cell stress

Research Article

www.cell-stress.com

Quantifying replication stress in cancer without proliferation confounding

Philipp Jungk[1] and Maik Kschischo[1,*]

[1] Institute for Computer Science, University of Koblenz, Koblenz, Rhineland-Palatinate, Germany.

* Corresponding Author:
Prof. Dr. Maik Kschischo, E-mail: kschischo@uni-koblenz.de

ABSTRACT Replication stress (RS) is a major driver of genomic instability and cancer development through impaired DNA replication that can lead to chromosomal instability (CIN). Although RS is mechanistically linked to CIN, its relationship with cellular proliferation is complex. Depending on the context, RS can either promote or suppress cell growth. Existing RS gene expression signatures overlook this complexity, relying on the overexpression of oncogenes such as *MYC*, which introduces a proliferation bias. To disentangle genuine RS from confounding cell cycle and proliferation transcriptional profiles, we developed and validated a novel gene expression signature that accurately predicts RS independently of oncogene activity. This tumorigenic RS signature (TRSS) captures RS-related transcriptional changes across diverse cellular contexts, enabling a more robust and proliferation-independent measure of RS in both experimental and clinical samples. Applying our signature to patient data, we discovered a link between RS and the non-homologous end-joining (NHEJ) DNA repair pathway. Specifically, we observed that *MSH2* and *MSH6* – core components of mismatch repair – are associated with elevated RS and may indicate a shift toward NHEJ-mediated repair under stress conditions. Our study provides a refined approach to quantify RS and sheds light on its broader impact on DNA repair network dynamics.

Received originally: 27.05.2025
in revised form: 08.08.2025,
Accepted 26.08.2025,
Published 28.10.2025

Keywords: replication stress, gene expression signatures, cell proliferation, chromosomal instability, non-homologous end-joining, mismatch repair.

Abbreviations:
APH – aphidicolin; **ARSS** – Acute Replication Stress Signature; **CIN** – chromosomal instability; **CPTAC** – Clinical Proteomic Tumor Analysis Consortium; **DSB** – double-strand break; **GSEA** – gene set enrichment analysis; **NHEJ** – non-homologous end-joining; **RS** – replication stress; **RSRD** – RS response defects; **S-CIN** – structural CIN; **SCS** – structural complexity score; **TCGA** – The Cancer Genome Atlas; **TRSS** – Tumorigenic RS Signature.

INTRODUCTION

Replication stress (RS) is a significant contributor to genomic instability and cancer development. It is defined as slowed or stalled DNA replication caused by factors like unusual DNA structures, impaired origin licensing, nucleotide pool depletion and transcription-replication collisions [1]. These disruptions stall or collapse replication forks, leading to under-replicated DNA, double-strand breaks (DSB) and ultimately to structural chromosomal instability (S-CIN). RS also links S-CIN and whole chromosomal instability (W-CIN), describing the continuous gain or loss of entire chromosomes [2], through mitotic errors caused by mechanisms like impaired kinetochore attachment and abnormal spindle assembly [3–5]. Chromosomal instability (CIN), a hallmark of cancers, is thought to promote tumor growth [6] by deregulating pathways via gains in oncogenes (e.g. *CCNE1*, *MYC*) or losses in tumor suppressors (e.g. *PTEN*, *RB1*) [7, 8].

Accurately quantifying RS based on omics-data is essential for linking it to other tumor-related processes, translating findings from biological experiments into primary tumor databases. In cell line experiments RS is commonly induced by drug treatment (e.g. aphidicolin (APH) or hydroxyurea), oncogene overexpression (e.g. *MYC*, or *CCNE1*), or gene knockouts (e.g. *DHX9*) [9–11]. While effective in controlled laboratory settings, each method also alters cellular programs unrelated to RS, skewing phenotypic correlations if examined in isolation. For instance, oncogene overexpression may bias the transcriptome toward genes linked to premature cell cycle progression, potentially obscuring RS-specific effects like the activation of DSB repair pathways in S-phase [1, 12–14]. Despite these limitations, current gene expression-based RS quantifications emphasize oncogene-related phenotypes [15, 16], likely due to their established association to RS in cell lines and frequent natural occurrence in primary tumors [10].

CIN was suggested to paradoxically affect proliferation, enhancing fitness in some context while impairing it in others [6, 17]. Further studies have extended upon these discussions, including evidence of a similar paradoxical relationship between RS and proliferation. For instance, CIN signatures were shown to positively correlate with the proliferation markers like *PCNA* and *MKI67*, as well as proliferation scores in primary tumors, suggesting that genomic instability might drive tumor growth in advanced stages of cancers [17–19]. CIN, assessed through comparative genomic hybridization and image cytometry, has been linked to gene expressions indicating cancer development and cellular growth in breast cancers [20]. In mouse models and murine embryonic fibroblasts, *BubR1*$^{+/-}$ *Apc*$^{Min/+}$ induced CIN led to the development of higher growth rates [21]. Similarly, a causal link to proliferation is suggested by nucleotide depletion in highly proliferative cells inducing RS [22, 23]. However, evidence for this positive association in cell lines is limited. Experiments suggesting that genomic instability increases proliferation required additional defects in cell cycle regulators (e.g. *ATR*, *ATM*), alongside oncogene-induced RS [24, 25]. Inhibiting the DNA damage checkpoint assuring faithfully replicated DNA in oncogene-driven proliferation environments may increase proliferation and overshadowing CIN/RS-dependent effects. Further experiments in cell lines do suggest that CIN and RS generally reduce cellular fitness. For example, even low RS levels induced by 100 nM APH were shown to decrease proliferation in cell lines [26], while increased karyotype heterogeneity in NCI60 cells correlates with slower doubling times [17]. Gene set enrichment analysis (GSEA) of CIN-associated genes in breast cancer cell lines indicated reduced cell cycle activity and a shift of their transcriptome toward a mesenchymal state [17, 27]. Interestingly, some mouse models harboring CIN were shown to reduce proliferation. For instance, *Pten*$^{+/-}$*Bub1*$^{-/H}$ mice exhibited increased abnormal chromosomal numbers, while *PCNA*-positive cells were reduced compared to *Pten*$^{+/-}$ mice with stable chromosome numbers [28]. Another study found a simultaneous decrease in expression of the proliferation marker *Ki67* and an increase in levels of DNA damage as well as aneuploidy in *SA1*-deficient mice [29]. While most studies on the association of genomic instability and proliferation focus on CIN, the limited evidence on RS displaying similar trends, suggests that the CIN-proliferation relationship extends to RS [2, 30–32].

To determine a suitable quantification of RS, we first systematically evaluate existing RS scores and analyze their association with a chromosomal complexity score while accounting for the influence of other cancer hallmarks. Using public gene expression data from various RS-inducing methods, we develop a new RS signature that performs well even on unseen experimental data, while minimizing the confounding effects of oncogene overexpression and capturing the complex relationship between proliferation to RS. Finally, we apply this novel RS signature to large genomic databases to uncover a proteomic link to the non-homologous end-joining pathway, demonstrating its utility for uncovering new insights into cancer biology.

RESULTS

Existing signatures fail to predict RS in cell lines

To estimate the performance of previously published RS signatures, we gathered public RNA-sequencing data from various studies using different methods for RS induction (**Table 1**): Three of the methods used RS-inducing oncogene overexpression (*CCNE1*, *CDC25A*, *MYC*) [33, 34], while three additional methods induced RS by silencing genes affecting RS-pathways (*DHX9*, *CDA*, *SMAD4*) [11, 35]. Samples were divided into those with and without RS.

TABLE 1 ● Samples with gene amplification or knockout to induce replication stress.

Study_ID	*Cell line*	*RS Technique*	*GEO series*
GSE171845	RPE_1	CCNE1	GSE171845
GSE185512	RPE_1	CDC25A, CCNE1, MYC	GSE185512
GSE185512	RPE_1_NOp53	CDC25A, CCNE1, MYC	GSE185512
GSE185512	BT549	CDC25A, CCNE1, MYC	GSE185512
GSE185512	MDA_MB231	CDC25A, CCNE1, MYC	GSE185512
GSE185512	HCC1806	CDC25A, CCNE1, MYC	GSE185512
GSE244103	H446	DHX9_KO	GSE244103
GSE244103	H82	DHX9_KO	GSE244103
GSE244103	H196	DHX9_KO	GSE244103
GSE252544	Mia_PACA_2	CDA_KO	GSE252544
GSE292334	HCT116	SMAD4_KO	GSE292334
GSE292334	HT29	SMAD4_KO	GSE292334

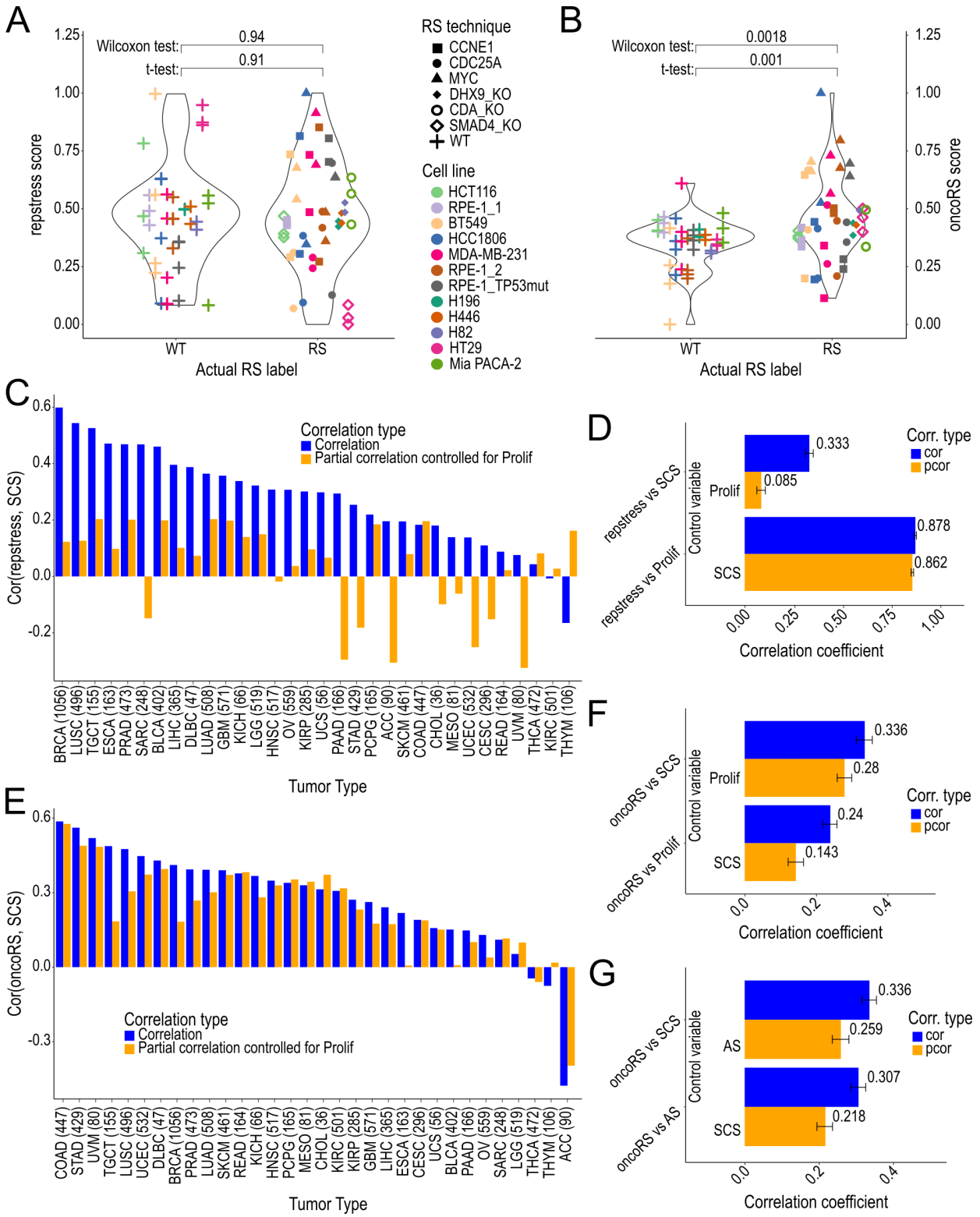


FIGURE 1 ● **Assessment of RS prediction scores and their robustness across multiple studies and associated phenotype scores.** Evaluation of two replication stress scores on a dataset including six different ways of inducing replication stress in multiple cell lines. The y-axis represents the replication stress levels as predicted by the repstress score **(A)** and the oncoRS score **(B)**, while the x-axis distinguishes between the actual status of samples either with replication stress (RS) or without (WT). **(C)** Spearman correlation of the repstress score and the SCS within tumors of TCGA. The y-axis visualizes the non-adjusted (blue) correlation coefficient and the partial (orange) correlation coefficient between repstress score and SCS accounting for proliferation in each tumor type (x-axis). **(D)** Pan-cancer correlation between repstress score and SCS and proliferation with (orange) and without (blue) correction for confounding effects of the annotated control variable using partial correlation. The y-axis shows the comparison pair, while the x-axis represents the (partial) correlation coefficient, which is additionally written next to the bar. **(E)** Spearman correlation of the oncoRS score and the SCS within tumors of TCGA. The y-axis visualizes the non-adjusted (blue) correlation coefficient and the partial (orange) correlation coefficient between oncoRS score and SCS accounting for proliferation in each tumor type (x-axis). **(F, G)** Pan-cancer correlation between oncoRS score and SCS, proliferation **(F)** and aneuploidy **(G)** with (orange) and without (blue) correction for confounding effects of the annotated control variable using partial correlation. The y-axis shows the comparison pair, while the x-axis represents the (partial) correlation coefficient, which is additionally written next to the bar. Error bars in **(D, F, G)** represent the 95% confidence interval. SCS: Structural Complexity Score; TCGA: The Cancer Genome Atlas; AS: Aneuploidy Score; Prolif: proliferation.

We applied two published RS signatures to estimate the RS in the dataset's samples: (i) The *repstress* score based on *MYC* overexpression, DNA damage checkpoint activation and sensitivity of checkpoint inhibitors and (ii) the *oncoRS* score based on oncogene overexpression in cell lines and oncogene amplification in primary tumors [15, 16]. Notably, samples with and without RS-induction displayed no difference in the repstress scores, indicating that the repstress score was not able to predict RS status (**Figure 1A**), while the oncoRS signature was only able to differentiate RS from non-RS in samples with *MYC*-based RS-induction (**Figure 1B**).

To further assess the physiological reliability of these signatures, we used the structural complexity score (SCS), which quantifies the heterogeneity of a sample's copy number landscape to approximate S-CIN [2]. Further, we used a gene expression-based quantification of the proliferation rate [36] and an aneuploidy score based on the absolute number of arm-level alterations of a sample [37].

We evaluated the association between RS scores and SCS in primary tumors of The Cancer Genome Atlas (TCGA) by calculating the Spearman correlation. Using partial correlation, we investigated the confounding of this RS-SCS association by proliferation, potentially resulting from bias toward oncogene and cell cycle activities.

The repstress score displayed a strong correlation with proliferation (Supplementary Figure S1) and a significant association with the SCS within (**Figure 1C**) and across (**Figure 1D**) TCGA tumor types. However, this association with the SCS was reduced significantly when controlling for proliferation. Additionally, more than half of the repstress signature genes are cell cycle genes (9/17, 53%) and most signature genes display a strong correlation to the proliferation score and gene expressions of the proliferation markers *PCNA* and *MKI67* (Supplementary Table S1). In combination this suggests that the repstress score-SCS association is largely mediated by proliferation.

The correlation of the oncoRS score with the SCS remained largely unaffected after accounting for proliferation (**Figure 1E, F**). Since the oncoRS score is partially based on oncogene amplification, we further investigated the signature for confounding by arm-level aneuploidies. A pan-cancer comparison of oncoRS and aneuploidy revealed a similar correlation as with the SCS. Partial correlation showed that the correlation of oncoRS and the S-CIN measure may be influenced by aneuploidies, as it resulted in a mild reduction in the oncoRS-SCS association (**Figure 1G**). Independent validation of the oncoRS signature genes remains limited with *NAT10* being the only gene confirmed to display an association with RS markers in tumors [16].

Moderate APH dosage leads to a non-physiological RS response with an accumulation of cells in S-phase

We aimed to derive a RS signature without intrinsic biases toward proliferation, copy number alterations (CNA), or oncogene-related processes. To achieve this, we expanded the dataset in **Table 1** to include drug-induced RS using APH and hydroxyurea (Supplementary Table S2). In our systematic approach (**Figure 2A**), we first calculated differentially expressed genes in RS samples using a bootstrap approach of the limma-voom pipeline [38, 39]. To further reduce the dimensionality of the selection, we integrated prior biological knowledge by including only genes involved in any processes potentially associated with RS, such as DNA replication and repair among others (Supplementary Table S3). The top-ranked genes were then selected for the signature. In the last step, batch corrected log-expressions of the signature genes were used to train a regularized logistic regression model.

TABLE 2 ● Genes in Acute Replication Stress Signature (ARSS) including their weights.

	Gene	coef
1	TFDP2	2.596
2	SF3A3	2.518
3	BTG2	2.013
4	RADX	1.063
5	BRME1	1.048
6	POLH	0.930
7	SUPT7L	0.575
8	RRM2B	0.534
9	CEP164	0.424
10	SUN2	0.259
11	REV3L	0.254
12	AOC2	0.193
13	PLK1	-0.039
14	PPP2R1A	-0.097
15	ERCC1	-0.097
16	VCL	-0.126
17	MCM6	-0.372
18	HINFP	-0.421
19	H2BC14	-0.728
20	NUCKS1	-0.861
21	POLR1D	-1.527
22	ERCC2	-1.924
23	FBXL7	-2.338
24	SPIDR	-2.440
25	PIF1	-2.577
26	H2AC14	-2.885
27	AKT3	-3.735
28	PPP2R5A	-4.588

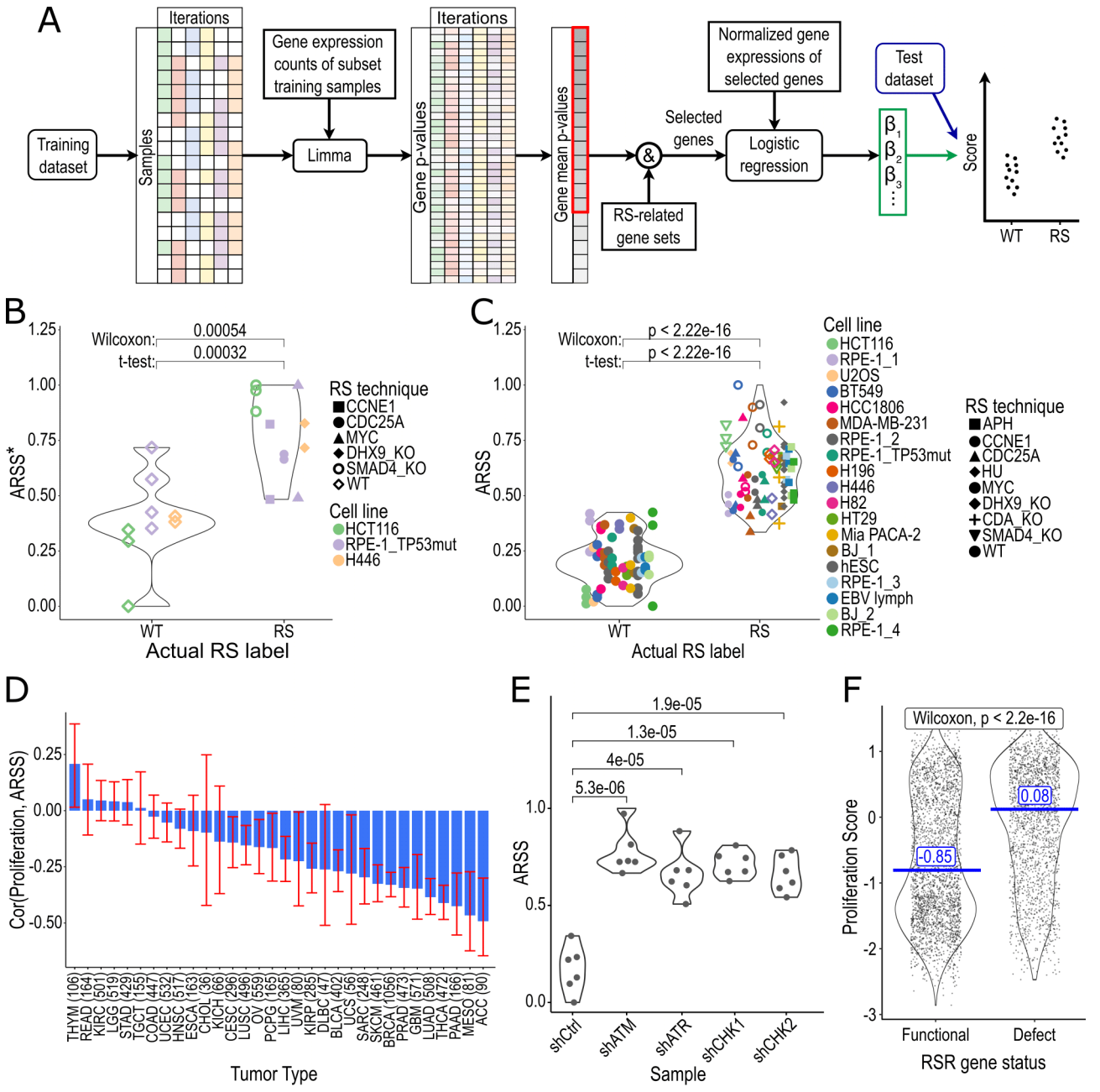


FIGURE 2 ● **Development and interpretation of a signature of acute replication stress.** **(A)** Workflow of the RS prediction model development. Estimating consistently differentially expressed genes using a permutation approach, limma-based differential gene expression determination and calculating gene weights using regularized logistic regression. **(B)** Comparison of the RS scores predicted by a signature for acute RS (ARSS⋆) between samples with (RS) and without (WT) replication stress on an unseen test data set. **(C)** Predictions of the ARSS on the entire dataset combining eleven studies and eight RS induction methods compared to the actual RS status grouped into samples with (RS) and without (WT) replication stress. **(D)** Spearman correlation between the ARSS and a proliferative score (y-axis) in different TCGA tumor types (x-axis). **(E)** Acute RS score in samples with functional RS response (shCtrl) and with silenced gene expression of RS response genes (shATM, shATR, shCHK1, shCHK2). **(F)** Comparison of proliferation between TCGA samples with (Defect) and without (Functional) mutations or copy number losses in RS response genes (ATR, ATM, CHEK1, CHEK2). RS: Replication stress; ARSS: Acute replication stress signature; TCGA: The Cancer Genome Atlas; RSR: Replication stress response

To assess the modeling approach and selected hyperparameters, we applied this signature – referred to as Acute Replication Stress Signature (ARSS⋆) – to a set of unseen test data. ARSS⋆ scores were computed by a linear combination of the model coefficients and the expression values of the signature genes, which resulted in a reliable prediction of RS status (Figure 2B).

After testing the performance on unseen data, we retrained the model on the full data set to leverage all available transcriptomic data for downstream analysis. We refer to this final model including all the updated coefficients as ARSS (Table 2) and used it for further analysis. The distribution of the ARSS across all samples (including training samples) is shown in Figure 2C.

Previously, using moderate APH dosages over 100 nM to induce RS has been criticized for halting cell cycle progression through checkpoint activation rather than a physiological response to RS [26]. To assess whether the ARSS, which uses training samples with APH levels of 200-400 nM, is physiologically feasible in primary tumors, we compared its association to proliferation. As displayed in **Figure 2D**, ARSS is negatively correlated with proliferation in TCGA tumor types, contrasting the previous observation of increased proliferation in tumors with RS and genomic instability [20]. To investigate the meaning of the ARSS, we used a dataset with defects in the RS response genes *CHK1/2*, *ATR*, *ATM* [24]. Notably, samples with RS response defects (RSRD) displayed significantly higher levels of ARSS (**Figure 2E**). Defects in checkpoint genes are associated with an increase in proliferation in TCGA primary tumors, as indicated by higher proliferation rates in samples with heterozygous RSRD (**Figure 2F**), and in cell line experiments with silenced RS-response genes [24, 25]. However, disruption of the cell cycle by hydroxyurea was reported to expose an inability to recover from stalled replication forks [24]. This suggests a high level of stress in these RSRD samples that might resemble samples with RS induced by moderate APH dosages as used in ARSS generation.

We performed Gene Set Enrichment Analysis [40] to identify pathways associated with both the ARSS scores and RSRD status. Samples were classified as RSRD if they contained a mutation or copy number loss in one of the RS response genes listed above. Common enrichment was observed in the *TP53* pathway ($q < 0.01$ for both signatures) and the gene set for downregulation of the *KRAS* signaling pathway ($q < 0.03$ for both signatures) [41]. Both describe the inhibition of cell cycle progression at the G1/S checkpoint [42, 43], which aligns with the observation of inhibited cell cycle progression in samples

treated with moderate APH dosages [26]. These results suggest that the samples with 200-400 nM APH treatment drive the ARSS to reflect conditions of severe RS leading to cell cycle arrest instead of tumorigenic RS.

A novel RS signature highlights elevated proliferative cell cycle activity in primary tumors versus cell lines

To capture physiological, tumorigenic RS, we developed a second signature by excluding samples subjected to severe RS conditions, such as those treated with high concentrations of APH and hydroxyurea from the dataset (**Table 1**). This signature was trained using the same methodology (**Figure 2A**) on the reduced dataset, resulting in the Tumorigenic RS Signature (TRSS$\star$). Applying the TRSS$\star$ to a held-out test set confirmed that the methodology was still able to differentiate RS and non-RS samples (**Figure 3A**). The final model was re-trained on the full dataset to obtain the final coefficients (**Table 3**). We will refer to this re-trained model as the TRSS. For comparison with **Figure 3A**, the distribution of the TRSS across all samples (including training samples) is shown in **Figure 3B**.

Next, we applied the TRSS to the aphidicolin-treated samples with the lowest dosage available in our dataset (**Table 1**). The TRSS retained its predictive capacity at the dosage of 200 nM APH (**Figure 3C**), while the two existing scores were again not able to predict treatment status (Supplementary Figure S1C and D). Further we found that molecular subtypes in gastrointestinal cancers (COAD, ESCA, READ, STAD) and UCEC tumors in the TCGA database associated with CIN or increased copy number alterations tend to display a higher TRSS than MSI or low CNA subtypes (**Figure 3D**; Supplementary Figure S2A).

To validate the biological relevance of this second RS signature, we examined its association to the SCS in TCGA tumors, while also testing whether this association was dependent on proliferation or aneuploidy using partial correlation. While proliferation had a minor effect on the association (**Figure 3E**), the confounding by aneuploidy was negligible (**Figure 3F**). Moreover, we found that the TRSS generally exhibited a weak correlation with proliferation (cor: 0.1) and aneuploidy (cor: 0.07), suggesting the TRSS captures an aspect of genomic instability independent of proliferative capacity or extensive, stable chromosomal changes.

TABLE 3 ● Genes in Tumorigenic Replication Stress Signature (TRSS) including their weights.

	Gene	coef
1	ARID2	3.252
2	H4C8	2.956
3	TOP3A	2.630
4	TAF12	2.358
5	AOC2	2.063
6	BTG2	1.906
7	TFDP2	1.848
8	PHF8	1.729
9	ATF2	1.591
10	H2BC12	1.415
11	H2AC6	0.235
12	GLI1	-0.085
13	POLD4	-0.144
14	TUBB	-0.390
15	BCAM	-0.793
16	H4C14	-1.026
17	UFL1	-1.210
18	CMPK2	-1.329
19	FBXL7	-1.626
20	TP53	-2.693
21	PPP2R5A	-3.138

In an additional analysis we included MSigDB hallmark gene set activities as covariates when modeling the association of TRSS and SCS. No single pathway accounted for a major portion of the association (Supplementary Figure S2B; Supplementary Methods).

Next, we investigated whether the association between TRSS and proliferation might vary depending on biological context by examining differences in cell cycle activity between primary tumors and cell lines. Using GSEA, we calculated the enrichment of cell cycle-related gene sets from the gene ontology biological process [44, 45] and hallmark collections [41] with respect to TRSS and SCS (**Figure 4A**; Supplementary Table S4). Interestingly, for both scores we found a significant upregulation of various cell cycle gene sets in TCGA primary tumors, whereas this upregulation was not observable in cell lines from the Cancer Cell Line Encyclopedia (CCLE). On the contrary, oncoRS and repstress scores displayed the same association with these cell cycle gene sets in cell lines and primary tumors. We asked if this enrichment implied a cell cycle-related increase of proliferation by examining the deregulation of leading-edge genes associated with the TRSS with a proliferation score and *PCNA* and *MKI67*. Indeed, in TCGA, leading-edge genes of the enriched pathways displayed a higher correlation with the three proliferation measures than non-leading-edge genes as determined by a two-sample Wilcoxon test (all p < 2.2e-16; Supplementary Figure S3A).

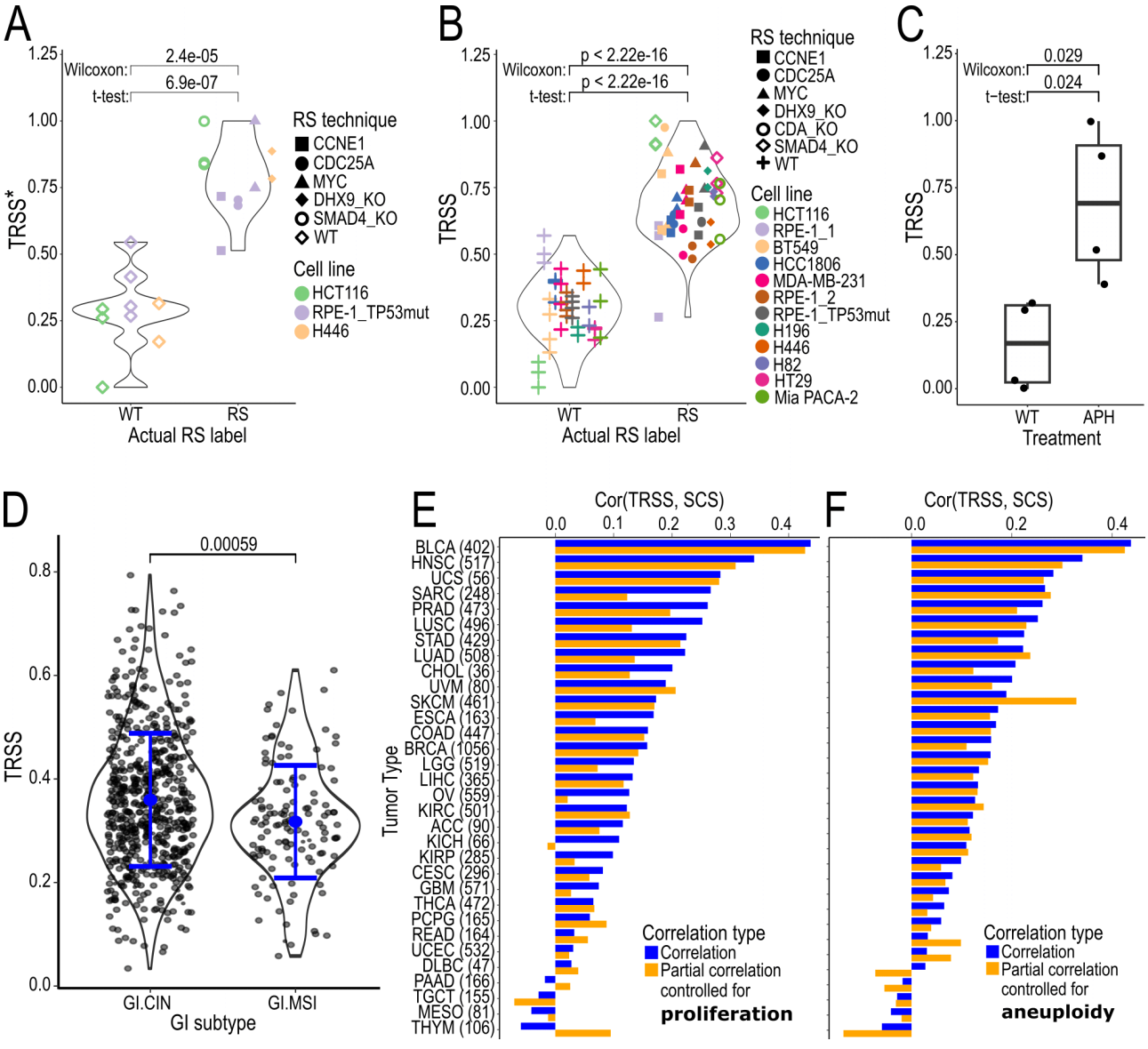


FIGURE 3 ● **Validation and application of a novel RS signature for tumorigenic RS not confounded by proliferation and aneuploidy.** **(A)** Comparison of the RS scores predicted by a signature for tumorigenic RS (TRSS*) between samples with (RS) and without (WT) replication stress on an unseen test data set. **(B)** Predictions of the TRSS on the entire dataset combining eleven studies and eight RS induction methods compared to the actual RS status grouped into samples with (RS) and without (WT) replication stress. **(C)** Comparison of TRSS between aphidicolin-treated (200 nM; APH) and untreated (WT) samples. **(D)** TRSS distribution across CIN and MSI molecular subtypes in TCGA gastrointestinal cancers. **(E, F)** Spearman correlation of the TRSS and the SCS within tumors of TCGA. The y-axis visualizes the non-adjusted (blue) correlation coefficient and the partial (orange) correlation coefficient between repstress score and SCS accounting for proliferation **(E)** and aneuploidy **(F)** in each tumor type (x-axis). RS: Replication Stress; TRSS: Tumorigenic replication stress signature; TCGA: The Cancer Genome Atlas SCS: Structural Complexity Score

In cell lines on the other hand, the leading-edge genes indicated a reduction in cell cycle progression (all p < 2.2e-16; Supplementary Figure S3B). Thus, the TRSS captures the context-dependent association to proliferation, indicating an increase in cell cycle activity promoting proliferation in primary tumors, but not in cell lines.

Prognostic and therapeutic associations of the TRSS

To assess the clinical relevance of the TRSS we first examined its association with overall survival across TCGA tumors. Notably, we did not observe a significant difference in survival between patients in the lowest (Q1) and highest (Q4) TRSS quartiles (Supplementary Figure S3C), suggesting that the TRSS does not independently stratify patient prognosis in this pan-cancer cohort.

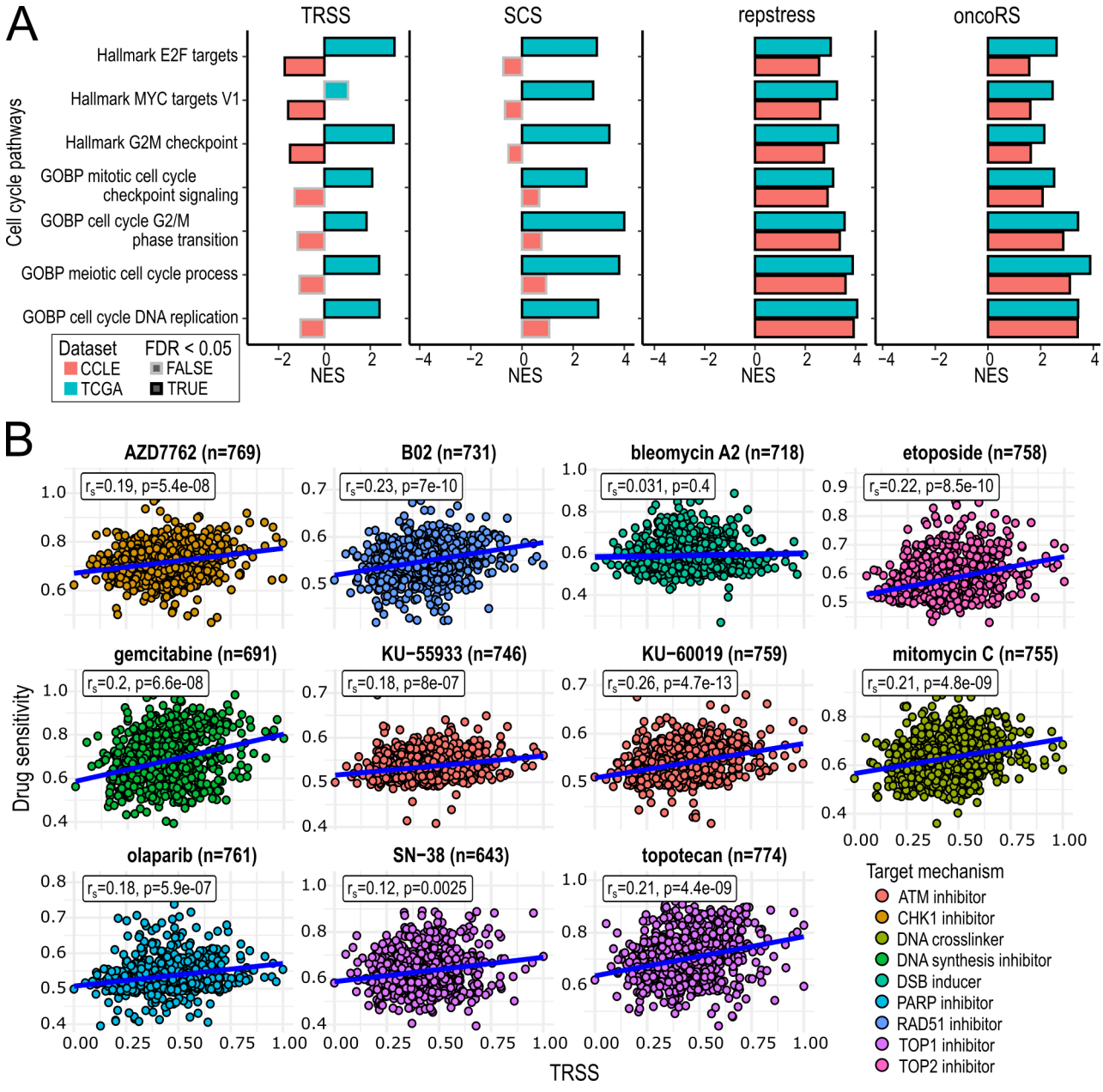


FIGURE 4 ● TRSS reflects context-dependent proliferative activity and predicts sensitivity to agents targeting faithful DNA replication. (A) Normalized enrichment score (NES) for gene sets representing distinct stages of the cell cycle shown in relation to the TRSS, and SCS, as well as the oncoRS, and repstress scores across TCGA (blue) and CCLE (red). Enrichment scores were calculated using gene set enrichment analysis. (B) Faceted scatterplots displaying the association of TRSS and drug sensitivity of eleven compounds available in the CTRP targeting DNA repair pathways or inducing DNA damage. RS: Replication Stress; TRSS: Tumorigenic replication stress signature; TCGA: The Cancer Genome Atlas; SCS: Structural Complexity Score; CCLE: Cancer cell line encyclopedia; NES: Normalized enrichment score; CTRP: Cancer Therapeutics Response Portal.

To evaluate whether the TRSS reflects sensitivity to agents that interfere with faithful DNA replication, we selected compounds known to target vulnerabilities induced by RS [46, 47] and analyzed the TRSS-associated drug sensitivity using data from the Cancer Therapeutics Response Portal (CTRP) [48–50]. The selected drugs included inhibitors targeting key components of the DNA damage response such as ATM inhibitors (KU-55933 and KU-60019), a CHK1 inhibitor (AZD7762) and a PARP inhibitor (olaparib), as well as compounds inducing DNA damage, including topoisomerase inhibitors (topotecan, etoposide), and a DNA crosslinker (mitomycin C).

We observed a modest but consistent increase in sensitivity to these agents in samples with higher TRSS levels (**Figure 4B**). An analysis of point mutations and copy number alterations further revealed that the TRSS is associated with alterations in genes involved in DNA repair, replication fork protection, and the TP53 pathway (Supplemental Material).

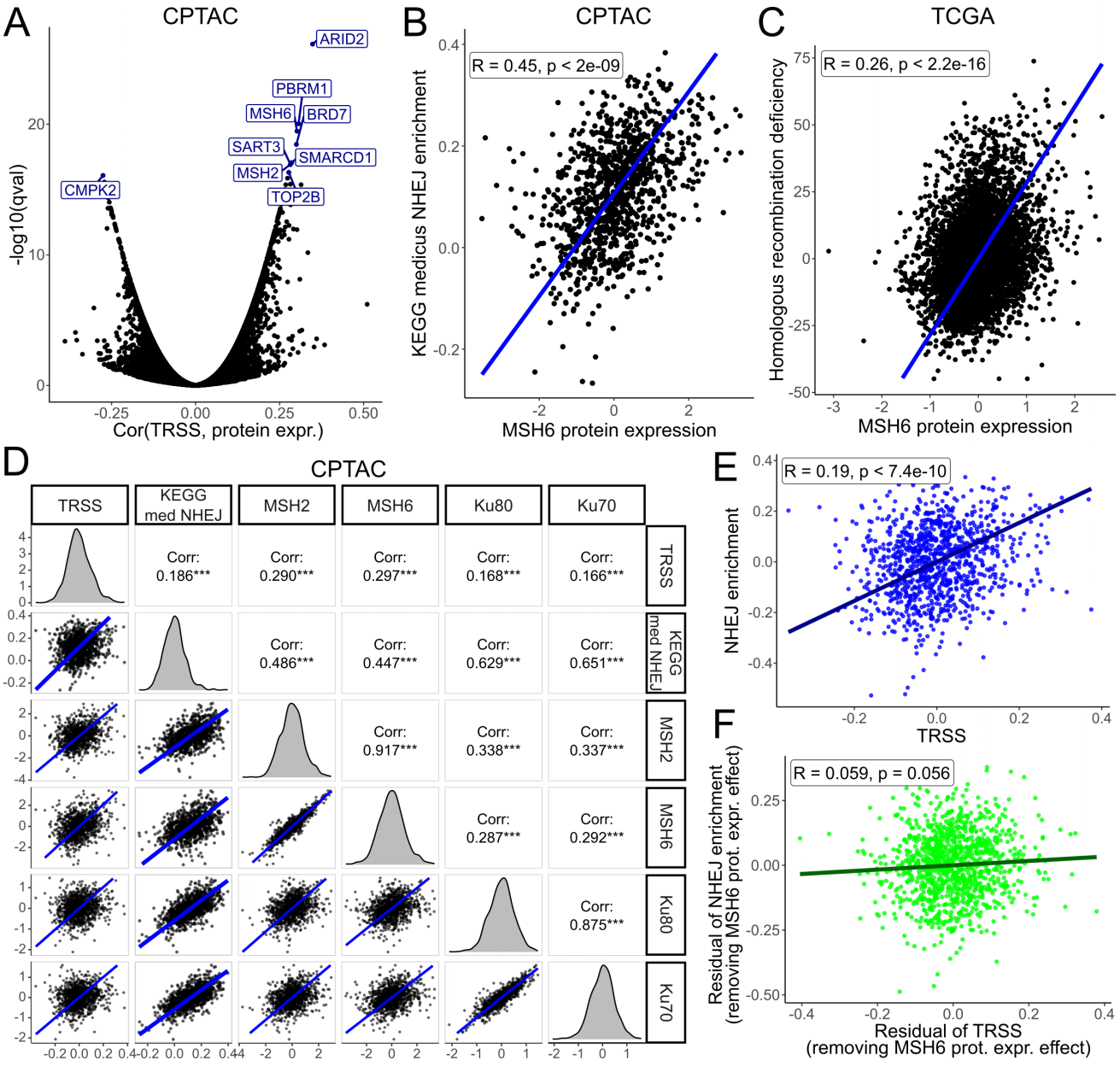


FIGURE 5 ● MSH6 protein expression is positively associated with the TRSS and NHEJ, potentially leading to low fidelity DNA repair and genomic scars resembling homologous recombination deficiency. **(A)** Volcano plot visualizing the spearman correlation and q-values of CPTAC protein expressions to the TRSS. **(B)** Activity in the KEGG medicus NHEJ gene set compared to the protein expression of MSH6 in CPTAC. **(C)** Amount of genomic scars typical for homologous recombination deficiencies compared to the protein expression of MSH6 in CPTAC. **(D)** A comparison of TRSS and NHEJ enrichment to protein expression of the proteins in the MSH2/6 complex and the Ku heterodimer in CPTAC. **(E)** Activity in the KEGG medicus NHEJ gene set compared to the TRSS in CPTAC. **(F)** Activity in the KEGG medicus NHEJ gene set compared to the TRSS after adjusting for the effects of MSH6 expression in CPTAC. TRSS: Tumorigenic replication stress signature; CPTAC: Clinical Proteomic Tumor Analysis Consortium; NHEJ: Non-homologous end-joining; TCGA: The Cancer Genome Atlas; KEGG: Kyoto Encyclopedia of Genes and Genomes

These observations show that TRSS levels correspond to DNA repair pathway alterations and a sensitivity to RS-targeting compounds.

Tumorigenic RS signature reveals MutSα-mediated regulation of double-stranded DNA recombination in proteomics data

A recent study in yeast suggests that cellular stress can alter mRNA translation patterns [51]. Using protein expression data from the Clinical Proteomic Tumor Analysis Consortium (CPTAC) database, we investigated proteomic effects of RS. Apart from ARID2, which is part of the TRSS signature (**Table 3**), two of the proteins most significantly correlated with the TRSS were *MSH6* and *MSH2*, which form the DNA mismatch repair complex MutSα (**Figure 5A**). *MSH6* protein expression and TRSS exhibited a Spearman correlation of 0.3, while *MSH2* protein expression displayed a correlation of 0.28 (q-value < 2e-14). Interestingly, no other MMR proteins demonstrated a comparably strong association with the TRSS (Supplementary Table S5). The two MutSα proteins were also among the top-ranking proteins in the TCGA proteomics dataset (both q-values < 2e-14), although it only includes about 300 (phospho-)proteins (Supplementary Figure S4A).

While sensing mismatches is the primary function of MutSα, its components are also implicated in the DSB repair pathways of non-homologues end-joining (NHEJ) and homologous recombination [52–54]. Correlation analyses revealed that *MSH2/6* protein expression was significantly associated with single sample GSEA enrichment scores [55] in the KEGG medicus reference NHEJ gene set [56] (**Figure 5B**, Supplementary Figure S4B). *MSH2/6* also correlated with a score quantifying genomic scars resulting from homologous recombination deficiency [57] in TCGA, indicating low fidelity DSB repair (**Figure 5C**, Supplementary Figure S4C). Further analysis in CPTAC demonstrated that MutSα components as well as the TRSS were associated with Ku70 and Ku80 (**Figure 5D**), crucial components of the double-strand break repair via NHEJ [58]. The relationship of MutSα and NHEJ also translated into an association between TRSS and NHEJ in CPTAC (**Figure 5E**, Supplementary Figure S4D). Interestingly, upon partial correlation to control for *MSH2/6* protein expression the association between TRSS and NHEJ activity was removed, suggesting that the association might be mediated by MutSα (**Figure 5F**, Supplementary Figure S4E).

DISCUSSION

This study sought to evaluate the proficiencies and limitations of existing RS gene expression signatures and improve upon these limitations to elucidate new insights into the role of RS in tumorigenesis.

Using a dataset of labeled samples from the gene expression omnibus, we found that both the repstress score as well as the oncoRS score fail to reliably differentiate RS samples. The repstress score demonstrated an exaggerated dependence on genes associated with cell cycle progression. This is attributable to the selection of *MYC*-associated genes from the G2M-checkpoint and E2F target hallmark pathways, as supported by the strong association to a proliferation score, which also decouples the expected association between the repstress score and a measure for structural aneuploidies. While there are contexts in which proliferation drives RS, for example by causing nucleotide deficiency [22, 23], RS can also induce cell cycle arrest, reducing proliferation [59]. Moreover, recent studies suggest that RS may also be reduced as a result of adaptation to aneuploidy, which is accompanied by an increase in proliferation [60, 61].

The oncoRS score is only targeted at identifying genes associated with oncogene-induced RS, as acknowledged by the original authors, likely limiting it from generalizing to RS originating from other sources. The signature may instead overemphasize other oncogene-specific processes like cell cycle regulation as essential causes of RS [1, 12–14]. Methodologically, the score lacks validation of the entire signature, as only one signature gene is explored further. While the present study establishes a link between the oncoRS score and the SCS, both the signature and the SCS rely on copy number alterations in TCGA tumors. The reduced association when controlling for arm-level aneuploidies might indicate that the signature captures aneuploidy-related gene expression, rather than directly measuring RS related to CIN.

Our first alternative approach leveraged oncogene, gene knockout and drug-based induction of RS to detect transcriptional changes and define the ARSS. However, our analysis suggests that the samples with high APH concentrations might contain non-physiological levels of RS, which manifest itself through increased levels of cell cycle arrest as evidenced by the lower proliferation rate in samples with high ARSS. While a slight increase in checkpoint activation might be expected upon DNA damage resulting from RS, the APH samples might represent a gene expression profile of a checkpoint response to excessive DNA RS and DNA damage [26]. The good predictive ability on samples with defective RS response indicates that RSRD samples are subjected to large stresses, as is evidenced by McGrail *et al.*'s observation that RSRD cells are unable to recover from disruption through HU treatment and accumulate in S-phase [24].

The TRSS was therefore designed to minimize these non-physiological biases by excluding samples that might display exaggerated RS. The association of TRSS and SCS across tumor types, independent of proliferation and aneuploidies, suggests that the TRSS reliably captures tumorigenic RS that goes hand in hand with structural complexity. While the TRSS shows a slightly weaker association with the SCS compared to previous scores, it avoids key confounders and is more generalizable. Specifically, the repstress score is heavily influenced by proliferation, while the oncoRS signature was partially trained on oncogene amplification, which overlaps with the SCS defini-

tion. In contrast, the TRSS captures broader RS-related features, potentially at the cost of reduced signal in specific contexts.

Applying the TRSS to large datasets, we observed differences in proliferative cell cycle activity between primary tumors and cell lines associated with the TRSS. This finding is consistent with findings summarized in the introduction, where primary tumors generally exhibit increased proliferation and tumor growth, while cell lines demonstrate reduced cell viability and proliferation upon induction of RS and CIN. These results underscore the ability of the TRSS to capture proliferation-independent RS, despite being trained solely on cell line data and not explicitly optimized to account for context-specific proliferation effects.

Considering this additional evidence for a significant difference in RS-related proliferation between primary tumors and cell lines, a key question arises: Why does RS appear to impair long-term proliferation in cell lines but not in primary tumors? One key difference is the tumor microenvironment, which may be shaped to support the proliferation associated with RS and CIN by (i) reducing the function of T-cells [62], (ii) recruiting pro-tumor M2 macrophages [63, 64], and (iii) signaling cancer-associated fibroblasts to provide energy-rich metabolites [65].

In TCGA, the TRSS was not associated with overall patient survival. A possible explanation for this could be the aneuploid adaptation mechanism associated with aggressive proliferation and hazardous chromosome arm alterations [60, 61]. This adaptation is also accompanied by a reduction in genomic instability, including RS, which may mask the impact of RS on patient survival. However, the TRSS identified a vulnerability to compounds inhibiting faithful DNA replication. The observed

correlations were consistent across multiple agents and are in line with the known role of RS in generating DNA lesions, which increase cellular dependence on intact DNA repair pathways [47]. At the same time, agents introducing additional DSBs may overwhelm the DNA damage response. Given the interconnected nature of DNA repair pathways, a stronger response might be achieved by combinatorial treatments, which might also prevent adaptation to single target therapies [46].

Leveraging protein expressions, we found the *MSH2-6* protein complex to be one of the DNA repair components that is most strongly associated with the TRSS. Our findings suggest that this complex connects RS to an increase in NHEJ activity, potentially exacerbating CIN through low fidelity DNA damage repair. This is consistent with broader understanding that DSBs are generally resolved by two main pathways, HR and NHEJ [66]. In context of HR deficient cells, which have been shown to also display reduced replication fork speeds [67, 68], NHEJ likely becomes the compensatory repair mechanism for cell survival. Recent findings support this by showing that depleting *LIG4*, a central NHEJ factor, in an RS context resulted in an increase in apoptotic areas, while *BRCA2* inhibition only resulted in a minor effect [69].

Importantly, it has also been shown that *MSH6* seems to modulate NHEJ activity directly with experimental evidence demonstrating that inhibition of *MSH6* leads to decreased NHEJ activity [54]. This relationship may be mediated via *MSH6*'s interaction with *Ku70* observed in both HeLa cells [54] and oocytes [70].

Extending this connection, Zhang et al. proposed that under RS conditions MutSα may block fork protecting factors causing DNA breaks and CIN upon misincorporations [71]. Another study found that prostate cancers with increased MMR gene expression, including *MSH6*, displayed a higher incidence of deletions at different chromosome loci [72]. Although the extent to which this RS-related NHEJ-activity is linked to CIN requires further investigation, these findings suggest that the TRSS will provide new insights into the role of RS in cancer.

MATERIALS AND METHODS

Development of the gene expression signature

1. Data quality and Processing of sequenced reads: We checked the RNA sequencing reads and removed any low-quality parts using FASTQC and Trimmomatic. We then quantified the gene expression counts using kallisto.

2. Normalization and batch correction: We shrinked variances toward a mean-variance trend, log2-normalized the counts, and corrected for batch effects of different study origins and cell lines.

3. Training and Testing Sets: We split the dataset into training and test data according to **Table 1** and Supplementary Table S2, while ensuring independent experiments by stratifying by treatment and cell line.

4. Identify most statistically RS-associated genes: We repeatedly (*n* times) selected a subset of *p*%. For each subset, we identified genes that changed in response to treatment. This differential gene expression was performed using the limma-voom pipeline [38, 39]. After these permutations we calculated the mean rank of the p-values of each gene.

5. Gene Selection: We selected the top *r* genes that were also assigned to RS-adjacent gene sets (Supplementary Table S3).

6. Train and validate model: Lastly, we used the selected genes to train a regularized logistic regression model with the elastic net coefficient $\boldsymbol{\alpha}$. We validated this approach using the unseen test dataset.

The hyperparameters (*n*, *p*, *r*, and $\boldsymbol{\alpha}$) as well as the choice of the mean rank for pre-selection were determined using 10-fold cross validation with the area under the receiver operating characteristic curve as performance metric.

Calculation of RS scores

The repstress [15], the oncoRS [16] and the RS scores in this article were calculated as weighted linear combination of gene expression values. All gene expression data were first normalized using the variance-stabilizing transformation (vst) function from the DESeq2 package [73] in the R programming language. To control for technical variability, batch effects were removed using the removeBatchEffect function from the limma

package [38]. Each RS score was calculated by multiplying the transform, batch-corrected expression values by a set of predefined weights:

$$score = w_1 \cdot x_1 + w_2 \cdot x_2 + \cdots + w_n \cdot x_n$$

where $x_1, x_2, \ldots, x_n$ are the normalized expression values for the respective signature genes, and $w_1, w_2 \ldots, w_n$ their corresponding weights.

- For oncoRS, all signature genes were weighted equally ($w_i = 1$).
- For repstress, we used the gene-specific weights defined in (Supplementary Table S1) of the original publication.
- For the ARSS and TRSS, we used the gene weights listed in Tables 2 and 3, respectively.

Statistical tests

Correlations were calculated using spearman correlation. Group differences were calculated using Wilcoxon rank-sum tests unless specified otherwise. Where applicable, p-values were adjusted for multiple testing using the Benjamini-Hochberg method.

Gene Set Enrichment Analysis

Differentially regulated pathways associated with gene signatures were determined using Gene Set Enrichment Analysis using the fgsea package on differentially expressed genes identified via the limma-voom pipeline [38, 39]. Single-sample GSEA was conducted using the gsva package to estimate gene set activity.

Clinical analysis

To focus on intra-tumor effects, TRSS values were mean-centered within each tumor type to adjust for baseline differences. Samples in the first and fourth quartile were classified as low and high TRSS samples, respectively. The effect of RS status on survival was compared using a Kaplan-Meier analysis with the survival and survminer R packages.

For the analysis of the association of TRSS and drug response, area under the curve (AUC) values from the CTRP were min-max scaled to [0,1]. Drug sensitivity was then defined as 1-scaledAUC.

Materials

The accession numbers for gene expressions used for generation and validation of RS signatures are listed in Table 1 & Supplementary Table S2 (accessed: 25. March 2024). Gene expression with RS samples with response defects are available at GSE59227.

Gene sets used were obtained from MsigDB [40, 41]: KEGG medicus gene set (MSigdb v2023.2) [56]; Gene Ontology [44, 45] and Hallmark (MSigdb v2023.1) [41]

Omics-data

Gene expression, DNA segments, mutational, proteomic, and clinical data in primary tumors generated by The Cancer Genome Atlas (TCGA) Research Network https://www.cancer.gov/tcga [74] were retrieved from the GDC data portal https://portal.gdc.cancer.gov/ (accessed: 03. July 2023).

Gene expression, and proteomic data in primary tumors generated by the National Cancer Institute Clinical Proteomic Tumor Analysis Consortium were retrieved from the PDC portal [75] https://proteomic.datacommons.cancer.gov/pdc/cptac-pancancer (accessed: 10. August 2024).

Gene expression data in cell lines compiled in the Cancer cell line encyclopedia [76] was retrieved through the DepMap data portal https://depmap.org/portal (accessed: 23. August 2023).

TCGA derivative data:

SCS: Distinct and Common Features of Numerical and Structural across Different Cancer Types [18].

Proliferation rate, HRD score, aneuploidy score: The immune landscape of cancer [77] (accessed: 01. July 2024).

Amplification and Deletion calls: Oncogenic Signaling Pathways in The Cancer Genome Atlas [78] (accessed: 27. June 2024).

CCLE derivative data:

We retrieved cell line drug sensitivities from the Cancer Therapeutics Response Portal (CTRP) version 2 through the National Cancer Institute's data portal https://ctd2-data.nci.nih.gov/Public/Broad/CTRPv2.0_2015_ctd2_ExpandedDataset/ (accessed: 23. August 2023).

ACKNOWLEDGMENTS

The authors thank Holger Bastians and Atmika Paul for providing additional RNA sequencing data for SMAD4 knockout-induced replication stress.

This work was funded by the German Research Foundation (FOR2800; Deutsche Forschungsgemeinschaft).

SUPPLEMENTAL MATERIAL

All supplemental data for this article are available online at www.cell-stress.com.

CONFLICT OF INTEREST

The authors declare no competing interests.

COPYRIGHT

© 2025 Jungk and Kschischo. This is an open-access article released under the terms of the Creative Commons Attribution (CC BY) license, which allows the unrestricted use, distribution, and reproduction in any medium, provided the original author and source are acknowledged.

Please cite this article as: Philipp Jungk and Maik Kschischo (2025). Quantifying replication stress in cancer without proliferation confounding. **Cell Stress** 9: 201-215. doi: 10.15698/cst2025.10.312

REFERENCES

1. Primo LMF, and Teixeira LK (**2020**). DNA replication stress: oncogenes in the spotlight. **Genet Mol Biol** 43(1 Suppl 1): e20190138. doi: 10.1590/1678-4685-GMB-2019-0138

2. Burrell RA, McClelland SE, Endesfelder D, Groth P, Weller M-C, Shaikh N, Domingo E, Kanu N, Dewhurst SM, Gronroos E, Chew SK, Rowan AJ, Schenk A, Sheffer M, Howell M, Kschischo M, Behrens A, Helleday T, Bartek J, Tomlinson IP, and Swanton C (**2013**). Replication stress links structural and numerical cancer chromosomal instability. **Nature** 494(7438): 492–496. doi: 10.1038/nature11935

3. Wilhelm T, Olziersky A-M, Harry D, De Sousa F, Vassal H, Eskat A, and Meraldi P (**2019**). Mild replication stress causes chromosome mis-segregation via premature centriole disengagement. **Nat Commun** 10: 3585. doi: 10.1038/s41467-019-11584-0

4. Böhly N, Schmidt A-K, Zhang X, Slusarenko BO, Hennecke M, Kschischo M, and Bastians H (**2022**). Increased replication origin firing links replication stress to whole chromosomal instability in human cancer. **Cell Rep** 41(11): 111836. doi: 10.1016/j.celrep.2022.111836

5. Ertych N, Stolz A, Stenzinger A, Weichert W, Kaulfuß S, Burfeind P, Aigner A, Wordeman L, and Bastians H (**2014**). Increased microtubule assembly rates mediate chromosomal instability in colorectal cancer cells. **Nat Cell Biol** 16(8): 779–791. doi: 10.1038/ncb2994

6. Bakhoum SF, and Cantley LC (**2018**). The Multifaceted Role of Chromosomal Instability in Cancer and Its Microenvironment. **Cell** 174(6): 1347–1360. doi: 10.1016/j.cell.2018.08.027

7. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, Lawrence MS, Zhang C-Z, Wala J, Mermel CH, Sougnez C, Gabriel SB, Hernandez B, Shen H, Laird PW, Getz G, Meyerson M, and Beroukhim R (**2013**). Pan-cancer patterns of somatic copy number alteration. **Nat Genet** 45(10): 1134–1140. doi: 10.1038/ng.2760

8. Marshall AE, Roes MV, Passos DT, DeWeerd MC, Chaikovsky AC, Sage J, Howlett CJ, and Dick FA (**2019**). RB1 Deletion in Retinoblastoma Protein Pathway-Disrupted Cells Results in DNA Damage and Cancer Progression. **Mol Cell Biol** 39(16): e00105-19. doi: 10.1128/MCB.00105-19

9. Vesela E, Chroma K, Turi Z, and Mistrik M (**2017**). Common Chemical Inductors of Replication Stress: Focus on Cell-Based Studies. **Biomolecules** 7(1): 19. doi: 10.3390/biom7010019

10. Guerrero Llobet S, van der Vegt B, Jongeneel E, Bense RD, Zwager MC, Schröder CP, Everts M, Fehrmann RSN, de Bock GH, and van Vugt MATM (**2020**). Cyclin E expression is associated with high levels of replication stress in triple-negative breast cancer. **NPJ Breast Cancer** 6: 40. doi: 10.1038/s41523-020-00181-w

11. Murayama T, Nakayama J, Jiang X, Miyata K, Morris AD, Cai KQ, Prasad RM, Ma X, Efimov A, Belani N, Gerstein ER, Tan Y, Zhou Y, Kim W, Maruyama R, Campbell KS, Chen L, Yang Y, Balachandran S, and Cañadas I (**2024**). Targeting DHX9 Triggers Tumor-Intrinsic Interferon Response and Replication Stress in Small Cell Lung Cancer. **Cancer Discov** 14(3): 468–491. doi: 10.1158/2159-8290.CD-23-0486

12. Nussinov R, Tsai C-J, Muratcioglu S, Jang H, Gursoy A, and Keskin O (**2015**). Principles of K-Ras effector organization and the role of oncogenic K-Ras in cancer initiation through G1 cell cycle deregulation. **Expert Rev Proteomics** 12(6): 669–682. doi: 10.1586/14789450.2015.1100079

13. García-Gutiérrez L, Delgado MD, and León J (**2019**). MYC Oncogene Contributions to Release of Cell Cycle Brakes. **Genes** 10(3): 244. doi: 10.3390/genes10030244

14. Fagundes R, and Teixeira LK (**2021**). Cyclin E/CDK2: DNA Replication, Replication Stress and Genomic Instability. **Front Cell Dev Biol** 9: 774845. doi: 10.3389/fcell.2021.774845

15. Takahashi N, Kim S, Schultz CW, Rajapakse VN, Zhang Y, Redon CE, Fu H, Pongor L, Kumar S, Pommier Y, Aladjem MI, and Thomas A (**2022**). Replication Stress Defines Distinct Molecular Subtypes Across Cancers. **Cancer Res Commun** 2(6): 503–517. doi: 10.1158/2767-9764.CRC-22-0168

16. Guerrero Llobet S, Bhattacharya A, Everts M, Kok K, van der Vegt B, Fehrmann RSN, and van Vugt MATM (**2022**). An mRNA expression-based signature for oncogene-induced replication-stress. **Oncogene** 41(8): 1216–1224. doi: 10.1038/s41388-021-02162-0

17. Sheltzer JM (**2013**). A Transcriptional and Metabolic Signature of Primary Aneuploidy Is Present in Chromosomally Unstable Cancer Cells and Informs Clinical Prognosis. **Cancer Res** 73(21): 6401–6412. doi: 10.1158/0008-5472.CAN-13-0749

18. Zhang X, and Kschischo M (**2022**). Distinct and Common Features of Numerical and Structural Chromosomal Instability across Different Cancer Types. **Cancers** 14(6): 1424. doi: 10.3390/cancers14061424

19. Carter SL, Eklund AC, Kohane IS, Harris LN, and Szallasi Z (**2006**). A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. **Nat Genet** 38(9): 1043–1048. doi: 10.1038/ng1861

20. Habermann JK, Doering J, Hautaniemi S, Roblick UJ, Bündgen NK, Nicorici D, Kronenwett U, Rathnagiriswaran S, Mettu RKR, Ma Y, Krüger S, Bruch H-P, Auer G, Guo NL, and Ried T (**2009**). The gene expression signature of genomic instability in breast cancer is an independent predictor of clinical outcome. **Int J Cancer J Int Cancer** 124(7): 1552–1564. doi: 10.1002/ijc.24017

21. Rao CV, Yang Y-M, Swamy MV, Liu T, Fang Y, Mahmood R, Jhanwar-Uniyal M, and Dai W (**2005**). Colonic tumorigenesis in BubR1+/–ApcMin/+ compound mutant mice is linked to premature separation of sister chromatids and enhanced genomic instability. **Proc Natl Acad Sci U S A** 102(12): 4365–4370. doi: 10.1073/pnas.0407822102

22. Bester AC, Roniger M, Oren YS, Im MM, Sarni D, Chaoat M, Bensimon A, Zamir G, Shewach DS, and Kerem B (**2011**). Nucleotide Deficiency Promotes Genomic Instability in Early Stages of Cancer Development. **Cell** 145(3): 435–446. doi: 10.1016/j.cell.2011.03.044

23. Do BT, Hsu PP, Vermeulen SY, Wang Z, Hirz T, Abbott KL, Aziz N, Replogle JM, Bjelosevic S, Paolino J, Nelson SA, Block S, Darnell AM, Ferreira R, Zhang H, Milosevic J, Schmidt DR, Chidley C, Harris IS, Weissman JS, Pikman Y, Stegmaier K, Cheloufi S, Su XA, Sykes DB, and Vander Heiden MG (**2024**). Nucleotide depletion promotes cell fate transitions by inducing DNA replication stress. **Dev Cell** 59(16): 2203-2221.e15. doi: 10.1016/j.devcel.2024.05.010

24. McGrail DJ, Lin CC-J, Dai H, Mo W, Li Y, Stephan C, Davies P, Lu Z, Mills GB, Lee J-S, and Lin S-Y (**2018**). Defective Replication Stress Response Is Inherently Linked to the Cancer Stem Cell Phenotype. **Cell Rep** 23(7): 2095–2106. doi: 10.1016/j.celrep.2018.04.068

25. Di Micco R, Fumagalli M, Cicalese A, Piccinin S, Gasparini P, Luise C, Schurra C, Garre' M, Giovanni Nuciforo P, Bensimon A, Maestro R, Giuseppe Pelicci P, and d'Adda di Fagagna F (**2006**). Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication. **Nature** 444(7119): 638–642. doi: 10.1038/nature05327

26. Böhly N, Kistner M, and Bastians H (**2019**). Mild replication stress causes aneuploidy by deregulating microtubule dynamics in mitosis. **Cell Cycle** 18(20): 2770–2783. doi: 10.1080/15384101.2019.1658477

27. Bakhoum SF et al. (**2018**). Chromosomal instability drives metastasis through a cytosolic DNA response. **Nature** 553(7689): 467–472. doi: 10.1038/nature25432

28. Baker DJ, Jin F, Jeganathan KB, and van Deursen JM (**2009**). Whole Chromosome Instability Caused by Bub1 Insufficiency Drives Tumorigenesis through Tumor Suppressor Gene Loss of Heterozygosity. **Cancer Cell** 16(6): 475–486. doi: 10.1016/j.ccr.2009.10.023

29. Remeseiro S, Cuadrado A, Carretero M, Martínez P, Drosopoulos WC, Cañamero M, Schildkraut CL, Blasco MA, and Losada A (**2012**). Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres. **EMBO J** 31(9): 2076–2089. doi: 10.1038/emboj.2012.11

30. Lengauer C, Kinzler KW, and Vogelstein B (**1997**). Genetic instability in colorectal cancers. **Nature** 386(6625): 623–627. doi: 10.1038/386623a0

31. Gaillard H, García-Muse T, and Aguilera A (**2015**). Replication stress and cancer. **Nat Rev Cancer** 15(5): 276–289. doi: 10.1038/nrc3916

32. Wilhelm T, Said M, and Naim V (2020). DNA Replication Stress and Chromosomal Instability: Dangerous Liaisons. Genes 11(6): 642. doi: 10.3390/genes11060642

33. Limas JC, Littlejohn AN, House AM, Kedziora KM, Mouery BL, Ma B, Fleifel D, Walens A, Aleman MM, Dominguez D, and Cook JG (2022). Quantitative profiling of adaptation to cyclin E overproduction. Life Sci Alliance 5(5): e202201378. doi: 10.26508/lsa.202201378

34. Zimmerli D, Brambillasca CS, Talens F, Bhin J, Linstra R, Romanens L, Bhattacharya A, Joosten SEP, Da Silva AM, Padrao N, Wellenstein MD, Kersten K, de Boo M, Roorda M, Henneman L, de Bruijn R, Annunziato S, van der Burg E, Drenth AP, Lutz C, Endres T, van de Ven M, Eilers M, Wessels L, de Visser KE, Zwart W, Fehrmann RSN, van Vugt MATM, and Jonkers J (2022). MYC promotes immune-suppression in triple-negative breast cancer via inhibition of interferon signaling. Nat Commun 13(1): 6579. doi: 10.1038/s41467-022-34000-6

35. Lumeau A, Bery N, Francès A, Gayral M, Labrousse G, Ribeyre C, Lopez C, Nevot A, El Kaoutari A, Hanoun N, Sarot E, Perrier M, Pont F, Cerapio J-P, Fournié J-J, Lopez F, Madrid-Mencia M, Pancaldi V, Pillaire M-J, Bergoglio V, Torrisani J, Dusetti N, Hoffmann J-S, Buscail L, Lutzmann M, and Cordelier P (2024). Cytidine Deaminase Resolves Replicative Stress and Protects Pancreatic Cancer from DNA-Targeting Drugs. Cancer Res 84(7): 1013–1028. doi: 10.1158/0008-5472.CAN-22-3219

36. Wolf DM, Lenburg ME, Yau C, Boudreau A, and Veer LJ van 't (2014). Gene Co-Expression Modules as Clinically Relevant Hallmarks of Breast Cancer Diversity. PLOS ONE 9(2): e88309. doi: 10.1371/journal.pone.0088309

37. Taylor AM, Shih J, Ha G, Gao GF, Zhang X, Berger AC, Schumacher SE, Wang C, Hu H, Liu J, Lazar AJ, Cherniack AD, Beroukhim R, and Meyerson M (2018). Genomic and Functional Approaches to Understanding Cancer Aneuploidy. Cancer Cell 33(4): 676-689.e3. doi: 10.1016/j.ccell.2018.03.007

38. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, and Smyth GK (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43(7): e47. doi: 10.1093/nar/gkv007

39. Law CW, Chen Y, Shi W, and Smyth GK (2014). voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 15(2): R29. doi: 10.1186/gb-2014-15-2-r29

40. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, and Mesirov JP (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci 102(43): 15545–15550. doi: 10.1073/pnas.0506580102

41. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, and Tamayo P (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 1(6): 417–425. doi: 10.1016/j.cels.2015.12.004

42. Coleman ML, Marshall CJ, and Olson MF (2004). RAS and RHO GTPases in G1-phase cell-cycle regulation. Nat Rev Mol Cell Biol 5(5): 355–366. doi: 10.1038/nrm1365

43. Lavin MF, and Gueven N (2006). The complexity of p53 stabilization and activation. Cell Death Differ 13(6): 941–950. doi: 10.1038/sj.cdd.4401925

44. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, and Sherlock G (2000). Gene Ontology: tool for the unification of biology. Nat Genet 25(1): 25–29. doi: 10.1038/75556

45. The Gene Ontology Consortium et al. (2023). The Gene Ontology knowledgebase in 2023. Genetics 224(1): iyad031. doi: 10.1093/genetics/iyad031

46. Lodovichi S, Cervelli T, Pellicioli A, and Galli A (2020). Inhibition of DNA Repair in Cancer Therapy: Toward a Multi-Target Approach. Int J Mol Sci 21(18): 6684. doi: 10.3390/ijms21186684

47. da Costa AABA, Chowdhury D, Shapiro GI, D'Andrea AD, and Konstantinopoulos PA (2023). Targeting replication stress in cancer therapy. Nat Rev Drug Discov 22(1): 38–58. doi: 10.1038/s41573-022-00558-5

48. Basu A, Bodycombe NE, Cheah JH, Price EV, Liu K, Schaefer GI, Ebright RY, Stewart ML, Ito D, Wang S, Bracha AL, Liefeld T, Wawer M, Gilbert JC, Wilson AJ, Stransky N, Kryukov GV, Dancik V, Barretina J, Garraway LA, Hon CS-Y, Munoz B, Bittker JA, Stockwell BR, Khabele D, Stern AM, Clemons PA, Shamji AF, and Schreiber SL (2013). An Interactive Resource to Identify Cancer Genetic and Lineage Dependencies Targeted by Small Molecules. Cell 154(5): 1151–1161. doi: 10.1016/j.cell.2013.08.003

49. Seashore-Ludlow B, Rees MG, Cheah JH, Cokol M, Price EV, Coletti ME, Jones V, Bodycombe NE, Soule CK, Gould J, Alexander B, Li A, Montgomery P, Wawer MJ, Kuru N, Kotz JD, Hon CS-Y, Munoz B, Liefeld T, Dančík V, Bittker JA, Palmer M, Bradner JE, Shamji AF, Clemons PA, and Schreiber SL (2015). Harnessing Connectivity in a Large-Scale Small-Molecule Sensitivity Dataset. Cancer Discov 5(11): 1210–1223. doi: 10.1158/2159-8290.CD-15-0235

50. Rees MG, Seashore-Ludlow B, Cheah JH, Adams DJ, Price EV, Gill S, Javaid S, Coletti ME, Jones VL, Bodycombe NE, Soule CK, Alexander B, Li A, Montgomery P, Kotz JD, Hon CS-Y, Munoz B, Liefeld T, Dančík V, Haber DA, Clish CB, Bittker JA, Palmer M, Wagner BK, Clemons PA, Shamji AF, and Schreiber SL (2016). Correlating chemical sensitivity and basal gene expression reveals mechanism of action. Nat Chem Biol 12(2): 109–116. doi: 10.1038/nchembio.1986

51. Coban I, Lamping J-P, Hirsch AG, Wasilewski S, Shomroni O, Giesbrecht O, Salinas G, and Krebber H (2024). dsRNA formation leads to preferential nuclear export and gene expression. Nature 631(8020): 432–438. doi: 10.1038/s41586-024-07576-w

52. Yang Q, Zhang R, Wang XW, Linke SP, Sengupta S, Hickson ID, Pedrazzi G, Perrera C, Stagljar I, Littman SJ, Modrich P, and Harris CC (2004). The mismatch DNA repair heterodimer, hMSH2/6, regulates BLM helicase. Oncogene 23(21): 3749–3756. doi: 10.1038/sj.onc.1207462

53. Smith JA, Bannister LA, Bhattacharjee V, Wang Y, Waldman BC, and Waldman AS (2007). Accurate Homologous Recombination Is a Prominent Double-Strand Break Repair Pathway in Mammalian Chromosomes and Is Modulated by Mismatch Repair Protein Msh2. Mol Cell Biol 27(22): 7816. doi: 10.1128/MCB.00455-07

54. Shahi A, Lee J-H, Kang Y, Lee SH, Hyun J-W, Chang I-Y, Jun J-Y, and You HJ (2011). Mismatch-repair protein MSH6 is associated with Ku70 and regulates DNA double-strand break repair. Nucleic Acids Res.39(6): 2130–2143. doi: 10.1093/nar/gkq1095

55. Barbie DA et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 462(7269): 108–112. doi: 10.1038/nature08460

56. Kanehisa M, Goto S, Furumichi M, Tanabe M, and Hirakawa M (2010). KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res 38(suppl_1): D355–D360. doi: 10.1093/nar/gkp896

57. Knijnenburg TA et al. (2018). Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. Cell Rep 23(1): 239-254.e6. doi: 10.1016/j.celrep.2018.03.076

58. Terasawa M, Shinohara A, and Shinohara M (2014). Canonical Non-Homologous End Joining in Mitosis Induces Genome Instability and Is Suppressed by M-phase-Specific Phosphorylation of XRCC4. PLOS Genet 10(8): e1004563. doi: 10.1371/journal.pgen.1004563.

59. Saxena S, and Zou L (2022). Hallmarks of DNA replication stress. Mol Cell 82(12): 2298–2314. doi: 10.1016/j.molcel.2022.05.004

60. Hintzen DC, Schubert M, Soto M, Medema RH, and Raaijmakers JA (2024). Reduction of chromosomal instability and inflammation is a common aspect of adaptation to aneuploidy. EMBO Rep 25(11): 5169-5193. doi: 10.1038/s44319-024-00252-0

61. Bökenkamp J-E, Keuper K, Redel S, Barthel K, Johnson L, Becker A, Wieland A, Räschle M, and Storchová Z (2025). Proteogenomic analysis reveals adaptive strategies for alleviating the consequences of aneuploidy in cancer. EMBO J 44(6):1829-1865. doi: 10.1038/s44318-025-00372-w

62. Shen R, Liu D, Wang X, Guo Z, Sun H, Song Y, and Wang D (2022). DNA Damage and Activation of cGAS/STING Pathway Induce Tumor Microenvironment Remodeling. Front Cell Dev Biol 9: 828657. doi: 10.3389/fcell.2021.828657

63. Kerneur C, Cano CE, and Olive D (2022). Major pathways involved in macrophage polarization in cancer. Front Immunol.13: 1026954. doi: 10.3389/fimmu.2022.1026954

64. Li J, Hubisz MJ, Earlie EM, Duran MA, Hong C, Varela AA, Lettera E, Deyell M, Tavora B, Havel JJ, Phyu SM, Amin AD, Budre K, Kamiya E, Cavallo J-A, Garris C, Powell S, Reis-Filho JS, Wen H, Bettigole S, Khan AJ, Izar B, Parkes EE, Laughney AM, and Bakhoum SF (2023). Non-cell-autonomous cancer

progression from chromosomal instability. **Nature** 620(7976): 1080–1088. doi: 10.1038/s41586-023-06464-z

65. Martinez-Outschoorn UE, Balliet RM, Rivadeneira DB, Chiavarina B, Pavlides S, Wang C, Whitaker-Menezes D, Daumer KM, Lin Z, Witkiewicz AK, Flomenberg N, Howell A, Pestell RG, Knudsen ES, Sotgia F, and Lisanti MP (**2010**). Oxidative stress in cancer associated fibroblasts drives tumor-stroma co-evolution: A new paradigm for understanding tumor metabolism, the field effect and genomic instability in cancer cells. **Cell Cycle** 9(16): 3256. doi: 10.4161/cc.9.16.12553

66. Scully R, Panday A, Elango R, and Willis NA (**2019**). DNA double strand break repair pathway choice in somatic mammalian cells. **Nat Rev Mol Cell Biol** 20(11): 698–714. doi: 10.1038/s41580-019-0152-0

67. Daboussi F, Courbet S, Benhamou S, Kannouche P, Zdzienicka MZ, Debatisse M, and Lopez BS (**2008**). A homologous recombination defect affects replication-fork progression in mammalian cells. **J Cell Sci** 121(2): 162–166. doi: 10.1242/jcs.010330

68. Wilhelm T, Magdalou I, Barascu A, Técher H, Debatisse M, and Lopez BS (**2014**). Spontaneous slow replication fork progression elicits mitosis alterations in homologous recombination-deficient mammalian cells. **Proc Natl Acad Sci** 111(2): 763–768. doi: 10.1073/pnas.1311520111

69. Molano-Fernández M, Hickson ID, and Herranz H (**2025**). Replication stress promotes cellular transformation in Drosophila epithelium. **Cell Death Discov** 11(1): 96. doi: 10.1038/s41420-025-02383-2

70. Ma Y, Mu X, Gao R, Zhang Y, Geng Y, Chen X, Yin X, Li F, and He J (**2023**). Maternal exposure to dibutyl phthalate regulates MSH6 crotonylation to impair homologous recombination in fetal oocytes. **J Hazard Mater** 455: 131540. doi: 10.1016/j.jhazmat.2023.131540

71. Zhang J, Zhao X, Liu L, Li H-D, Gu L, Castrillon DH, and Li G-M (**2022**). The mismatch recognition protein MutSα promotes nascent strand degradation at stalled replication forks. **Proc Natl Acad Sci** 119(40): e2201738119. doi: 10.1073/pnas.2201738119

72. Wilczak W, Rashed S, Hube-Magg C, Kluth M, Simon R, Büscheck F, Clauditz TS, Grupp K, Minner S, Tsourlakis MC, Möller-Koop C, Graefen M, Adam M, Haese A, Wittmer C, Sauter G, Izbicki JR, Huland H, Schlomm T, Steurer S, Krech T, and Lebok P (**2017**). Up-regulation of mismatch repair genes MSH6, PMS2 and MLH1 parallels development of genetic instability and is linked to tumor aggressiveness and early PSA recurrence in prostate cancer. **Carcinogenesis** 38(1): 19–27. doi: 10.1093/carcin/bgw116.

73. Love MI, Huber W, and Anders S (**2014**). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. **Genome Biol** 15(12): 550. doi: 10.1186/s13059-014-0550-8

74. Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, and Stuart JM (**2013**). The Cancer Genome Atlas Pan-Cancer analysis project. **Nat Genet** 45(10): 1113–1120. doi: 10.1038/ng.2764

75. Li Y et al. (**2023**). Proteogenomic data and resources for pan-cancer analysis. **Cancer Cell** 41(8): 1397–1406. doi: 10.1016/j.ccell.2023.06.009

76. Barretina J et al. (**2012**). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. **Nature** 483(7391): 603–607. doi: 10.1038/nature11003

77. Thorsson V et al. (**2018**). The Immune Landscape of Cancer. **Immunity** 48(4): 812-830.e14. doi: 10.1016/j.immuni.2018.03.023

78. Sanchez-Vega F et al. (**2018**). Oncogenic Signaling Pathways in The Cancer Genome Atlas. **Cell** 173(2): 321-337.e10. doi: 10.1016/j.cell.2018.03.035

79. Siri SO, Martino J, and Gottifredi V (**2021**). Structural Chromosome Instability: Types, Origins, Consequences, and Therapeutic Opportunities. **Cancers** 13(12): 3056. doi: 10.3390/cancers13123056

80. Kim W, Zhao F, Gao H, Qin S, Hou J, Deng M, Kloeber JA, Huang J, Zhou Q, Guo G, Gao M, Zeng X, Zhu S, Tu X, Wu Z, Zhang Y, Yin P, Kaufmann SH, Luo K, and Lou Z (**2021**). USP13 regulates the replication stress response by deubiquitinating TopBP1. **DNA Repair** 100: 103063. doi: 10.1016/j.dnarep.2021.103063

81. Hodroj D, Recolin B, Serhal K, Martinez S, Tsanov N, Abou Merhi R, and Maiorano D (**2017**). An ATR-dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism. **EMBO J** 36(9): 1182–1198. doi: 10.15252/embj.201695131

82. Flanagan A-M, Stavenschi E, Basavaraju S, Gaboriau D, Hoey DA, and Morrison CG (**2017**). Centriole splitting caused by loss of the centrosomal linker protein C-NAP1 reducesncentriolar satellite density and impedes centrosome amplification. **Mol Biol Cell** 28(6): 736–745. doi: 10.1091/mbc.E16-05-0325

83. Shi J, Hauschulte K, Mikicic I, Maharjan S, Arz V, Strauch T, Heidelberger JB, Schaefer JV, Dreier B, Plückthun A, Beli P, Ulrich HD, and Wollscheid H-P (**2023**). Nuclear myosin VI maintains replication fork stability. **Nat Commun** 14: 3787. doi: 10.1038/s41467-023-39517-y

84. Olivieri M, Cho T, Álvarez-Quilón A, Li K, Schellenberg MJ, Zimmermann M, Hustedt N, Rossi SE, Adam S, Melo H, Heijink AM, Sastre-Moreno G, Moatti N, Szilard RK, McEwan A, Ling AK, Serrano-Benitez A, Ubhi T, Feng S, Pawling J, Delgado-Sainz I, Ferguson MW, Dennis JW, Brown GW, Cortés-Ledesma F, Williams RS, Martin A, Xu D, and Durocher D (**2020**). A genetic map of the response to DNA damage in human cells. **Cell** 182(2): 481-496.e21. doi: 10.1016/j.cell.2020.05.040

85. Pan X, Zhao J, Zhang W-N, Li H-Y, Mu R, Zhou T, Zhang H-Y, Gong W-L, Yu M, Man J-H, Zhang P-J, Li A-L, and Zhang X-M (**2009**). Induction of SOX4 by DNA damage is critical for p53 stabilization and function. **Proc Natl Acad Sci U S A** 106(10): 3788–3793. doi: 10.1073/pnas.0810147106

86. Hur W, Rhim H, Jung CK, Kim JD, Bae SH, Jang JW, Yang JM, Oh S-T, Kim DG, Wang HJ, Lee SB, and Yoon SK (**2010**). SOX4 overexpression regulates the p53-mediated apoptosis in hepatocellular carcinoma: clinical implication and functional analysis in vitro. **Carcinogenesis** 31(7): 1298–1307. doi: 10.1093/carcin/bgq072

87. Perks CM, Keith AJ, Goodhew KL, Savage PB, Winters ZE, and Holly JMP (**2004**). Prolactin acts as a potent survival factor for human breast cancer cell lines. **Br J Cancer** 91(2): 305–311. doi: 10.1038/sj.bjc.6601947

88. O'Leary KA, Rugowski DE, Sullivan R, and Schuler LA (**2014**). Prolactin cooperates with loss of p53 to promote claudin low mammary carcinomas. **Oncogene** 33(23): 3075–3082. doi: 10.1038/onc.2013.278