing protease inhibitor cocktail (Roche). After incubation at 4°C for 1 h with rotation, beads were washed 3 times with binding buffer and resuspended in binding buffer before adding *in vitro* transcribed/translated proteins at 4°C for 2 h with rotation. The beads were then washed 3 times with binding buffer and bound proteins were eluted by boiling in SDS sample buffer, resolved on SDS-PAGE, and analyzed by western blotting.

Peptide pull-down assay

GST or GST-fused TRPS1 mutant proteins were expressed in BL21 *E. coli* and resolved on SDS-PAGE for Coomassie brilliant blue staining. Peptide pull-down assay (Zhang et al., 2011) was performed by incubating 2.5 μg histone peptide with streptavidin beads in binding buffer (50 mM Tris-HCl, pH7.5, 150 mM NaCl, 0.05% Nonidet P-40, 0.3 mg/mL BSA, and protease inhibitor cocktail (Roche)) containing 0.1% Nonidet P-40 and 300 mM NaCl at 4°C overnight. Recovered peptides were examined by dot blotting assay using anti-biotin or H3K9me3 antibodies. Beads were washed 5 times with binding buffer and 0.5 μg GST alone or GST-fused TRPS1 mutants were added to determine the binding affinity for H3K9me3 with unmodified H3 peptides (1-21) as the control. The beads were washed and bound proteins were eluted by boiling in SDS sample buffer and analyzed by western blotting using anti-GST antibody.

Chromatin immunoprecipitation

ChIP assay was performed as described previously (Zhang et al., 2016). Briefly, ~2 × 10^7 cells were crosslinked using 1% formaldehyde at RT for 10 min and then quenched by 125mM glycine buffer for 5min. The fixed cells were resuspended in lysis

buffer (1% SDS, 5mM EDTA, and 50mM Tris-HCl, pH 8.1) containing protease inhibitors cocktail (Roche) and sonicated with Bioruptor® Plus sonication device (Bioruptor) to generate chromatin fragments. Lysates were diluted with dilution buffer (1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, pH 8.1, and 150mM NaCl) containing protease inhibitor cocktail (Roche). The diluted lysates were incubated with normal IgG or specific antibodies at 4°C for 12 h with rotation. Protein A or G beads (Thermo Scientific) were added, and the incubation was continued for an additional 2 h. Beads were sequentially washed with TSE I buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, pH 8.1, and 150mM NaCl), TSE II buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, pH 8.1, and 500mM NaCl), buffer III (0.25 M LiCl, 1% Nonidet P-40, 1% sodium deoxycholate, 1mM EDTA, and 10mM Tris-HCl, pH 8.1), and TE buffer(1mM EDTA, 10mM Tris-HCl, pH 8.0). The immunoprecipitated DNA-protein complex or input was heated at 65°C for 6 h in elution buffer (1% SDS, 0.1M NaHCO3) to reverse the formaldehyde crosslinking. DNA was purified with QIAquick PCR Purification Kit (QIAGEN) and analyzed by qPCR using primers described in Table S3.

ChIP-seq

Genomic DNAs were prepared as described above. For sequencing library construction, the purified DNA was subjected to blunt-ending, dA addition to 3′ end, and adaptor ligation. In-depth whole genome DNA sequencing was performed by the Beijing Genome Institute (BGI). Sequence reads were aligned to Genome Reference Consortium GRCh37 of human genome through bowtie (Langmead et al., 2009). TRPS1 peaks were called using MACS (Zhang et al., 2008) with a *P value* cutoff of 1e-5. ChIP-seq reads density was normalized with the total number of mappable reads in relevant sequencing library. Chromatin-state was annotated by ChromHMM (Ernst and Kellis, 2012) using ChIP-seq data on H3K4me1, H3K4me3, H3K9me3, H3K27me3, H3K36me3, and H3K27ac in MCF-7 cells (Consortium, 2012). Seven classes of chromatin states were distinguished and referred as: E1: polycomb repressed state enriched with H3K27me3, E2: heterochromatin state without enrichment of any of these marks, E3: repressive state enriched with H3K9me3, E4: transcriptionally active state enriched with H3K36me3 at gene bodies; E5: enhancer state strongly enriched with enhancer marker H3K27ac and H3K4me1; E6: enhancer state weakly enriched with H3K27ac, and E7: promoter state strongly enriched with H3K4me3 and moderately enriched with H3K27ac (Zhang et al., 2016). HOMER package (Heinz et al., 2010) was utilized to determine transcription factor motifs surrounding TRPS1 binding peaks, BEDOPS (Neph et al., 2012) and GenomicRanges (Lawrence et al., 2013) were utilized in R (http://www.R-project.org) to handle annotated genomic ranges, and ggplot2 (http://ggplot2.tidyverse.org) was used to produce heatmaps and statistical plots. For TRPS1 binding profiling in relation to replication origins, the actually fired origins in MCF-7 cells were obtained from previous massive sequencing of RNA-primed nascent DNAs (Martin et al., 2011; Weddington et al., 2008). These origins were classified into different epigenomic states based on their overlaps with the ChromHMM-identified 7 epigenetic states. HOMER package (Heinz et al., 2010) was used to project the TRPS1 binding and input signals onto the epigenetically classified origins. For false discovery rate (FDR) analysis of TRPS1 binding surrounding H3K9me3-marked origins, Poisson distribution was used to capture the statistical significance of TRPS1 enrichment relative to input signals and FDR method (Benjamini and Hochberg, 1995) was used to adjust *P value*s from multiple comparisons.

Replication timing analysis

For analysis of replication timing of TRPS1-bound DNAs, MCF-7 cells were pulse-incorporated with EdU, sorted into fractions spanning the full DNA synthesis phase, and captured with Dynabeads MyOne Streptavidin C1 (Invitrogen). DNA replication pattern was determined and reads density from the published (Hansen et al., 2010) or our Repli-seq data were plotted within TRPS1 binding summit-centered windows. The background replication efficiency was measured by averaging the Repli-seq reads density over 100 groups of randomly shuffled genomic sites with a similar peak length distribution and the same epigenomic state as the relevant TRPS1 subgroup. The mean and standard derivations of nascent DNA production were calculated. The number of active origins in the windows centered at the same TRPS1 binding peaks but with different sizes ranging from 0.5 Kb to 1 Mb was countered. For assessment of the relative contribution of TRPS1-associated origins stratified by different epigenetic states, the observed number of active origins was normalized to the expected fractions of epigenomic states within TRPS1 binding summit-centered windows.

DNA structural variation analysis

For analysis of TRPS1-associated DNA structural variations, MCF-10A cells were infected with lentiviruses carrying control or wild-type TRPS1, TRPS1C896Y, or TRPS1R908Q. After 14 days of puromycin selection, cells were further cultured for 30-90 days to induce transmissible genome aberrations. Genomic DNAs were extracted and analyzed for copy number alteration using Affymetrix SNP arrays. Copy number gain or loss was determined using hidden Markov model with hapmap270.na35.r1.a5.ref as the reference. The segments larger than 100 Kb and with more than 5 markers were called. Recurrent TRPS1 overexpression-associated CNA sites were defined as ones present in both replicates in wild-type TRPS1-transfected cells but absent from any CNA lists called from the control or TRPS1C896Y- or TRPS1R908Q-transduced cells. TRPS1 overexpression-associated genome structural variations in clinical samples were analyzed by intersecting the candidate genes within experimental CNAs with clinically identified TRPS1 co-amplified/deleted genes retrieved through cBioPortal (Cancer Genome Atlas, 2012; Cerami et al., 2012; Curtis et al., 2012; Gao et al., 2013; Lefebvre et al., 2016; Pereira et al., 2016; Stephens et al., 2012) from 4 large cohorts of breast cancers in METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) and TCGA (The Cancer Genome Atlas).

Extrachromosomal DNA analysis

Extrachromosomal DNA was prepared using HiSpeed Plasmid MidiKit (QIAGEN) and sequenced as described previously with minor modifications (Shibata et al., 2012). Briefly, nuclei were suspended in buffer P1 and lysed with buffer P2. Genomic DNA was precipitated with buffer P3 and removed with QIAfilter Cartridge. DNA in supernatant was loaded onto a HiSpeed Midi Tip, washed with buffer QC and eluted with buffer QF, and then precipitated by isopropanol. After being washed with 70% ethanol and dissolved in buffer EB, DNA pellet was treated sequentially with RNase A and proteinase K, followed by phenol-chloroform extraction and ethanol precipitation. DNA pellet was digested with Exonuclease VIII (NEB) to remove linear DNA and then purified by QIAquick PCR Purification Kit (QIAGEN) to further remove oligonucleotides and DNA of > 10 kb (including mitochondrial DNA). The eccDNA was then amplified by using of REPLI-g Mini Kit (QIAGEN) according to the manufacturer's instruction. Briefly, eccDNAs was denatured with buffer D1 and denaturation was stopped by buffer N1. Reaction mixture containing denatured eccDNA, REPLI-g Mini Reaction Buffer and REPLI-g Mini DNA Polymerase was incubated at 30°C for16 hours for multiple displacement amplification, and then heated at 65°C for 3 min to stop the reaction. The Microcon DNA Fast Flow Centrifugal Filter Unit with Ultracel membrane (Millipore) was used to exchange the buffer. The amplified DNA was then purified and used for library construction followed by high-throughput sequencing. The genomes from TRPS1-overexpressing versus control cells were tiled into 10 Kb windows and the significantly enriched bins (SEBs) or significantly depleted bins (SDBs) were called as ones with reads density ratio greater than 2 (or less than 0.5) and *P value* less than 1e-5 using Poisson test. TRPS1-associated CNAs in extrachromosomes were then analyzed by comparing the number of CNA sites overlapped with at least one SEB (or SDB) to overlaps in null models consisting of 100 groups of randomly shuffled genomic sites with a similar length distribution.

Drug sensitivity test

MCF-10A or MDA-MB-231 cells stably transduced with empty vector or TRPS1 were seeded in 96-well plates overnight prior to the addition of antineoplastic drugs. Drug-free controls were included in each assay. Plates were incubated at 37°C for additional 3 days in a humidified atmosphere with 5% CO_2 followed by cell viability assessment using MTS assays in triplicate experiments. IC_{50} was calculated using GraphPad Prism software and *P value* was calculated using extra sum-of-squares F-test.

Whole genome sequencing and CNA analysis

Drug-resistant cells were established by treatment of MCF-10A cells stably expressing TRPS1 with various chemical drugs for 2 weeks. The drug concentrations used were determined according to the IC_{50} of TRPS1-overexpressing groups measured from drug sensitivity test. Genomic DNAs were extracted and subjected to whole genome sequencing. The copy number variations were analyzed using the CNVnator (Abyzov et al., 2011) read-depth algorithm. The algorithm divides the genome into non-overlapping bins of equal size and uses the count of mapped reads in each bin as the Read-Depth signal. Standard settings with a bin size of 100 bp were used.

Colony formation assay

MCF-10A cells seeded in 6-well plate were maintained in culture media containing a variety of chemical compounds with concentration determined according to the IC_{50} of drug sensitivity test for 2-3 weeks. The resulting cells were fixed with 4% paraformaldehyde before staining with 0.5% crystal violet. The experiments were repeated 3 times and representative images were presented.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis was performed with GraphPad or R software. For t test, two-tailed unpaired parameters were employed unless otherwise indicated.